Stochastic homogenization for functionals with anisotropic rescaling and non-coercive Hamilton-Jacobi equations.

Nicolas Dirr* Federica Dragoni † Paola Mannucci ‡ Claudio Marchi §

Abstract

We study the stochastic homogenization for a Cauchy problem for a first-order Hamilton-Jacobi equation whose operator is not coercive w.r.t. the gradient variable. We look at Hamiltonians like $H(x, \sigma(x)p, \omega)$ where $\sigma(x)$ is a matrix associated to a Carnot group. The rescaling considered is consistent with the underlying Carnot group structure, thus anisotropic. We will prove that under suitable assumptions for the Hamiltonian, the solutions of the ε-problem converge to a deterministic function which can be characterized as the unique (viscosity) solution of a suitable deterministic Hamilton-Jacobi problem.

Keywords: Stochastic homogenization, non-coercive Hamilton-Jacobi equations, Carnot groups, Hörmander condition, Heisenberg group, anisotropic functionals.

1 Introduction

Homogenization problems have been studied for many years for both their intrinsic mathematical interest and the many applications in different sciences (e.g. the study of heterogeneous media). In particular stochastic homogenization arises whenever at the microscopic level the system depends on some random variable but at the macroscopic level one can expect a deterministic behaviour.

In this paper, we study asymptotics of a special class of *degenerate* (i.e. non-coercive) first-order Hamilton-Jacobi equations with random coefficients taking the form

$$\begin{cases} u_t + H\left(x, \sigma(x)\nabla u, \omega\right) = 0, \ t > 0, \ x \in \mathbb{R}^N, \ \omega \in \Omega, \\ u(0, x, \omega) = g(x), \ x \in \mathbb{R}^N, \ \omega \in \Omega, \end{cases} \tag{1.1}$$

*Cardiff School of Mathematics, Cardiff University, Cardiff, UK, e-mail: dirrnp@cardiff.ac.uk.

†Cardiff School of Mathematics, Cardiff University, Cardiff, UK, e-mail: DragoniF@cardiff.ac.uk.

‡Dipartimento di Matematica "Tullio Levi-Civita" , Università di Padova, Padova, Italy, e-mail: mannucci@math.unipd.it

§Dipartimento di Ingegneria dell'Informazione, Università di Padova, Padova, Italy, e-mail: claudio.marchi@unipd.it

where $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space, and $\sigma : \mathbb{R}^N \to \mathbb{R}^{m\times N}$ with $m \leq N$. Even though $H(x, q)$ is coercive and convex in the variable $q = \sigma(x)p \in \mathbb{R}^m$, the map $p \mapsto H(x, \sigma(x)p, \omega)$ is in general not coercive because $\sigma(x)$ may have a nontrivial kernel. The illustrating example is the Heisenberg group, which is topologically $\mathbb{R}^3$ but with a different algebraic structure (see Section 2, e.g. (2.2)).

Equations as (1.1) can be understood in the framework of Carnot groups, i.e. non-commutative stratified nilpotent Lie groups (see Section 2 for more details). In particular these groups satisfy the Hörmander condition: they are endowed with a family of vector fields that, together with all their associated commutators, span the whole tangent space at any point of the original manifold. For the associated homogenization problem, the Carnot group structure suggests a natural anisotropic rescaling of $\mathbb{R}^N$, denoted by $\delta_{1/\varepsilon}(x)$ for $x \in \mathbb{R}^N$. Then the homogenization problem can be formulated as follows: under some assumptions made precise later (see Section 3), find the equation solved by the (locally uniform) limit of $u^\varepsilon(t, x, \omega)$ where u^ε are viscosity solutions of

$$\begin{cases} u^\varepsilon_t + H\left(\delta_{1/\varepsilon}(x), \sigma(x)\nabla u^\varepsilon, \omega\right) = 0, \ t > 0, \ x \in \mathbb{R}^N, \ \omega \in \Omega, \\ u^\varepsilon(0, x, \omega) = g(x), \ x \in \mathbb{R}^N, \ \omega \in \Omega. \end{cases} \tag{1.2}$$

In other words, the aim is to identify $\overline{H} : \mathbb{R}^m \to \mathbb{R}$ such that the viscosity solutions of (1.2) converge, locally uniformly in t and x and almost surely in ω, to a deterministic function $u(t, x)$ which can be characterized as the unique viscosity solution of a problem of the form

$$\begin{cases} u_t + \overline{H}\left(\sigma(x)\nabla u\right) = 0, \ t > 0, \ x \in \mathbb{R}^N, \\ u(0, x) = g(x), \ x \in \mathbb{R}^N. \end{cases} \tag{1.3}$$

In the case of the Heisenberg group, the anisotropic rescaling is $\delta_{1/\varepsilon}(x_1, x_2, x_3) = (\varepsilon^{-1}x_1, \varepsilon^{-1}x_2, \varepsilon^{-2}x_3)$. This is consistent with the geometric structure of the Heisenberg group, but the anisotropy can be understood heuristically in another way: at each point, some directions are "forbidden", i.e. paths of the associated control problem can move only on a two-dimensional subspace. By varying their direction often (i.e. by the use of non-trivial commutators from the Hörmander condition) they are able to reach any given point but the cost for "zig-zagging" to get in the forbidden direction is higher, so typically they move slower in these directions, which makes a faster rescaling necessary.

Note that, in (1.2), $\sigma(x)$ is not rescaled so this is in principle a problem with a fast and a slow variable, but the equation is degenerate if the slow variables are frozen. Obviously, general non-coercive equations have no homogenization, so considering a cell problem with a frozen variable is not the way to tackle this problem.

Instead, our approach is based on the use of a variational formulation for the viscosity solutions of (1.2), that has been introduced in the coercive case by Souganidis [37] and Rezakhanlou-Tarver [36]. This variational approach is motivated by Γ-convergence methods for the random Lagrangian. In order to define the associated variational problem from the Hamilton-Jacobi equation, some form of convexity is needed, but it should be noted that due to the degeneracy the relation is more subtle than the Euclidean Legendre transform, see [8].

Moreover the approach developed in [36, 37] fails since the idea of using the Subadditive Ergodic Theorem indirectly requires the existence of curves invariant under translation and rescaling (as straight lines are w.r.t. the Euclidean translations). In our anisotropic geometries, this property is true only for curves that have constant horizontal speed (i.e. velocity constant w.r.t. a given family of left-invariant vector fields, see Section 2 for more details). Unfortunately those lines are too few to cover the whole space (they only generate a m-dimensional submanifold in $\mathbb{R}^N$ where usually $N > m$).

We would like to mention that non-coercive Hamilton-Jacobi problems have also been studied by Ciomaga, Souganidis and Tran [19] and Cardaliaguet and Souganidis [18]. Note that in these cases the source of non-coercivity is different: in our case, we have at each point some deterministic directions in which the Hamiltonian grows superlinearly, and some directions in which it does not grow at all, not even linearly. Hence the techniques are necessarily quite different.

It is an interesting topic for further research whether PDE techniques like the approximate correctors as in [31] or the metric problem as e.g. in [6] could be of use in the degenerate setting at hand. While the notion of "path", which is at the heart of the variational approach, has an obvious analogue in the sub-Riemannian case, many standard viscosity techniques are problematic in the degenerate case, leading to technical difficulties. These technical difficulties can often be traced back to two sources. First, the vector fields may have in general coefficients with superlinear growth at infinity, thus making global estimates difficult to obtain. Second, while in the Euclidean setting $|x - y|^2$ is smooth, this is in general not true for the square of the Carnot-Caratheodory distance, which has consequences for regularization via inf/sup convolution or doubling of variables techniques.

The main idea of the proof for the convergence theorem is to apply the techniques from [36, 37] (for the periodic case see also [24]) to a lower dimensional constrained variational problem (Section 5), the constraint being to belong to the m-dimensional manifold mentioned above. Then by an approximating argument (Section 6) we write the original variational problem (3.4) as limit of sum of lower dimensional constrained variational problems. The key role in the whole argument will be to approximate any horizontal curve by a suitable family of piecewise horizontal lines with constant speed and the use of the Hörmander condition to move everywhere in the space.
Here our a priori bounds on the Lagrangian ensure that the cost of connecting any two points can be bounded by a function of the geodesic distance. This allows to estimate the difference in cost for connecting nearby points, a property which makes up for the lack of uniform continuity of the Lagrangian due to the rescaling in space.

This is to our knowledge the first paper which connects two previously separate branches of homogenization theory: Stochastic homogenization on the one hand, which so far has not been considered in sub-Riemannian geometries, and homogenization in the sub-elliptic setting, which so far has been restricted to a suitable generalization of periodic environments, i.e. essentially in a compact setting. For homogenization in subelliptic settings in the periodic case see for example [13, 15, 27, 28, 34, 38], and for homogenization with singular perturbation see [2, 3].

Since the first results on stochastic homogenization for first-order Hamilton-Jacobi equations ([37, 36]), it has been a difficult question which are the neces-

sary conditions on the deterministic structure of the Hamiltonian, with convexity and coercivity being sufficient. The case of non-convexity has been understood better recently, see e.g. [26, 39]. Instead this paper gives a very general class of examples which are convex (but not strictly) but non-coercive. This homogenization result is in line with the folk theorem that, in order to have homogenization, characteristics have to be able to go everywhere: our degeneracy is related to Hörmander vector fields, which have the property that admissible paths (see Section 2) can connect any two given points.

Γ-convergence for random functionals, which is used here, has in a general setting first been studied by Dal Maso and Modica, [20] and recently been extended to non-convex integrands, [23]. Alternatives to the variational approach for obtaining stochastic homogenization results in the Euclidean setting for both first and second order equations and the simultaneous effect of homogenization and vanishing viscosity (i.e. singular perturbation) have been developed subsequently, see for example [30, 32, 31]. Extending these methods to the sub-Riemannian setting will be a challenge for further research.

This paper is organized as follows.

In Section 2 we introduce some basic notions for Carnot groups, in particular the dilations in the group and some norms and distances related to both the geometric and the algebraic structure of Carnot groups. In this section we also introduce horizontal curves, horizontal velocity and study some properties, which will be very useful in later proofs.

In Section 3 we state the problem and we explain the meaning of some assumptions on the Hamiltonian; in particular the stationary ergodic assumption which is crucial in order to get a deterministic limit problem. In the same section we also introduce the variational formulation for the solutions of the ε-problem.

In Section 4 we study several properties for the variational problem. In particular we prove local uniform continuity.

In Section 5 we prove the convergence for the constrained variational problem, i.e. for the minimizing problem for an integral cost under the additional m-dimensional constraint.

In Section 6 we prove our main convergence result for the unconstrained variational problem by the introduction of a suitable approximation argument.

In Section 7 we apply the convergence proved in Section 6 to the family of non-coercive Cauchy-Hamilton-Jacobi problems (1.2) via variational formula.

In the Appendix (Section 8) we give a proof for the well-posedness of the ε-problem (1.2) in the viscosity sense.

2 Preliminaries: Carnot groups.

Carnot groups are non-commutative Lie groups: they are endowed both with a non-commutative algebraic structure and with a manifold structure. The lack of commutativity in the algebraic structure reflects on the manifold structure as restrictions on the admissible motions. This means that the allowed curves are constrained to have their velocities in a lower dimensional subspace of the tangent space of the manifold. Then the associated manifold structure is not Riemannian but sub-Riemannian. We refer the reader to [16] for an overview on Carnot groups and sub-Riemannian manifolds. Here we only recall the definitions and some of the main properties, which will be crucial in the later

proofs.

Definition 2.1 (Carnot group)**.** *A Carnot group $(\mathbb{G}, \circ)$ of step r is a simply connected, nilpotent Lie group whose Lie algebra g of left-invariant vector fields admits a stratification, i.e. there exist non zero subspaces $\{V_i\}$, $i = 1, \dots r$ such that $g = \bigoplus_{i=1}^{r} V_i$, $[V_1, V_i] = V_{i+1} \neq 0$, for $i = 1, \dots r-1$, $[V_1, V_r] = 0$. V_1 is called the first layer.*
Any such group is isomorphic to a homogeneous Carnot group in $\mathbb{R}^N$, that is a triple $(\mathbb{R}^N, \circ, \delta_\lambda)$ where $\mathbb{R}^N = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_k}$, $\circ$ is a group operation whose identity is e *and such that $(x, y) \to y^{-1} \circ x$ is smooth (where y^{-1} denote the inverse of y), and $\delta_\lambda : \mathbb{R}^N \to \mathbb{R}^N$ is the dilation:*

$$\delta_\lambda(x) = \delta_\lambda\left(x^{(1)}, x^{(2)}, \cdots, x^{(r)}\right) := \left(\lambda\, x^{(1)}, \lambda^2\, x^{(2)}, \cdots, \lambda^r x^{(r)}\right),\ x^{(i)} \in \mathbb{R}^{n_i}, \tag{2.1}$$

is an automorphism of the group $(\mathbb{R}^N, \circ)$ for all $\lambda > 0$ and there are $m := n_1$ smooth vector fields $X_1, \cdots, X_m$ on $\mathbb{R}^N$ invariant with respect to the left translation

$$L_\beta(x) := \beta \circ x$$

for all $\beta \in \mathbb{R}^N$ and such that they generate a Lie algebra with rank N at every point $x \in \mathbb{R}^N$. The vector fields $X_1, \cdots, X_m$ are called the generators of the Carnot group or horizontal vector fields and the $n \times m$ matrix whose columns are these vector fields is denoted by σ.

For $x \in \mathbb{R}^N$, we shall also use the notation: $x = (x^1, x^2)$ with $x^1 \in \mathbb{R}^m$, $x^2 \in \mathbb{R}^{N-m}$ and $x^1 := \pi_m(x)$.

The definition of dilations (that replace the role of product of a point by a scalar in the Euclidean case) gives good notions of rescaling in these geometries.

Note that we are interested only in the case where $\mathbb{G} = \mathbb{R}^N$ for some $N \geq 3$ (in fact Carnot groups with dimension less than 3 do not exist).

Example 2.1. *The simplest example of a Carnot group is the so called Heisenberg group. The N-dimensional Heisenberg group $\mathbb{H}^N$ is a Carnot group of step 2 (i.e. $r = 2$ in the stratification) defined in $\mathbb{R}^{2N+1}$ (with $N \geq 1$). In particular if $N = 1$ the stratification is $V_1 \bigoplus V_2$, where $V_1 = \mathbb{R}^2$ and $V_2 = \mathbb{R}$. In this last case the group operation is*

$$x \circ y := \left(x_1 + y_1, x_2 + y_2, x_3 + y_3 + \frac{x_1 y_2 - x_2 y_1}{2}\right)$$

where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ are two points in $\mathbb{R}^3$ and the generators are the two vector fields

$$X_1(x) = \begin{pmatrix} 1 \\ 0 \\ \frac{x_2}{2} \end{pmatrix},\ X_2(x) = \begin{pmatrix} 0 \\ 1 \\ \frac{-x_1}{2} \end{pmatrix}$$

In the Heisenberg group $\mathbb{H}^1$ the dilations that give the natural rescaling are

$$\delta_\lambda(x) = \delta_\lambda(x_1, x_2, x_3) = (\lambda\, x_1, \lambda\, x_2, \lambda^2\, x_3).$$

To make the paper more easily readable for mathematicians not used to work in Carnot groups we will explain most of the notions and properties of Carnot groups, using the 1-dimensional Heisenberg group $\mathbb{H} = \mathbb{H}^1$ as referring model.

Another family of algebraic objects which will play a crucial role in our homogenization problem are the translations. Since the group law is not commutative, in general left translations and right translations will be different. We will always translate points using only the left translations.
Using the stratification, a Carnot group can be endowed with a homogeneous norm that induces a homogeneous distance. The homogeneous norm and the homogeneous distance are very important in homogenization problems since they are compatible with rescaling under dilations (as we will see in the properties below).

Definition 2.2 (Homogeneous norm and homogeneous distance)**.** *A homogeneous norm $\|\cdot\|_h$ is a continuous function from $\mathbb{G}$ to $[0,+\infty)$ such that*

1. $\|x\|_h = 0 \iff x = \mathrm{e}$
2. $\|x^{-1}\|_h = \|x\|_h$
3. $\|\delta_\lambda(x)\|_h = \lambda\|x\|_h, \forall x \in \mathbb{G}, \lambda > 0$
4. $\|x+y\|_h \le \|x\|_h + \|y\|_h, \forall x, y \in \mathbb{G}$.

The homogeneous distance between two points $x, y \in \mathbb{G}$ is

$$d_h(x,y) = \|y^{-1} \circ x\|_h.$$

From $\|x\|_h = \|x^{-1}\|_h$ we have that $d_h(x,y) = d_h(y,x)$ and obviously $d_h(x,x) = 0$ for all $x, y \in \mathbb{G}$.
Moreover, given two points $x, y \in \mathbb{G} \equiv \mathbb{R}^N$, $\big(\delta_\lambda(x)\big)^{-1} = \delta_\lambda(x^{-1})$ and $\delta_\lambda(x) \circ \delta_\lambda(y) = \delta_\lambda(x \circ y)$. This implies that

$$d_h\big(\delta_\lambda(x), \delta_\lambda(y)\big) = \lambda\, d_h(x,y).$$

In the case of the 1-dimensional Heisenberg group $\mathbb{H}$ we have

$$\|x\|_h = \|(x_1,x_2,x_3)\|_h = \big((x_1^2+x_2^2)^2 + x_3^2\big)^{1/4}.$$

Moreover it is easy to check that $\mathrm{e} = (0,0,0)$ and $x^{-1} = (-x_1,-x_2,-x_3)$ so

$$d_h(x,y) = \big(((x_1-y_1)^2 + (x_2-y_2)^2)^2 + (x_3-y_3)^2\big)^{1/4}.$$

One can easily check all the properties listed above in the case of the 1-dimensional Heisenberg group.

For later use, it is very useful to introduce the $m \times n$ matrix associated to the vector fields

$$\sigma(x) := \big(X_1(x), \dots, X_m(x)\big)^T$$

e.g. in $\mathbb{H}^1$ the matrix $\sigma(\cdot)$ is the 2×3-matrix given by

$$\sigma(x_1,x_2,x_3) = \begin{pmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{pmatrix}. \tag{2.2}$$

From now on we will always consider the Carnot groups, written in exponential coordinates (or canonical coordinates). In fact in exponential coordinates the vector fields (and so the associated matrix $\sigma(x)$) assume a special form, as shown in the following lemma.

Lemma 2.1. *Given a Carnot group in exponential (or canonical) coordinates, then the vector fields can be considered as the columns of a $m \times N$ matrix $\sigma(x)$ of this form*

$$\sigma(x) = \begin{pmatrix} Id_{m\times m} & A(x) \end{pmatrix} \tag{2.3}$$

where $Id_{m\times m}$ is the identity matrix $m \times m$ and $A(x)$ is a $m \times (N-m)$ matrix whose coefficients are smooth functions depending only on $x_1, \dots, x_m$.
Moreover the non-vanishing coefficients of $A(x) = (a_{j,i}(x))$ with $i = 1, \dots, N-m$ and $j = 1, \dots, m$ are polynomial functions of degree $k-1$ whenever the $(m+i)$-th component rescale as λ^k in the dilations δ_λ defined in (2.1).

For a proof we refer the reader to [16]; in particular see [16, Proposition 1.3.5, Corollary 1.3.19] for the polynomial structure and the corresponding homogeneity degree. Remember that δ_λ-homogeneity corresponds to Euclidean homogeneity whenever the functions depend only on the first m components.

The previous lemma is easy to check in the 1-dimensional Heisenberg group (see (2.2)), in fact $a_{1,1}(x_1, x_2) = -\frac{x_2}{2}$ and $a_{2,1}(x_1, x_2) = \frac{x_1}{2}$ are both polynomials of degree 2-1=1. We now give another example for a step 3 Carnot group.

Example 2.2 (Engel group in exponential coordinates)**.** *The Engel group is Carnot group of step 3 defined on $\mathbb{R}^4$. It can be written as extension of the Heisenberg group but for us it is crucial to write it in exponential coordinates (see e.g. [12]). The rescaling in the Engel group is given by*

$$\delta_\lambda(x_1, x_2, x_3, x_4) = \big(\lambda x_1, \lambda x_2, \lambda^2 x_3, \lambda^3 x_4\big).$$

In exponential coordinate the vector fields generating V_1 can be written as

$$X_1(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_1} \text{ and } X_2(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3} + \frac{x_1^2}{2} \frac{\partial}{\partial x_4}.$$

In this case the corresponding 2×4-matrix has the form of a 2×2-identity matrix and a 2×2 matrix $A(x)$ whose coefficients are $a_{1,1}(x) = 0 = a_{1,2}(x)$, while $a_{2,1}(x) = x_1$ which is a polynomial of degree 1 (in fact the component $2+1=3$ rescales with $k=2$), and $a_{2,2}(x) = \frac{x_1^2}{2}$ which is a polynomial of degree 2 (in fact the component $2+2=4$ rescales with $k=3$). Then Lemma 2.1 is easily verified.

So far, we have briefly recalled the algebraic structure of Carnot groups. Since Carnot groups are also sub-Riemannian manifolds there is also another important distance to consider: the so called Carnot-Carathéodory distance. Before defining the Carnot-Carathéodory distance and its relations with the homogeneous distance and the Euclidean distance, we need to introduce the sub-Riemannian manifold structure associated to a Carnot group. Consider the left-invariant vector fields $X_1, \dots, X_m$ introduced above on $\mathbb{R}^N$, then by identifying the tangent space at the origin with the Lie algebra g (see Definition 2.1) and in any other point by left-translation, then $X_1, \dots, X_m$ satisfy the

Hörmander condition with step r. We remind that the Hörmander condition states that the Lie algebra induced by the vector fields has to be at any point equal to the whole tangent space at that point.
Denoted by $\mathcal{H}_x = \mathrm{Span}\big(X_1, \dots, X_m\big)$ the distribution spanned by the given left-invariant vector fields, then it is possible to define a Riemannian metric on $\mathcal{H}_x$ induced by the vector fields, by taking $< v, w >= \alpha \cdot \beta$ where α and β are m-valued vectors, corresponding to the coordinates of v and w respectively, w.r.t. the given vector fields.
The triple $\big(\mathbb{R}^N, \mathcal{H}_x, < \cdot, \cdot > \big)$ is a sub-Riemannian manifold. For more details on sub-Riemannian manifolds in general and the manifold structure associated to Carnot groups in particular, we refer respectively to [35] and [16].
Next we recall the notion of horizontal (or admissible) curve that will play a crucial role in defining the Carnot-Carathéodory distance and later in the variational formulas.

Definition 2.3. *An absolutely continuous curve $\xi : [0, T] \to \mathbb{R}^N$ is called horizontal if there exists $\alpha^\xi : [0, T] \to \mathbb{R}^m$ measurable such that*

$$\dot{\xi}(s) = \sum_{i=1}^{m} \alpha_i^\xi(s) X_i(\xi(s)), \quad \text{a.e. } s \in (0, T), \tag{2.4}$$

where the vector fields X_i are those introduced in Definition 2.1.
The vector α^ξ is called horizontal velocity *of the curve.*

Remark 2.1. *Note that whenever $X_1, \dots, X_m$ are linearly independent, as they are always in the case of Carnot groups (see e.g. [16, Ch. 1]), the vector α^ξ is unique up to a measure zero set.*

Let us define the Carnot-Carathéodory distance (briefly C-C distance) associated to a family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$.

Definition 2.4. *Given two points $x, y \in \mathbb{R}^N$ and a family of smooth vector fields on $X_1, \dots, X_m$, we define the Carnot-Carathéodory distance as the minimal length distance (or geodesic distance) among all horizontal curves joining x to y, that is*

$$d_{CC}(x, y) = \inf \left\{ \int_0^T |\alpha^\xi(t)|\, dt \,\middle|\, \xi(0) = x,\ \xi(T) = y \text{ and } \xi \text{ is horizontal} \right\},$$

where $|\alpha^\xi(t)|$ is the Euclidean norm of the m-valued horizontal velocity.

Whenever $X_1, \dots, X_m$ satisfy the Hörmander condition (as in our case of Carnot groups), then $d_{CC}(x, y) < +\infty$ for all $x, y \in \mathbb{R}^N$ and it is continuous w.r.t. the Euclidean topology on $\mathbb{R}^N$.
We denote by $\|x\|_{CC} := d_{CC}(x, 0)$ the Carnot-Carathéodory norm.

Remark 2.2. *The Carnot-Carathéodory distance is globally equivalent to the so-called minimal-time (or control) distance that is defined as*

$$\hat{d}(x, y) := \inf\{T \geq 0 | \exists\ \xi \text{ subunit horizontal in } [0, T] \text{ with } \xi(0) = x, \xi(T) = y\},$$

where an absolutely continuous curve $\xi : [0, T] \to \mathbb{R}$ is called subunit horizontal *if satisfies (2.4) and $|\alpha^\xi(t)| \leq 1$ for a.e. $t \in [0, T]$.*

Note that, even if it is possible to give an explicit formulation for the Carnot-Carathéodory distance in $\mathbb{H}$ (by computing the geodesics), this is extremely complicated so we omit that.
Thus we will need to use both the Carnot-Carathéodory distance and the homogeneous distance, so it is important to recall the relation between these distances and between them and the standard Euclidean distance in $\mathbb{R}^N$.

Lemma 2.2. *Let d_h and d_{CC} be the homogeneous distance and the Carnot-Carathéodory distance defined respectively in Definitions 2.2 and 2.4. Then for any compact $K \subset \mathbb{R}^N$ there exists a positive constant C_K such that*

$$C_K^{-1}|x-y| \leq d_{CC}(x,y) \leq C_K|x-y|^{1/r},$$

where r is the step of the Carnot group and $|x-y|$ denotes here the standard Euclidean distance in $\mathbb{R}^N$.
The same statement holds also replacing d_h and d_{CC}.
Moreover d_h and d_{CC} are equivalent distance on compact sets, i.e. for any compact $K \subset \mathbb{R}^N$ there exists a positive constant c_K such that

$$c_K^{-1} d_h(x,y) \leq d_{CC}(x,y) \leq c_K d_h(x,y).$$

For the proof we refer to the monograph [16].

In the following Lemma we collect several properties of horizontal curves that will be very useful later.

Lemma 2.3. *Let ξ be a horizontal curve with velocity $\alpha^\xi(s)$ such that $\xi(0) = x$ and $\xi(t) = y$. Then the following properties hold:*

(i) For any $z \in \mathbb{R}^N$, $\widetilde{\xi}(s) := z \circ \xi(s)$ is still horizontal with $\alpha^{\widetilde{\xi}}(s) = \alpha^\xi(s)$, $\widetilde{\xi}(0) = z \circ x$ and $\widetilde{\xi}(t) = z \circ y$.

(ii) For any $C > 0$, $\eta(s) := \xi(Cs)$ is still horizontal with $\alpha^\eta(s) = C\alpha^\xi(Cs)$, $\eta(0) = x$ and $\eta(t/C) = y$.

(iii) For any $\lambda > 0$, $\hat{\xi}(s) := \delta_\lambda(\xi(s))$ is still horizontal with $\alpha^{\hat{\xi}}(s) = \lambda\alpha^\xi(s)$.

Proof.

(i) Denote by L_z the left translation w.r.t. z (i.e. $L_z(x) = z \circ x$) and by DL_z the differential of the left translation L_z. We have

$$\begin{aligned}
\dot{\widetilde{\xi}}(s) &= DL_z(\xi(s))\,\dot{\xi}(s) = DL_z(\xi(s))\left(\sum_{i=1}^m \alpha_i^\xi(s) X_i(\xi(s))\right) = \\
&= \sum_{i=1}^m \alpha_i^\xi(s) DL_z(\xi(s))\, X_i(\xi(s)) = \sum_{i=1}^m \alpha_i^\xi(s) X_i(z \circ \xi(s)) = \\
&= \sum_{i=1}^m \alpha_i^\xi(s) X_i(\widetilde{\xi}(s)),
\end{aligned}$$

where we have used the fact that the vector fields X_i are left-invariant by definition, i.e. $DL_z(\xi(s))\, X_i(\xi(s)) = X_i(z \circ \xi(s))$, for all z.

(ii) For any $C \in \mathbb{R}$, given $\eta(s) = \xi(Cs)$, then

$$\dot{\eta}(s) = C\dot{\xi}(Cs) = C\bigg(\sum_{i=1}^{m} \alpha_i^{\xi}(Cs) X_i(\xi(Cs))\bigg) = \sum_{i=1}^{m} C\alpha_i^{\xi}(Cs) X_i(\eta(s)), \tag{2.5}$$

so η is horizontal with $\alpha^{\eta}(s) = C\alpha^{\xi}(Cs)$.

(iii) Using the fact that we are in exponential coordinates and the definition of dilations as automorphisms of the group by the exponential map, that is:

$$\delta_\lambda\left(\exp\left(\sum_{i=1}^{r}\sum_{j=1}^{m_i} g_{j,i}X_{j,i}\right)\right) = \exp\left(\sum_{i=1}^{r}\sum_{j=1}^{m_i} \lambda^i g_{j,i}X_{j,i}\right),$$

where $X_{j,i}$ for $j = 1, \dots, m_i$ are a basis for the layer V_i, and $g_{j,i}$ are the associated exponential coordinates for the point $g \in \mathbb{G} = \mathbb{R}^N$. From the previous formula written for horizontal curves, that means $i = 1$ and $j = 1, \dots, m_1 = m$, it follows immediately that $\hat{\xi}(s) := \delta_\lambda(\xi(s))$ is horizontal and $\alpha^{\hat{\xi}}(s) = \lambda\alpha^{\xi}(s)$.

□

The following lemma proves that we can control the supremum norm of two curves by the L^1-norm of the associated horizontal velocity.

Lemma 2.4. *Consider two measurable functions $\alpha, \beta : [0,T] \to \mathbb{R}^m$ and the associated horizontal curves ξ^α, ξ^β starting from the same initial point, i.e.*

$$\dot{\xi}^\alpha(s) = \sum_{i=1}^{m} \alpha_i(s) X_i\big(\xi^\alpha(s)\big), \quad \dot{\xi}^\beta(s) = \sum_{i=1}^{m} \beta_i(s) X_i\big(\xi^\beta(s)\big), \quad \xi^\alpha(0) = \xi^\beta(0).$$

If α, β are equi-bounded in $L^1(0,T)$, then there exists a positive constant $C > 0$ such that

$$\left\|\xi^\alpha - \xi^\beta\right\|_\infty \le C \left\|\alpha - \beta\right\|_{L^1(0,T)}.$$

Proof. The proof is trivial. In fact:

$$\xi^\alpha(t) - \xi^\beta(t) = \int_0^t \left[\dot{\xi}^\alpha(s) - \dot{\xi}^\beta(s)\right] ds = \sum_{i=1}^{m} \int_0^t \left[\alpha_i(s) X_i\big(\xi^\alpha(s)\big) - \beta_i(s) X_i\big(\xi^\beta(s)\big)\right] ds$$

At this point we add $\pm\alpha_i(s) X_i\big(\xi^\beta(s)\big)$, we use that the vector fields are smooth (so in particular locally Lipschitz continuous) and that α, β are equibounded. ξ^α, ξ^β are equi-bounded, Hence by Gronwall's inequality one can easily conclude.

□

The manifold structure is crucial when one works with PDEs. In fact vector fields allow us to define naturally derivatives of any order, just considering how a vector field acts on smooth functions. Since we are interested in first-order Hamilton-Jacobi equations we introduce only the first derivatives.

Definition 2.5 (Horizontal gradient)**.** *For a Carnot group defined on $\mathbb{R}^N$, consider the family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$ associated to the Carnot group. The horizontal gradient is defined as*

$$\mathcal{X}u := \big(X_1 u\big)X_1 + \cdots + \big(X_m u\big)X_m.$$

Remark 2.3. *Note that $\mathcal{X}u$ is always an element of the distribution $\mathcal{H}$ since it is defined as a linear combination of the vector fields that span the distribution. For sake of simplicity we will often identify the horizontal gradient (which is a N-dimensional object in $\mathcal{H}$) with its coordinate vector $\nabla_{\mathcal{X}} u = (X_1 u, \dots, X_m u)^T$ which is instead an element in $\mathbb{R}^m$. Trivially $\nabla_{\mathcal{X}} u = \sigma(x)\nabla u$ where ∇u denotes the standard (Euclidean) gradient of u and σ is defined in* (2.3)*.*

Example 2.3. *In the case of $\mathbb{H}$ the horizontal gradient can be explicitly written as*

$$\nabla_{\mathcal{X}} u = \begin{pmatrix} u_{x_1} - \frac{x_2}{2} u_{x_3} \\ u_{x_2} + \frac{x_1}{2} u_{x_3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{x_2}{2} \\ 0 & 1 & \frac{x_1}{2} \end{pmatrix} \nabla u \in \mathbb{R}^2,$$

while $\mathcal{X}u = \big(u_{x_1} - \frac{x_2}{2} u_{x_3}\big)\, X_1(x_1, x_2, x_3) + \big(u_{x_2} + \frac{x_1}{2} u_{x_3}\big)\, X_2(x_1, x_2, x_3) \in \mathbb{R}^3$.

3 Statement of the problem.

For any $\varepsilon > 0$, we look at the following family of randomly perturbed problems:

$$\begin{cases} u_t^\varepsilon + H\left(\delta_{1/\varepsilon}(x), \sigma(x)\nabla u^\varepsilon, \omega\right) = 0, \ x \in \mathbb{R}^N, \ \omega \in \Omega, \ t > 0, \\ u^\varepsilon(0, x, \omega) = g(x), \ x \in \mathbb{R}^N, \ \omega \in \Omega, \end{cases} \tag{3.1}$$

where $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ is a probability space, $\sigma(x)$ is a smooth $m \times N$ matrix (with $m \leq N$), whose columns are vector fields $X_1, X_2, \dots, X_m$ associated to a Carnot group, $\delta_\lambda(\cdot)$ are the anisotropic dilations defined by the Carnot group structure. We assume that the Hamiltonian $H : \mathbb{R}^N \times \mathbb{R}^m \times \Omega \to \mathbb{R}$ satisfies the following assumptions w.r.t. $x \in \mathbb{R}^N$, $q = \sigma(x)p \in \mathbb{R}^m$ and $\omega \in \Omega$:

(H1) $q \mapsto H(x, q, \omega)$ is convex in q;

(H2) $\exists\, \overline{C}_1 > 0$, $\overline{\lambda} > 1$ such that

$$\overline{C}_1^{-1}(|q|^{\overline{\lambda}} - 1) \leq H(x, q, \omega) \leq \overline{C}_1(|q|^{\overline{\lambda}} + 1), \ \forall\, (x, q, \omega) \in \mathbb{R}^N \times \mathbb{R}^m \times \Omega;$$

(H3) there exists $\overline{m} : [0, +\infty) \to [0, +\infty)$ concave, monotone increasing with $\overline{m}(0^+) = 0$ and such that for all $x, y \in \mathbb{R}^N, q \in \mathbb{R}^m, \omega \in \Omega$

$$|H(x, q, \omega) - H(y, q, \omega)| \leq \overline{m}\big(\| - x \circ y\|_h (1 + |q|^{\overline{\lambda}})\big);$$

(H4) $\forall\, q \in \mathbb{R}^m$ the function $(x, \omega) \mapsto H(x, q, \omega)$ is a stationary, ergodic random field on $\mathbb{R}^N \times \Omega$ w.r.t. the unitary translation operator, associated to the Carnot group structure.

Recall that $|q|$ is the usual Euclidean norm in $\mathbb{R}^m$ while $\|x\|_h$ is the homogeneous distance in Definition 2.2, and $x^{-1} = -x$ in exponential coordinates.

Example 3.1 (Main model)**.** The main model that we have in mind is:

$$H(x,q,\omega) = a(x,\omega)\frac{|q|^\beta}{\beta} + V(x,\omega), \tag{3.2}$$

where $\beta > 1$, the functions V and a are bounded and satisfy **(H3)-(H4)** and a is also uniformly strictly positive.

Remark 3.1 (Non-coercivity of the Hamiltonian)**.** *Note that the main difference between these assumptions and the assumptions in [37] is that* the Hamiltonian is not anymore coercive w.r.t. the gradient but only w.r.t. the lower dimensional horizontal gradient *(assumption* **(H2)***). This lack of coercivity w.r.t. the total gradient variable p is what will make the approach in [37] and [36] failing and will lead to the main technical difficulties.*

Remark 3.2. *Assumption* **(H3)** *is adapted to the anisotropic dilations in the group. Nevertheless using Lemma 2.2 this assumption can be rewritten in terms of the standard Euclidean distance (with a power depending on the step of the group). In particular if $H(x,p,\omega)$ is Lipschitz continuous in x w.r.t. the homogeneous distance then it is only Hölder continuous in x w.r.t. the standard Euclidean distance with power $1/r$ and r =step of the Carnot group. E.g. in the Heisenberg group it would be $1/2$-Hölder continuous.*

We next write more explicitly assumption **(H4)** to show how this adapts to the algebraic structure of the Carnot group.
Assumption **(H4)** means that there exists a family of measure-preserving maps $\tau_x : \Omega \to \Omega$, indexed by either x in the Carnot group or x in a discrete version of the Carnot group ($\mathbb{Z}^N$ as subset of the Carnot group equipped with the group operation of the Carnot group) with the following properties:

- $\tau_0 = id$
- $\tau_x(\tau_y(\omega)) = \tau_{x\circ y}(\omega)$ for all $x, y \in \mathbb{R}^N$ (or $\mathbb{Z}^N$, respectively) and almost all $\omega \in \Omega$.
- (Stationarity) $H(x,q,\omega) = H(0,q,\tau_x(\omega))$ for all $x \in \mathbb{R}^N$ (or $\mathbb{Z}^N$, respectively) and almost all $\omega \in \Omega$.
- (Ergodicity) If $A \subseteq \Omega$ is such that $\tau_x(A) = A$ for all $x \in \mathbb{R}^N$ (or $\mathbb{Z}^N$, respectively), then $\mathbb{P}(A) \in \{0,1\}$.

Examples are short-correlated Euclidean-stationary random fields (by Borel-Cantelli) or (for the discrete Heisenberg group) Heisenberg-periodic sets where an independent identically distributed random variable is chosen for each cell.

Remark 3.3. *We would like to add some remarks concerning the assumptions on ergodicity and stationarity.*

1. *Note that we apply the ergodic theorem only to a one-dimensional* Abelian *subgroup of the Carnot group ($\mathcal{X}$-lines), so for convergence we are completely in the framework of classical ergodic theory, see proof of Theorem 5.2. The ergodicity with respect to actions of the full group is only used to establish that the limit is deterministic.*

2. *In case of the Example 3.1, and short-correlated random coefficients, the convergence to a deterministic limit in Theorem 5.2 follows already from the law of large numbers.*

3. *For examples of ergodic actions of the Heisenberg group on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ see e.g. [21]. In order to obtain from this a model like Example 3.1 with $a(x,\omega) = 1$, take a bounded random variable $V : \Omega \to \mathbb{R}$ and set $V(x,\omega) = V(\tau_x(\omega))$.*

Example 3.2. *An explicit example for a model satisfying* **(H4)** *in the special case of Example 3.1 for the Heisenberg group $\mathbb{H}$ can be constructed in the following way:*
Take three independent random fields on $\mathbb{R}$, $f_i(x,\omega) : \mathbb{R} \times \Omega \to \mathbb{R}$, for $i = 1, 2, 3$, such that they are stationary ergodic w.r.t. the action of $\mathbb{R}$. Then for a Borel-measurable bounded function $G : \mathbb{R}^3 \to \mathbb{R}$ the random potential

$$V(x_1, x_2, x_3, \omega) := G(f_1(x_1, \omega), f_2(x_2, \omega), f_3(x_3, \omega))$$

is Heisenberg-stationary. Indeed, by independence and one-dimensional stationarity we have for any open intervals $(a_1, a_2), (b_1, b_2), (c_1, c_2)$ that

$$\begin{aligned}
&\mathbb{P}\{(x_1 + r, x_2 + s, x_3 + t + 1/2(x_1 s - x_2 r)) \in (a_1, a_2) \times (b_1, b_2) \times (c_1, c_2)\} \\
&= \mathbb{P}\{f_1(x_1 + r) \in (a_1, a_2)\}\mathbb{P}\{f_2(x_2 + s) \in (b_1, b_2)\} \\
&\times \mathbb{P}\{f_3(x_3 + t + 1/2(x_1 s - x_2 r)) \in (c_1, c_2)\} \\
&= \mathbb{P}\{f_1(x_1) \in (a_1, a_2)\}\mathbb{P}\{f_2(x_2) \in (b_1, b_2)\}\mathbb{P}\{f_3(x_3) \in (c_1, c_2)\} \\
&= \mathbb{P}\{(x_1, x_2, x_3) \in (a_1, a_2) \times (b_1, b_2) \times (c_1, c_2)\}.
\end{aligned}$$

Since open rectangles generate the Borel-σ-algebra, the result follows.

We introduce

$$u^\varepsilon(t, x, \omega) := \inf_{y \in \mathbb{R}^N} \left\{ g(y) + L^\varepsilon(x, y, t, \omega) \right\} \tag{3.3}$$

with

$$L^\varepsilon(x, y, t, \omega) := \inf_{\xi \in \mathcal{A}^t_{y,x}} \int_0^t H^* \left(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega \right) ds \tag{3.4}$$

where

$$\mathcal{A}^t_{y,x} := \left\{ \xi \in W^{1,\infty}\big((0,t)\big) \,\middle|\, \xi \text{ horizontal curve such that } \xi(0) = y \text{ and } \xi(t) = x \right\},$$

while H^* denotes the Legendre-Fenchel transform of H w.r.t. $q \in \mathbb{R}^m$, that is

$$H^*(x, q, \omega) := \sup_{p \in \mathbb{R}^m} \{p \cdot q - H(x, p, \omega)\}$$

and $\alpha^\xi(s)$ is the m-valued measurable function corresponding to the horizontal velocity of $\xi(s)$ defined in (2.4).

Theorem 3.1. *We assume g bounded and Euclidean Lipschitz continuous, and that the Hamiltonian H satisfies* **(H1)-(H3)**. *Then, for all fixed $\omega \in \Omega$, $u^\varepsilon(t, x, \omega)$ given by formula* (3.3) *is the unique BUC viscosity solution of problem* (3.1).

The proof of this theorem will be given in the Appendix.
From now on we use the notation:

$$L(x,q,\omega) := H^*(x,q,\omega) = \sup_{p\in\mathbb{R}^m}\{p\cdot q - H(x,p,\omega)\}; \tag{3.5}$$

$L(x,\cdot,\omega)$ is the Legendre-Fenchel transform of the Hamiltonian $H(x,\cdot,\omega)$ taken w.r.t. $q\in\mathbb{R}^m$, for each $x\in\mathbb{R}^N$ and $\omega\in\Omega$ fixed, and it is called *Lagrangian* associated to the given Hamiltonian. In the following lemma we show how the properties of the Hamiltonian pass to the associated Lagrangian.

Lemma 3.1. *If $H(x,q,\omega)$ satisfies* **(H1)-(H4)** *and $L(x,q,\omega)$ is the associated Lagrangian defined by* (3.5)*, then L satisfies*

(L1) *$q\mapsto L(x,q,\omega)$ is convex, for all $(x,\omega)\in\mathbb{R}^N\times\Omega$;*

(L2) *there exists $C_1>0$ such that*

$$C_1^{-1}(|q|^\lambda - 1) \le L(x,q,\omega) \le C_1(|q|^\lambda+1), \quad \forall\,(x,q,\omega)\in\mathbb{R}^N\times\mathbb{R}^m\times\Omega,$$

where $\lambda = \overline{\lambda}^* := \frac{\overline{\lambda}}{\overline{\lambda}-1}$ (i.e., the conjugate of the constant $\overline{\lambda}$ in assumption **(H2)***);*

(L3) *for all $R>0$, there exists $m_R:[0,+\infty)\to[0,+\infty)$ concave, monotone increasing, with $m_R(0^+)=0$ and such that for all $x,y\in\mathbb{R}^N,\omega\in\Omega$*

$$|L(x,q,\omega)-L(y,q,\omega)| \le m_R\big(\|-x\circ y\|_h\big), \quad \forall\, q\in B_R(0),$$

where $B_R(0)$ is the (Euclidean) ball in $\mathbb{R}^m$ of radius R centered in 0;

(L4) *$\forall\, q\in\mathbb{R}^m$ the function $(x,\omega)\mapsto L(x,q,\omega)$ is stationary, ergodic random field on, $\mathbb{R}^N\times\Omega$ w.r.t. the unitary translation operator associated to the Carnot group structure;*

(L5) *$L^* = H$.*

Proof. Properties **(L1)** and **(L5)** follow immediately by definition of $L=H^*$. Property **(L2)** comes trivially from **(H2)** and (3.5) taking $\lambda = \overline{\lambda}^{\,*}$, the conjugate exponent of $\overline{\lambda}$. The proofs of **(L3)** and **(L4)** are also immediate: one can find a detailed proof of **(L3)** in [17, Theorem A.2.6] while **(L4)** comes directly from the definition of $L=H^*$. □

Remark 3.4. *We note that w.l.o.g. we can replace assumption* **(L2)** *with*

$$C_1^{-1}(|q|^\lambda+1) \le L(x,q,\omega) \le C_1(|q|^\lambda+1). \tag{3.6}$$

Indeed, if we replace $L(x,q,\omega)$ by $\hat{L}((x,q,\omega) = L(x,q,\omega)+2C_1^{-1}$ then $\hat{u}^\varepsilon(t,x) = u^\varepsilon(t,x)+2C_1^{-1}t$ making no relevant change in the homogenisation problem.

Example 3.3. *In the case of model* (3.2) *in Example 3.1, the associated Lagrangian is*

$$L(x,q,\omega) = b(x,\omega)\frac{|q|^{\beta^*}}{\beta^*} + V(x,\omega), \tag{3.7}$$

with $\beta^* = \frac{\beta}{\beta-1}$ and $b(x,\omega) = \frac{1}{a(x,\omega)^{\frac{1}{\beta-1}}}$.

Here we state the main results of this paper. The proof will be given in Section 7.

Theorem 3.2. *Consider the problem* (3.1) *with* $g : \mathbb{R}^N \to \mathbb{R}$ *bounded and (Euclidean) Lipschitz continuous. Assume that the Hamiltonian* $H(x, p, \omega)$ *satisfies assumptions* **(H1)-(H4)** *and for* $L = H^*$ *the following additional assumption holds:*

(L6) *there exist constants* $C > 0$ *and* $\lambda > 1$ *such that*

$$|L(x, p, \omega) - L(x, s\,p, \omega)| \leq C\big|1 - |s|^{\lambda}\big|\,|L(x, p, \omega)|,$$

for all $s \in \mathbb{R},\ x \in \mathbb{R}^N, p \in \mathbb{R}^m,\ \omega \in \Omega$.

Then the viscosity solutions $u^{\varepsilon}(t, x, \omega)$ *of problem* (3.1) *converge locally uniformly in* x *and* $t > 0$ *and a.s. in* $\omega \in \Omega$ *to the unique solution* $u(t, x)$ *of the deterministic problem* (1.3)*, where the effective Hamiltonian* $\overline{H}(q)$ *is defined as* $\overline{H}(q) = \overline{L}^*(q)$ *and* $\overline{L}(q)$ *is the effective Lagrangian defined by limit* (5.13)*.*

Now we give a class of operators where we can apply the previous result.

Corollary 3.1. *Consider the problem* (3.1) *with* $g : \mathbb{R}^N \to \mathbb{R}$ *bounded and Lipschitz continuous. Assume that the Hamiltonian* $H(x, p, \omega)$ *satisfies assumptions* **(H1)-(H4)** *and moreover*

$$H(x, p, \omega) = H_1(x, p, \omega) + H_2(x, \omega), \tag{3.8}$$

with $H_1(x, p, \omega)$ λ*-homogeneous in* p *(namely* $H_1(x, s\,p, \omega) = |s|^{\lambda} H_1(x, p, \omega)$*). Then the viscosity solutions* $u^{\varepsilon}(t, x, \omega)$ *of* (3.1) *converge locally uniformly in* x *and* $t > 0$ *and a.s. in* $\omega \in \Omega$ *to the unique solution* $u(t, x)$ *of the deterministic problem* (1.3)*, where the effective Hamiltonian* $\overline{H}(q)$ *is defined as* $\overline{H}(q) = \overline{L}^*(q)$ *and* $\overline{L}(q)$ *is the effective Lagrangian defined by limit* (5.13)*.*

Proof of Corollary 3.1. We need only to remark that, whenever H satisfies (3.8), then the associated Lagrangian has the same structure (by taking $\lambda = \overline{\lambda}^*$), and such a structure for L implies assumption **(L6)**. Hence Theorem 3.2 and Remark 3.4 immediately imply the result. □

Example 3.4. *Our main model of Hamiltonian* (3.2) *satisfies the assumptions of Corollary 3.1.*

4 Properties of $L^{\varepsilon}(x, y, t, \omega)$.

In this section we will investigate several properties for the variational problem $L^{\varepsilon}(x, y, t, \omega)$ defined by (3.4).

Lemma 4.1. *Under assumption* **(L2)** *then*

$$C_1^{-1} t \leq L^{\varepsilon}(x, y, t, \omega) \leq C_1 t + C_1 \frac{d_{CC}(x, y)^{\lambda}}{t^{\lambda - 1}}, \qquad \forall x, y \in \mathbb{R}^N, t > 0, \omega \in \Omega,$$

where C_1 *and* λ *are the constants introduced in assumption* **(L2)**.

Proof. By assumption **(L2)** and Remark 3.4 we get the first inequality of the statement. We consider a geodesic η from y to x in time t (parametrized by arc-length), then $\eta \in \mathcal{A}^t_{y,x}$ with $|\alpha^\eta(s)| = d_{CC}(x,y)/t$. By assumption **(L2)** we have

$$L^\varepsilon(x,y,t,\omega) \le \int_0^t L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega)\, ds \le \int_0^t C_1(1 + d_{CC}(x,y)^\lambda t^{-\lambda})\, ds,$$

which easily entails the last inequality of the statement. □

Proposition 4.1. *Under assumption* **(L2)** *then*

$$L^\varepsilon(x,y,t,\omega) = \inf_\eta \left\{ \int_0^t L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega)\, ds \right\}, \tag{4.1}$$

where the infimum is taken over the curves

$$\eta \in \mathcal{A}^t_{y,x} \qquad \text{such that } \|\alpha^\eta\|_{L^\lambda(0,t)} \le \widetilde{C}, \tag{4.2}$$

where $\widetilde{C} = [(C_1^2 + C_1 + 1)t^\lambda + C_1^2 d_{CC}(x,y)^\lambda]^{1/\lambda} t^{\frac{1}{\lambda}-1}$.

Proof. Using the definition of $L^\varepsilon(x,y,t,\omega)$ as greatest lower bound, then there exists a curve $\eta \in \mathcal{A}^t_{y,x}$ such that

$$\int_0^t L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega)\, ds \le L^\varepsilon(x,y,t,\omega) + t.$$

By the bound from below in **(L2)** and using Lemma 4.1, we have

$$C_1^{-1} \int_0^t \left(|\alpha^\eta(s)|^\lambda - 1 \right) ds \le L^\varepsilon(x,y,t,\omega) + t \le (C_1+1)t + C_1 \frac{d_{CC}(x,y)^\lambda}{t^{\lambda-1}},$$

and consequently

$$\|\alpha^\eta\|^\lambda_{L^\lambda(0,t)} \le (C_1^2 + C_1 + 1)t + C_1^2 \frac{d_{CC}(x,y)^\lambda}{t^{\lambda-1}},$$

which is equivalent to relation (4.2). □

Corollary 4.1. *Under assumption* **(L2)**, *locally uniformly for* $(x,y,t) \in \mathbb{R}^N \times \mathbb{R}^N \times (0,T]$ *the infimum in* $L^\varepsilon(x,y,t,\omega)$ *is attained over admissible curves* η *such that*

(i) $\|\alpha^\eta\|_{L^\gamma(0,t)} \le C_2$, *for all* $1 \le \gamma \le \lambda$, *where* λ *is the constant given in assumption* **(L2)** *and* C_2 *depends only on the constants in assumption* **(L2)** *and the compact sets;*

(ii) $\|\eta\|_\infty \le C_3$, *where* C_3 *depends only on* $\sigma(x)$, *the constants in assumption* **(L2)** *and the compact sets.*

Proof. This follows easily by Proposition 4.1, the standard embedding properties of $L^\lambda(0,t)$ and $\xi(t) = \xi(0) + \int_0^t \sigma(\xi(s))\alpha^\xi(s) ds$. □

We now prove the uniform continuity of L^ε, uniformly w.r.t. $\varepsilon > 0$. To this purpose, we adapt some arguments from [36].

Lemma 4.2. *Under assumption* **(L2)**, *then*

$$L^\varepsilon(x, y, t + h, \omega) \le L^\varepsilon(x, y, t, \omega) + C_1 h, \tag{4.3}$$

for all $\varepsilon > 0, x, y \in \mathbb{R}^N, t > 0$, $h \ge 0$, $\omega \in \Omega$ *and where* C_1 *is the constant introduced in* **(L2)**.

Proof. We proceed as in [36]. Let ξ be an admissible path for $L^\varepsilon(x, y, t, \omega)$; we introduce

$$\xi_1(s) := \begin{cases} \xi(s) & 0 \le s \le t, \\ x & t \le s \le t + h. \end{cases}$$

Note that ξ_1 is an admissible path for $L^\varepsilon(x, y, t + h, \omega)$. Hence, we have

$$\begin{aligned} L^\varepsilon(x, y, t + h, \omega) &\le \int_0^{t+h} L(\delta_{1/\varepsilon}(\xi_1(s)), \alpha^{\xi_1}(s), \omega) ds \\ &\le \int_0^t L(\delta_{1/\varepsilon}(\xi(s)), \alpha^{\xi}(s), \omega) ds + \int_t^{t+h} L(\delta_{1/\varepsilon}(x), 0, \omega) ds. \end{aligned}$$

Taking the infimum over ξ and by assumption **(L2)**, we obtain the statement. □

Lemma 4.3. *Under assumption* **(L2)**, *then*

$$L^\varepsilon(x \circ v, y, t + \|v\|_{CC}, \omega) \le L^\varepsilon(x, y, t, \omega) + 2C_1 \|v\|_{CC}, \tag{4.4}$$

for all $\varepsilon > 0, x, y, v \in \mathbb{R}^N, t > 0$, $\omega \in \Omega$ *and where* C_1 *is the constant introduced in assumption* **(L2)**. *Moreover, for each compact* $K \subset\subset \mathbb{R}^N$, *there exists a constant* C *(depending only on* K *and on the assumptions of the problem, so in particular independent of* ε*) such that*

$$L^\varepsilon(x \circ v, y, t + \|v\|_h, \omega) \le L^\varepsilon(x, y, t, \omega) + C\|v\|_h \qquad \forall v \in K, \tag{4.5}$$

and for all $\varepsilon > 0, x, y \in \mathbb{R}^N, t > 0$, $\omega \in \Omega$.

Proof. For any curve $\xi \in \mathcal{A}^t_{y,x}$, we define

$$\eta_\xi(s) := \begin{cases} \xi(s) & 0 \le s \le t \\ \tilde\gamma(s) & t \le s \le t + \|v\|_{CC} \end{cases}$$

where $\tilde\gamma(s) = \gamma(s - t)$ and γ is a geodesic from x to $x \circ v$ in time $\|v\|_{CC}$. Since $\eta_\xi \in \mathcal{A}^{t+\|v\|_{CC}}_{y,x\circ v}$, then we have

$$\begin{aligned} L^\varepsilon(x \circ v, y, t + \|v\|_{CC}, \omega) &\le \int_0^{t+\|v\|_{CC}} L(\delta_{1/\varepsilon}(\eta_\xi(s)), \alpha^{\eta_\xi}(s), \omega) ds \\ &= \int_0^t L(\delta_{1/\varepsilon}(\xi(s)), \alpha^{\xi}(s), \omega) ds + \int_t^{t+\|v\|_{CC}} L(\delta_{1/\varepsilon}(\tilde\gamma(s)), \alpha^{\tilde\gamma}(s), \omega) ds \\ &\le \int_0^t L(\delta_{1/\varepsilon}(\xi(s)), \alpha^{\xi}(s), \omega) ds + 2C_1 \|v\|_{CC}, \end{aligned}$$

where the last inequality is due to assumption **(L2)** and the fact that by arc-length parametrisation we can assume $|\alpha^{\tilde{\gamma}}(s)| = 1$. Taking the infimum over ξ, we get the bound (4.4).

It remains to prove (4.5). We argue as before defining

$$\eta_\xi(s) := \begin{cases} \xi(s) & 0 \le s \le t \\ \tilde{\gamma}_1(s) & t \le s \le t + \|v\|_h \end{cases}$$

where $\tilde{\gamma}_1(s) := \gamma_1(s-t)$ and γ_1 is a geodesic from x to $x \circ v$ in time $\|v\|_h$. We have $|\alpha^{\gamma_1}(s)| = \|v\|_{CC}/\|v\|_h$. Hence, as before, we get

$$L^\varepsilon(x \circ v, y, t + \|v\|_h, \omega) \le \int_0^t L(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega) ds + C_1 \|v\|_h \left(1 + \frac{\|v\|_{CC}^\lambda}{\|v\|_h^\lambda}\right).$$

Recall that for each compact $K \subset\subset \mathbb{R}^N$, there exists a constant c such that $c^{-1}\|v\|_h \le \|v\|_{CC} \le c\|v\|_h$ (see Lemma 2.2); hence

$$L^\varepsilon(x \circ v, y, t + \|v\|_h, \omega) \le \int_0^t L(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega) ds + C_1(1 + c^\lambda)\|v\|_h$$

and, taking the infimum over ξ, we get (4.5). □

Lemma 4.4. *Under assumptions* **(L2)** *and* **(L6)**, $L^\varepsilon(x, y, t, \omega)$ *is locally uniformly continuous in* t *away from* 0, *locally uniformly w.r.t.* x *and* y *and uniformly w.r.t.* ε. *More precisely for any* $\delta \in (0, 1)$ *there exists* $C_\delta > 0$ *(depending only on the constants in assumptions* **(L2)** *and* **(L6)** *and going to* $+\infty$ *as* $\delta \to 0^+$*) such that*

$$|L^\varepsilon(x, y, t + h, \omega) - L^\varepsilon(x, y, t, \omega)| \le C_\delta h \quad \forall \varepsilon > 0, \tag{4.6}$$

for any $t, t + h \in [\delta, 1/\delta] =: I_\delta$, $h \ge 0$ *and for any* $(x, y) \in A_\delta$ *where*

$$A_\delta := \{(x, y) \in \mathbb{R}^N \times \mathbb{R}^N \,|\, d_{CC}(x, y) < 1/\delta\}.$$

Proof. By Lemma 4.2 we know that for any $h \ge 0$

$$L^\varepsilon(x, y, t + h, \omega) - L^\varepsilon(x, y, t, \omega) \le C_1 h.$$

It remains to show the opposite inequality, i.e.

$$L^\varepsilon(x, y, t + h, \omega) - L^\varepsilon(x, y, t, \omega) \ge -C_\delta h. \tag{4.7}$$

Take $(x, y) \in A_\delta$ and a curve η admissible for $L^\varepsilon(x, y, t + h, \omega)$ and such that $\|\alpha^\eta\|_{L^\lambda(0,t+h)} \le \overline{C}$ where $\overline{C} = \overline{C}(\delta)$ is the constant introduced in Proposition 4.1. We define $\xi_\eta(s) := \eta(\frac{t+h}{t}s)$. By Lemma 2.3, $\xi_\eta(s)$ is still horizontal with

$$\alpha^{\xi_\eta}(s) = \frac{t+h}{t}\alpha^\eta\left(\frac{t+h}{t}s\right),\ \xi_\eta(0) = \eta(0) = y,\ \xi_\eta(t) = \eta\left(\frac{t+h}{t}t\right) = x;$$

so, $\xi_\eta(s)$ is admissible for $L^\varepsilon(x, y, t, \omega)$. We observe that

$$\begin{aligned} L^\varepsilon(x, y, t, \omega) &\le \int_0^t L(\delta_{1/\varepsilon}(\xi_\eta(s)), \alpha^{\xi_\eta}(s), \omega) ds \\ &= \frac{t}{t+h}\int_0^{t+h} L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) ds. \end{aligned}$$

Define

$$I := \frac{t}{t+h}\int_0^{t+h} L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) ds - \int_0^{t+h} L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega) ds,$$

so that

$$\int_0^{t+h} L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega) ds - L^\varepsilon(x, y, t, \omega) \geq -I. \tag{4.8}$$

Assume for the moment that

$$I \leq C_\delta h. \tag{4.9}$$

Then passing to the infimum over η in (4.8) we obtain relation (4.7).
Let us now prove (4.9). Writing $\frac{t}{t+h} = 1 - \frac{h}{t+h}$, we have

$$I = \int_0^{t+h} \left(L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) - L\left(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega\right)\right) ds + \\ - \frac{h}{t+h}\int_0^{t+h} L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) ds. \tag{4.10}$$

To get (4.9) we estimate $|I|$. We start estimating the modulus of the latter integral in the right hand side. Note that **(L2)** implies that

$$|L(x, q, \omega)| \leq C_1\big(|q|^\lambda + 1\big) + C_1^{-1}\big(|q|^\lambda - 1\big) \leq C(|q|^\lambda + 1) \tag{4.11}$$

where we have used that $\max\{A, B\} \leq |A| + |B|$ and $C := C_1 + C_1^{-1}$.
Then by (4.11) and for h sufficiently small (so that $\delta < t + h < \frac{1}{\delta}$), we have

$$\begin{aligned}
&\frac{h}{t+h}\int_0^{t+h} \left| L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right)\right| ds \\
&\quad \leq \frac{h}{t+h}\int_0^{t+h} C\left(1 + \left(\frac{t+h}{t}\right)^\lambda |\alpha^\eta(s)|^\lambda\right) ds \\
&\quad \leq \frac{h}{\delta} C\left(\frac{1}{\delta} + \left(\frac{t+h}{t}\right)^\lambda \|\alpha^\eta\|^\lambda_{L^\lambda(0,t+h)}\right) \\
&\quad \leq h\frac{C}{\delta}\left(\frac{1}{\delta} + \left(\frac{\widetilde{C}}{\delta^2}\right)^\lambda\right) = C_\delta h.
\end{aligned} \tag{4.12}$$

where the last inequality is due to Proposition 4.1. On the other hand, using first **(L6)** and then **(L2)**, we have

$$\begin{aligned}
&\int_0^{t+h} \left| L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) - L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega)\right| ds \\
&\quad \leq \int_0^{t+h} C\left|\left(\frac{t+h}{t}\right)^\lambda - 1\right| L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega) ds \\
&\quad \leq C\, C_1 \int_0^{t+h} \left|\left(\frac{t+h}{t}\right)^\lambda - 1\right| (|\alpha^\eta(s)|^\lambda + 1) ds.
\end{aligned}$$

Hence, by our choice of t and h, we get

$$\begin{aligned}\int_0^{t+h} \left| L\left(\delta_{1/\varepsilon}(\eta(s)), \frac{t+h}{t}\alpha^\eta(s), \omega\right) - L(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega)\right| ds \\ \leq C_1\delta^{-\lambda}\big|(t+h)^\lambda - t^\lambda\big|(\|\alpha^\eta\|^\lambda_{L^\lambda(0,t+h)} + t + h) \\ \leq C_\delta h \end{aligned} \tag{4.13}$$

where the last inequality is due to Proposition 4.1. Using (4.12) and (4.13) in (4.10), we get (4.9). □

For later use, we collect some consequences of the previous estimates.

Lemma 4.5. *Assuming* **(L2)** *and* **(L6)**, *there exists a positive contant C, depending only on the assumptions on the Lagrangian L, such that*

1. $L^\varepsilon(x \circ v, x, \|v\|_{CC}, \omega) \leq C\|v\|_{CC}$,
2. $|L^\varepsilon(x, y, t, \omega) - L^\varepsilon(x, y, (1+\rho)t, \omega)| \leq C\rho|L^\varepsilon(x, y, t, \omega)|$, *for* $0 < \rho \ll 1$,
3. $L^\varepsilon(x, z, t+s, \omega) \leq L^\varepsilon(x, y, t, \omega) + L^\varepsilon(y, z, s, \omega)$,

for all $x, v, y, z \in \mathbb{R}^N$, $\omega \in \Omega$ and $t, s > 0$.

Proof. Point 1 follows from Lemma 4.3 by choosing $x = y$ and $t = 0$.
We prove now point 2: Applying Lemma 4.4 with $h = \rho\, t$ we get

$$|L^\varepsilon(x, y, t, \omega) - L^\varepsilon(x, y, (1+\rho)t, \omega)| \leq C_\delta\, \rho\, t.$$

We accomplish the proof recalling from Lemma 4.1 that $C_1^{-1}t \leq L^\varepsilon(x, y, t, \omega)$.
Point 3 is obvious since the combination of two minimisers for $L^\varepsilon(x, y, t, \omega)$ and $L^\varepsilon(y, z, s, \omega)$, respectively, is an admissible path for $L^\varepsilon(x, z, t+s, \omega)$. □

As a direct consequence we have the following lemma:

Lemma 4.6. *Assume* **(L2)** *and* **(L6)** *and consider $x_1, y_1, x_2, y_2 \in \mathbb{R}^N$ and $t > 0$, with $\|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC} \ll t$. Then*

$$\begin{aligned}&|L^\varepsilon(x_1, y_1, t, \omega) - L^\varepsilon(x_2, y_2, t, \omega)| \\ &\leq C\big(|L^\varepsilon(x_1, y_1, t, \omega)| + |L^\varepsilon(x_2, y_2, t, \omega)|\big)\left(\| - x_1 \circ x_2\|_{CC} + \| - y_1 \circ y_2\|_{CC}\right).\end{aligned}$$

Proof. By applying twice Lemma 4.5 part 3, we deduce

$$\begin{aligned}&L^\varepsilon\big(x_1, y_1, t + \|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}, \omega\big) \\ &\leq L^\varepsilon\big(x_1, x_2, \|-x_1 \circ x_2\|_{CC}, \omega\big) + L^\varepsilon(x_2, y_2, t, \omega) + L^\varepsilon(y_2, y_1, \|-y_1 \circ y_2\|_{CC}, \omega).\end{aligned}$$

Next we apply Lemma 4.5 part 1 to $L^\varepsilon\big(x_1, x_2, \|-x_1 \circ x_2\|_{CC}, \omega\big)$ and to $L^\varepsilon(y_2, y_1, \|-y_1 \circ y_2\|_{CC}, \omega)$ taking respectively $v = -x_2 \circ x_1$ and $v = -y_1 \circ y_2$, and we deduce:

$$\begin{aligned}L^\varepsilon\big(x_1, y_1, t + \|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}, \omega\big) \\ \leq C\big(\|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}\big) + L^\varepsilon(x_2, y_2, t, \omega)\end{aligned}$$

(recall $\|-x_2 \circ x_1\|_{CC} = \|-x_1 \circ x_2\|_{CC}$). Finally, Lemma 4.5 part 2 with $\rho = \frac{\|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}}{t}$, implies

$$L^\varepsilon\big(x_1, y_1, t + \|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}, \omega\big) \\ \geq L^\varepsilon(x_1, y_1, t, \omega) - C\,\rho |L^\varepsilon(x_1, y_1, t, \omega)|.$$

The last two inequalities entail

$$L^\varepsilon(x_1, y_1, t, \omega) - L^\varepsilon(x_2, y_2, t, \omega) \leq C\,\rho |L^\varepsilon(x_1, y_1, t, \omega)| + C\rho t.$$

To conclude, from Lemma 4.1 we get $L^\varepsilon(x, y, t, \omega) \geq C_1^{-1} t$ and we recall that $t > 0$ is fixed. Then, up to changing the constant, the previous inequality can be written as

$$L^\varepsilon(x_1, y_1, t, \omega) - L^\varepsilon(x_2, y_2, t, \omega) \leq C |L^\varepsilon(x_1, y_1, t, \omega)| (\|-x_1 \circ x_2\|_{CC} + \|-y_1 \circ y_2\|_{CC}).$$

Reversing the role of x_1, y_1 with that of x_2, y_2, the claim follows. □

Theorem 4.1. *Under assumptions* **(L2)** *and* **(L6)**, $L^\varepsilon(x, y, t, \omega)$ *is locally uniformly continuous w.r.t.* $x, y \in \mathbb{R}^N$ *and* t *away from 0, uniformly w.r.t.* $\varepsilon > 0$ *independently of* $\omega \in \Omega$.

Proof. By Lemma 4.4 we have the local uniform continuity with respect to t. It remains to show the local uniform continuity with respect to x (the one w.r.t y is analogous so it is omitted). We need to show that, for every $\delta > 0$, there exists a constant $\overline{C}_\delta > 0$ such that

$$|L^\varepsilon(x, y, t, \omega) - L^\varepsilon(\tilde{x}, y, t, \omega)| \leq \overline{C}_\delta \, \|-x \circ \tilde{x}\|_h \qquad \forall \varepsilon > 0 \tag{4.14}$$

and for any $t \in [\delta, 1/\delta]$ and for any $x, \tilde{x}, y$ with CC-norm smaller than $\frac{1}{\delta}$. Indeed, we have

$$L^\varepsilon(x, y, t, \omega) - L^\varepsilon(\tilde{x}, y, t, \omega) = [L^\varepsilon(x, y, t, \omega) - L^\varepsilon(x, y, t + \| - \tilde{x} \circ x\|_h, \omega)] + \\ [L^\varepsilon(x, y, t + \| - \tilde{x} \circ x\|_h, \omega) - L^\varepsilon(\tilde{x}, y, t, \omega)].$$

We observe that Lemma 4.3 (with $v = -\tilde{x} \circ x$) and Lemma 4.4 give

$$L^\varepsilon(x, y, t, \omega) - L^\varepsilon(\tilde{x}, y, t, \omega) \leq (2C_1 + C_\delta) \, \|-x \circ \widetilde{x}\|_h \,,$$

where C_1 is the constant introduced in **(L2)** while C_δ is the constant introduced in Lemma 4.4. Reversing the role of x and $\tilde{x}$ we accomplish the proof. □

Lemma 4.7. *Under assumption* **(L2)**, *then*

$$L^\varepsilon(x, y, t, \omega) \geq C_1^{-1} t^{1-\lambda} (d_{CC}(x, y))^\lambda - C_1^{-1} t, \tag{4.15}$$

for all $\varepsilon > 0$, $t > 0$ *and* $x, y \in \mathbb{R}^N$, *where* C_1 *and* λ *are the constants introduced in* **(L2)**.

Proof. By the definition of $L^\varepsilon(x,y,t,\omega)$, assumption **(L2)** and Jensen's inequality, we obtain

$$\begin{aligned} L^\varepsilon(x,y,t,\omega) &\geq C_1^{-1} \inf_{\xi\in\mathcal{A}^t_{y,x}} \left\{ \int_0^t (|\alpha^\xi(s)|^\lambda - 1) ds \right\} \\ &= C_1^{-1} t \inf_{\xi\in\mathcal{A}^t_{y,x}} \left\{ \left(\frac{1}{t} \int_0^t |\alpha^\xi(s)|^\lambda ds \right) \right\} - C_1^{-1} t \\ &\geq C_1^{-1} t^{1-\lambda} \left(\inf_{\xi\in\mathcal{A}^t_{y,x}} \left\{ \int_0^t |\alpha^\xi(s)| ds \right\} \right)^\lambda - C_1^{-1} t \end{aligned}$$

which is equivalent to the statement because of the definition of $d_{CC}(x,y)$. □

Proposition 4.2. *Assume that the Lagrangian L satisfies* **(L2)** *and that the initial datum g satisfies*

$$g(x) \geq -C(1 + d_{CC}(x,0)) \qquad \forall x \in \mathbb{R}^N. \tag{4.16}$$

Then the $\inf_{y\in\mathbb{R}^N} \{g(y) + L^\varepsilon(x,y,t,\omega)\}$ is attained in a CC-ball centred at x with radius depending only on the constants in **(L2)** *and on T for all $t \in (0,T)$. Recall that by Lemma 2.2 CC-balls and standard Euclidean balls can be included into each other up to changing the radius.*

Proof. As in [36, Lemma 3.4], we want to prove that the infimum outside a suitable ball is greater than the infimum over the entire space. Fix $(t,x) \in (0,T) \times \mathbb{R}^N$. From **(L2)** we have that $L(x,0,\omega) \leq C_1$, so $L^\varepsilon(x,x,t,\omega) \leq C_1 t$, that implies

$$\inf_{y\in\mathbb{R}^N} \{g(y) + L^\varepsilon(x,y,t,\omega)\} \leq g(x) + L^\varepsilon(x,x,t,\omega) \leq g(x) + C_1 t, \tag{4.17}$$

where the second inequality is obtained choosing the constant curve $\xi(s) = x$ for any $s \in (0,t)$ in the definition of $L^\varepsilon(x,x,t,\omega)$.
From (4.16) and the triangle inequality, we have

$$\begin{aligned} g(y) + C_1^{-1} t^{1-\lambda} d_{CC}(x,y)^\lambda - C_1^{-1} t \geq & - C - C\, t \frac{d_{CC}(x,y)}{t} - C d_{CC}(x,0) \\ & + C_1^{-1} t \left(\frac{d_{CC}(x,y)}{t} \right)^\lambda - C_1^{-1} t. \end{aligned}$$

Since the right-hand side goes to $+\infty$, as $d_{CC}(x,y) \to +\infty$, then there exist $R > 0$ such that

$$g(y) + C_1^{-1} t^{1-\lambda} d_{CC}(x,y)^\lambda - C_1^{-1} t \geq g(x) + C_1 t \qquad \forall y \in \mathbb{R}^N \backslash D_R \tag{4.18}$$

where $D_R := \{y \in \mathbb{R}^N : d_{CC}(x,y) \leq RT\}$. By using both inequalities (4.17) and (4.18), we get

$$\begin{aligned} \inf_{y\in\mathbb{R}^N} \{g(y) + L^\varepsilon(x,y,t,\omega)\} &\leq \inf_{y\in\mathbb{R}^N\backslash D_R} \left\{ g(y) + C_1^{-1} t \left(\frac{d_{CC}(x,y)}{t} \right)^\lambda - C_1^{-1} t \right\} \\ &\leq \inf_{y\in\mathbb{R}^N\backslash D_R} \left\{ g(y) + L^\varepsilon(x,y,t,\omega) \right\} \end{aligned}$$

where the last inequality is due to Lemma 4.7. To conclude we just remark that, by the Hörmander condition, D_R is contained in an Euclidean ball (up to a different radius), then the lemma is proved. □

5 A lower dimensional constrained problem to determine the effective Lagrangian.

Inspired by the approach of [37], we now pass to study the convergence of the functional $L^\varepsilon(x, y, t, \omega)$ as $\varepsilon \to 0^+$, by using the Subadditive Ergodic Theorem. First we introduce a special family of horizontal curves which can be used as initial condition to build a subadditive stationary process. At this purpose we use curves which have constant horizontal velocity w.r.t. the given family of vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$, namely $\mathcal{X}$-lines. For more details on those curves one can see [9, 10].

Definition 5.1. *We call* $\mathcal{X}$-line *any absolute continuous curve* $\xi : [0, t] \to \mathbb{R}^N$, *satisfying*

$$\dot{\xi}(s) = \sum_{i=1}^{m} q_i X_i(\xi(s)) = \sigma(\xi(s))q, \quad \text{a.e. } s \in (0, t), \tag{5.1}$$

for some constant vector $q \in \mathbb{R}^m$. *Using notation coherent with [37] we denote by* $l_q^{\mathcal{X}}(s)$ *the* $\mathcal{X}$*-line starting from the origin associated to the horizontal constant velocity* $q \in \mathbb{R}^m$.

Remark 5.1.

1. *Since the vector fields associated to Carnot groups are smooth, the* $\mathcal{X}$*-lines are smooth curves so relation* (5.1) *holds for all* $s \in (0, t)$.

2. *Since the vector fields are linearly independent at any points (see Remark 2.1), for any fixed* $q \in \mathbb{R}^m$ *and for any fixed starting point* x, *there is a unique* $\mathcal{X}$*-line starting from the point* x *and associated to the horizontal constant velocity* q.

3. $\mathcal{X}$*-lines starting from a given point* x *are curves in* $\mathbb{R}^N$ *depending only on* m *parameters with* $m < N$. *Then, while there always exists an horizontal curve joining two given points* x *and* y, *in general a* $\mathcal{X}$*-line joining* x *to* y *may not exist in a Carnot group.*

To study the convergence of $L^\varepsilon(x, y, t, \omega)$ we need to use $\mathcal{X}$-lines so we first restrict our attention to the points x and y that can be connected by using a $\mathcal{X}$-line. Following the notations in [10], we define the $\mathcal{X}$-plane associated to a point x which is, roughly speaking the union of all the $\mathcal{X}$-lines starting from x.

Definition 5.2. *We call* $\mathcal{X}$*-plane associated to the point* x *the set of all the points that one can reach from* x *through a* $\mathcal{X}$*-line, i.e.*

$$V_x := \{y \in \mathbb{R}^N \mid \exists\, q \in \mathbb{R}^m \text{ and } \xi^q \ \mathcal{X}\text{-line such that } \xi^q(0) = x, \xi^q(1) = y\}.$$

Remark 5.2 ($\mathcal{X}$-lines in Carnot groups)**.** *Note that in the Heisenberg group the* $\mathcal{X}$*-lines form a subset of Euclidean straight lines but in general the structure of* $\mathcal{X}$*-lines can look very different from the Euclidean straight lines (see [9, 10] for some examples). Still, if we assume that the vector fields are associated to a Carnot group in exponential coordinates (see* (2.3)*), then at least the first* m*-components are Euclidean affine. This implies that in Carnot groups, whenever* $y \in V_x$, *then the unique horizontal velocity* q *such that* $\xi^q(0) = x$ *and* $\xi^q(1) = y$ *is given by* $q = \pi_m(y - x)$.

Let us define

$$\mathcal{B}^q_{a,b}:=\left\{\xi:[a,b]\to\mathbb{R}^N\,|\xi\in W^{1,\infty}\big((a,b)\big)\text{ horizontal},\xi(a)=l^{\mathcal{X}}_q(a),\,\xi(b)=l^{\mathcal{X}}_q(b)\right\}. \tag{5.2}$$

For any interval $[a,b)$ we define the following stochastic process (similar to [37]):

$$\mu_q([a,b),\omega):=\inf_{\xi\in\mathcal{B}^q_{a,b}}\int_a^b L(\xi(s),\alpha^\xi(s),\omega)ds. \tag{5.3}$$

To use the Sub-additive Ergodic Theorem, fixed the slope $q\in\mathbb{R}^m$, we need to consider the action of $\mathbb{Z}$ on the process μ_q as additive translation in time. More precisely:

Definition 5.3. *Given $a,b\in\mathbb{R}$ with $a<b$, $q\in\mathbb{R}^m$ and $z\in\mathbb{Z}$, we define*

$$\tau_z\,\mu_q([a,b),\omega):=\mu_q(z+[a,b),\omega)=\inf_{\xi\in\mathcal{B}^q_{a+z,b+z}}\int_{a+z}^{b+z}L(\xi(s),\alpha^\xi(s),\omega)ds.$$

Lemma 5.1. *For every $q\in\mathbb{R}^m$, $z\in\mathbb{Z}$, let μ_q be defined by (5.3) and τ_z the additive action introduced in Definition 5.3. Under assumptions* **(L1)-(L4)***, we have*

$$\tau_z\,\mu_q([a,b),\omega)=\mu_q([a,b),\tau_{z_q}\omega) \tag{5.4}$$

where $z_q=l^{\mathcal{X}}_q(z)$ and $l^{\mathcal{X}}_q$ is the $\mathcal{X}$-line defined in Definition 5.1.

Proof. For any $\xi\in\mathcal{B}^q_{a+z,b+z}$ we consider

$$\widetilde{\xi}(s):=[l^{\mathcal{X}}_q(z)]^{-1}\circ\xi(s+z).$$

By Lemma 2.3 part (i), $\widetilde{\xi}(s)$ is still horizontal and $\alpha^{\widetilde{\xi}}(s)=\alpha^\xi(s+z)$.
Moreover note that $\widetilde{\xi}(a)=[l^{\mathcal{X}}_q(z)]^{-1}\circ l^{\mathcal{X}}_q(a+z)$ and $\widetilde{\xi}(b)=[l^{\mathcal{X}}_q(z)]^{-1}\circ l^{\mathcal{X}}_q(b+z)$.
We claim that

$$\widetilde{\xi}(a)=l^{\mathcal{X}}_q(a)\quad\text{and}\quad\widetilde{\xi}(b)=l^{\mathcal{X}}_q(b). \tag{5.5}$$

In fact, consider the two curves $l^{\mathcal{X}}_q(s)$ and $\widetilde{l}^{\mathcal{X}}_q(s):=[l^{\mathcal{X}}_q(z)]^{-1}\circ l^{\mathcal{X}}_q(s+z)$: both the curves start from the origin since $\widetilde{l}^{\mathcal{X}}_q(0)=[l^{\mathcal{X}}_q(z)]^{-1}\circ l^{\mathcal{X}}_q(z)=0=l^{\mathcal{X}}_q(0)$. Moreover they have the same horizontal velocity since $\alpha^{l^{\mathcal{X}}_q}(s)=q$ (by definition) and $\alpha^{\widetilde{l}^{\mathcal{X}}_q}(s)=\alpha^{l^{\mathcal{X}}_q}(s+z)=q$ (by Lemma 2.3). Hence by standard uniqueness for ODEs with smooth data, the two curves coincide and in particular (5.5) holds. This implies that for each $\xi\in\mathcal{B}^q_{a+z,b+z}$, the curve $\widetilde{\xi}\in\mathcal{B}^q_{a,b}$. Then, by the change of variable $\widetilde{s}=s-z$,

$$\begin{aligned}\tau_z\,\mu_q([a,b),\omega)&=\inf_{\widetilde{\xi}\in\mathcal{B}^q_{a,b}}\int_{a+z}^{b+z}L([l^{\mathcal{X}}_q(z)]\circ\widetilde{\xi}(s-z),\alpha^{\widetilde{\xi}}(s-z),\omega)ds\\&=\inf_{\widetilde{\xi}\in\mathcal{B}^q_{a,b}}\int_a^b L([l^{\mathcal{X}}_q(z)]\circ\widetilde{\xi}(s),\alpha^{\widetilde{\xi}}(s),\omega)ds.\end{aligned}$$

We set $z_q=l^{\mathcal{X}}_q(z)$ and use assumption **(L4)** to conclude

$$\tau_z\,\mu_q([a,b),\omega)=\inf_{\xi\in\mathcal{B}^q_{a,b}}\int_a^b L(\xi(s),\alpha^\xi(s),\tau_{z_q}\omega)ds=\mu_q([a,b),\tau_{z_q}\omega).$$

□

Lemma 5.2 (Subadditivity)**.** *For $n \in \mathbb{N}$, there holds*

$$\mu_q([0,n),\omega) \leq \sum_{k=1}^{n} \mu_q(k-1+[0,1),\omega).$$

Proof. If $\xi_k \in \mathcal{B}^q_{k-1,k}$, then we can construct a continuous horizontal curve ξ in $\mathcal{B}^q_{0,n}$, such that $\xi(s) = \xi_k(s)$ for $s \in [k-1,k]$, $k \in \{1,\dots,n\}$. The claim follows from the definition of the infimum. □

Under assumptions **(L1)-(L4)** the Subadditive Ergodic Theorem applies to the process defined in (5.3). We will use it in the following form, which is taken from [20, Prop. 1], based on Akcoglu and Krengel's theorem, [1]. We state it for one dimension. First, we get the existence of a limit which may still depend on ω and then we use ergodicity to show independence of ω. From [20] we recall:

Definition 5.4. *We denote by $\mathcal{U}_0$ the family of all bounded measurable subsets of $\mathbb{R}$. For $A \in \mathcal{U}_0$, its Lebesgue measure is $|A|$. We denote by $\mathcal{M}$ the family of subadditive functions $m : \mathcal{U}_0 \to \mathbb{R}$ such that, for some $c > 0$, there holds*

$$0 \leq m(A) \leq c|A| \qquad \forall A \in \mathcal{U}_0.$$

Theorem 5.1 (**Subadditive Ergodic Theorem, [20]**)**.** *Let $\mu : \Omega \to \mathcal{M}$ be a subadditive process. If μ and $\tau_x\mu$ have the same law for every $x \in \mathbb{Z}$, then there exists a set of full measure Ω' and a measurable function ϕ such that on Ω'*

$$\lim_{t\to\infty} t^{-1}|I|^{-1}\mu(\omega)(tI) = \phi(\omega)$$

for every interval I, where $|I|$ denotes its length.

We look at the points $y \in V_0 \subset \mathbb{R}^N$. Let us recall from Definition 2.1 that we write $y = (y^1, y^2) \in \mathbb{R}^m \times \mathbb{R}^{N-m}$ and $y^1 = \pi_m(y)$. If $y \in V_0$ then $y^2 = y^2(y^1) \in \mathbb{R}^{N-m}$ where $y^2(\cdot)$ is a $(N-m)$-dimensional valued function associated to the vector fields. E.g. in the case of the n-dimensional Heisenberg group $N = 2n+1$ and $m = 2n$ so $y^2 = 0 \in \mathbb{R}$, for all $y^1 \in \mathbb{R}^{2n}$. (See [10, Lemma 2.2] for more details).
We are now ready to give a first pointwise convergence result. Note that, differently from the Euclidean case, the following theorem gives the asymptotic behaviour of $L^\varepsilon(0,y,t,\omega)$ only under the additional $(N-m)$-dimensional constraint expressed by $y \in V_0$.

Theorem 5.2. *Under assumptions* **(L1)-(L4)**, *for each $t > 0$ and $y \in V_0$ fixed,*

1. *The following limit exists a.s. in ω*

$$L^\varepsilon(y,0,t,\omega) \longrightarrow^{\varepsilon\to 0^+} t\,\overline{\mu}\left(\frac{y^1}{t},\omega\right), \tag{5.6}$$

   where $\overline{\mu} : \mathbb{R}^m \times \Omega \to \mathbb{R}$ is a measurable function.

2. *The limit value $\overline{\mu}$ in* (5.6) *is constant in ω.*

Proof. We first prove part 1 by applying the Subadditive Ergodic Theorem 5.1. To this end let us observe that for q fixed, μ_q belongs to $\mathcal{M}$. Actually by Remark 3.4 we have $\mu_q(I,\omega) \geq 0$ for any interval I. On the other hand, since $l_q^\chi \in \mathcal{B}_{a,b}^q$, we have

$$\mu_q([a,b),\omega) \leq \int_a^b L(l_q^\chi(s), \alpha^{l_q^\chi}(s), \omega)ds \leq C_1(1+|q|^\lambda)(b-a),$$

where the last inequality is due to assumption **(L2)** and to the relation $\alpha^{l_q^\chi}(s) = q$. Hence the Subadditive Ergodic Theorem 5.1 implies

$$\frac{1}{\tau}\mu_q([0,\tau),\omega) \longrightarrow^{\tau\to+\infty} \overline{\mu}(q,\omega),\ a.s.\ \omega \in \Omega. \tag{5.7}$$

Note that the definition of μ_q involves only a one-dimensional subgroup of translations, $\{\tau_{\ell_q^\chi(z)}\}_{z\in\mathbb{Z}}$, the subgroup that leaves invariant the $\mathcal{X}$-line with direction q, passing through the origin.

Now we rewrite the functional $L^\varepsilon(y,0,t,\omega)$ defined by (3.4) in terms of $\mu_q\big([a,b),\omega\big)$: let us prove

$$L^\varepsilon(y,0,t,\omega) = \varepsilon\mu_{\frac{y^1}{t}}([0,\varepsilon^{-1}t),\omega).$$

For any $\xi \in \mathcal{A}_{0,y}^t$, we define $\widetilde{\xi}(s) := \delta_{1/\varepsilon}(\xi(s))$. Using Lemma 2.3, part (iii)

$$L^\varepsilon(y,0,t,\omega) = \inf_{\widetilde{\xi}\in\mathcal{A}_{0,y_\varepsilon}^t} \int_0^t L(\widetilde{\xi}(s), \varepsilon\alpha^{\widetilde{\xi}}(s), \omega)ds,$$

where $y_\varepsilon = \delta_{1/\varepsilon}(y)$. By the change of variable $\widetilde{s} = s/\varepsilon$ (which we call again s) the previous identity becomes

$$L^\varepsilon(y,0,t,\omega) = \varepsilon \inf_{\widetilde{\xi}\in\mathcal{A}_{0,y_\varepsilon}^{t/\varepsilon}} \int_0^{t/\varepsilon} L(\widetilde{\xi}(\varepsilon s), \varepsilon\alpha^{\widetilde{\xi}}(\varepsilon s), \omega)ds.$$

Take now $\eta(s) := \widetilde{\xi}(\varepsilon\, s)$ and note that by Lemma 2.3 $\eta(s)$ is still a horizontal curve with $\alpha^\eta(s) = \varepsilon\,\alpha^{\widetilde{\xi}}(\varepsilon s)$, $\eta(0) = \delta_{1/\varepsilon}(y)$ and $\eta(t/\varepsilon) = 0$. Then

$$L^\varepsilon(y,0,t,\omega) = \varepsilon \inf_{\eta\in\mathcal{A}_{0,y_\varepsilon}^{t/\varepsilon}} \int_0^{t/\varepsilon} L(\eta(s), \alpha^\eta(s), \omega)\, ds. \tag{5.8}$$

Now fix $t > 0$, $y \in V_0$ and $y^1 = \pi_m(y)$. To use the convergence result in (5.7) it remains to show that

$$\mathcal{A}_{0,y_\varepsilon}^{t/\varepsilon} = \mathcal{B}_{a,b}^q \text{ with } q = \frac{y^1}{t},\ a = 0 \text{ and } b = \frac{t}{\varepsilon}.$$

For this purpose, consider the $\mathcal{X}$-line joining 0 to y with constant horizontal velocity $q_i = \frac{y_i}{t}$ for $i = 1,\dots,m$, i.e. $l_q^{\mathcal{X}}$ is the unique solution of

$$\dot{l}_q^{\mathcal{X}}(s) = \sum_{i=1}^m \frac{y_i}{t} X_i(l_q^{\mathcal{X}}(s)), \quad l_q^{\mathcal{X}}(0) = 0,$$

(recall that $l_q^{\mathcal{X}}(t) = y$).
CLAIM: for all constant $C > 0$

$$l_q^{\mathcal{X}}(Ct) = \delta_C\big(l_q^{\mathcal{X}}(t)\big) = \delta_C(y). \tag{5.9}$$

To prove claim (5.9), let us introduce the two curves $l_1(s) := l_q^{\mathcal{X}}(Cs)$ and $l_2(s) := \delta_C(l_q^{\mathcal{X}}(s))$. Note that by Lemma 2.3, parts (ii) and (iii), we have

$$l_1(0) = l_2(0) = 0, \quad \alpha^{l_1}(s) = C\alpha^l = C\frac{y^1}{t} \quad \text{and} \quad \alpha^{l_2}(s) = C\alpha^l = C\frac{y^1}{t}.$$

This means that $l_1(\cdot)$ and $l_2(\cdot)$ both solve the ODE problem

$$\dot{x}(s) = \sum_{i=1}^{m} C\frac{y_i^1}{t} X_i(x(s)), \qquad x(0) = 0.$$

By standard uniqueness for ODEs with smooth data, we deduce $l_1(s) = l_2(s)$. This implies in particular $l_1(t) = l_2(t)$ which gives (5.9). Note that here is crucial that the horizontal velocity of the two curves l_1 and l_2 is constant in time.

The claim (5.9) implies that $\mathcal{A}_{0,y_\varepsilon}^{t/\varepsilon} = \mathcal{B}_{0,t/\varepsilon}^{q}$ with $q = \frac{y^1}{t}$, thus equation (5.8) gives

$$L^\varepsilon(y, 0, t, \omega) = \varepsilon\mu_{\frac{y^1}{t}}([0, t/\varepsilon), \omega). \tag{5.10}$$

For $t > 0$, $y \in V_0$, $y^1 = \pi_m(y)$ and $q = \frac{y^1}{t}$ fixed, we can rewrite (5.7) as

$$L^\varepsilon(y, 0, t, \omega) = \varepsilon\mu_{\frac{y^1}{t}}([0, t/\varepsilon), \omega) \longrightarrow^{\varepsilon\to 0^+} t\overline{\mu}\left(\frac{y^1}{t}, \omega\right), \ a.s.\ \omega \in \Omega. \tag{5.11}$$

We now prove part 2: we show the independence of $\overline{\mu}$ from ω. Fix $z \in \mathbb{R}^N$ and define $l_q^z(s) := -z \circ l_q^{\mathcal{X}}(s)$, then by Lemma 2.3 this is still an $\mathcal{X}$-line. We have by stationarity of the coefficients

$$t\overline{\mu}\left(\frac{y^1}{t}, \tau_z(\omega)\right) = \lim_{\varepsilon\to 0^+} \varepsilon\mu_{\frac{y^1}{t}}([0, t/\varepsilon), \tau_z(\omega)),$$

and by (5.3) and stationarity

$$\begin{aligned}
\mu_{\frac{y^1}{t}}([0, t/\varepsilon), \tau_z(\omega)) &= \inf_{\xi\in\mathcal{B}_{0,t/\varepsilon}^q} \int_0^{t/\varepsilon} L(\xi(s), \alpha^\xi(s), \tau_z(\omega))ds \\
&= \inf_{\xi\in\mathcal{B}_{0,t/\varepsilon}^q} \int_0^{t/\varepsilon} L(-z\circ\xi(s), \alpha^\xi(s), \omega)ds \\
&= \inf_{\overline{\xi}\in\mathcal{B}_{0,t/\varepsilon}^{q,z}} \int_0^{t/\varepsilon} L(\overline{\xi}(s), \alpha^{\overline{\xi}}(s), \omega)ds
\end{aligned}$$

where

$$\mathcal{B}_{0,t/\varepsilon}^{q,z} :=$$
$$\left\{\overline{\xi} : [a,b] \to \mathbb{R}^N \,|\overline{\xi} \in W^{1,+\infty}\big((a,b)\big) \text{ horiz}, \overline{\xi}(a) = -z\circ l_q^{\mathcal{X}}(a),\ \overline{\xi}(b) = -z\circ l_q^{\mathcal{X}}(b)\right\}.$$

We have to show

$$\lim_{\varepsilon\to 0^+} \varepsilon \inf_{\xi\in\mathcal{B}^q_{0,t/\varepsilon}} \int_0^{t/\varepsilon} L(\xi(s),\alpha^{\xi}(s),\omega)ds = \lim_{\varepsilon\to 0^+} \varepsilon \inf_{\bar{\xi}\in\mathcal{B}^{q,z}_{0,t/\varepsilon}} \int_0^{t/\varepsilon} L(\bar{\xi}(s),\alpha^{\bar{\xi}}(s),\omega)ds \tag{5.12}$$

for all $z\in\mathbb{R}^N$, then $\overline{\mu}\left(\frac{y^1}{t},\tau_z(\omega)\right) = \overline{\mu}\left(\frac{y^1}{t},\omega\right)$, so by ergodicity w.r.t to the group action $\overline{\mu}\left(\frac{y^1}{t},\omega\right)$ does not depend on ω.
We show that both infima have the same limit by connecting the endpoints $x_1 := l^{\mathcal{X}}_q(a)$ to $y_1 := -z\circ l^{\mathcal{X}}_q(a)$, $x_2 := l^{\mathcal{X}}_q(b)$ to $y_2 := -z\circ l^{\mathcal{X}}_q(b)$ by geodesics of length of order $C(q)\|z\|_{CC}$. Indeed, by Lemma 4.6 the difference of the cost disappears in the limit $\varepsilon\to 0$. (See Figure 1.)
This means any path in $\mathcal{B}^{q,z}_{0,t/\varepsilon}$ can be made into a path in $\mathcal{B}^q_{0,t/\varepsilon}$ by paying a cost of order $|z|$ (for a similar argument, we refer the reader also to the proof of Lemma 6.3). This extra cost vanishes in the limit after multiplication by ε. □

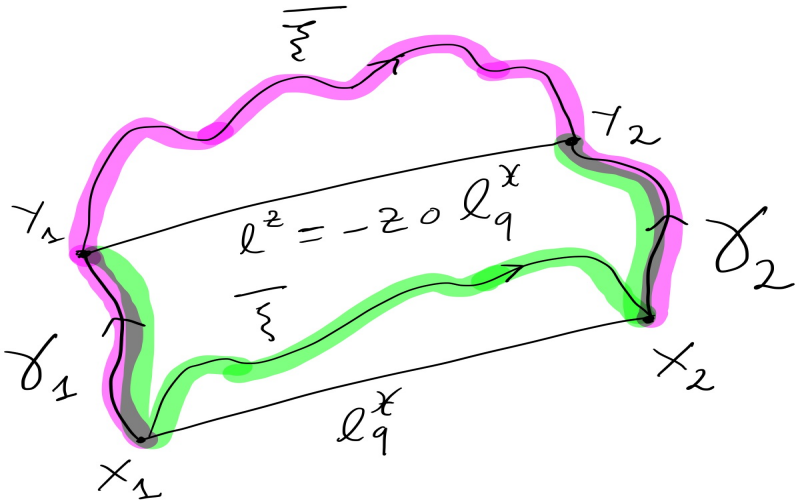

Figure 1: In the picture are drawn two $\mathcal{X}$-lines with constant horizontal velocity q connecting respectively x_1 with x_2 and $y_1 = -z\circ x_1$ with $y_2 = -z\circ x_2$. Then ξ and $\bar{\xi}$ are respectively admissible curves touching at the two couple of points.

Remark 5.3. *Note that the convergence result* (5.6) *means that, for each* $t>0$ *and* $y\in V_0$ *fixed, there exists* $\Omega^{t,y}\subset\Omega$ *with* $\mathbb{P}(\Omega^{t,y})=1$ *such that* $L^{\varepsilon}(y,0,t,\omega)\to t\overline{\mu}\left(\frac{\pi_m(y)}{t}\right)$, *for all* $\omega\in\Omega^{t,y}$. *This convergence result is enough to define the effective Lagrangian but it is still too weak to obtain the convergence of the solutions of the homogenization problem.*

Definition 5.5 (Effective Lagrangian)**.** *We define the effective Lagrangian* $\overline{L}:\mathbb{R}^m\to\mathbb{R}$ *as*

$$\overline{L}(q) := \overline{\mu}(q) = \lim_{\varepsilon\to 0^+} L^{\varepsilon}\big((q,y^2_q),0,1,\omega\big), \tag{5.13}$$

where the point $y_q^2 \in \mathbb{R}^{N-m}$ *is uniquely determined by* $y^1 = q \in \mathbb{R}^m$ *for all points* $(q, y_q^2) \in V_0 \subset \mathbb{R}^N$.

Example 5.1 (Heisenberg group)**.** *In the 1-dimensional Heisenberg group there holds:* $\overline{L}(q_1, q_2) := \lim_{\varepsilon \to 0^+} L^\varepsilon\big((q_1, q_2, 0), 0, 1, \omega\big)$.

Using the definition of effective Lagrangian introduced in (5.13) we can rewrite the limit in Theorem 5.2 as follows: for each $y \in V_0$ and $t > 0$ fixed, there exists a set $\Omega^{t,y} \subset \Omega$ with $\mathbb{P}(\Omega^{t,y}) = 1$ such that

$$\lim_{\varepsilon \to 0^+} L^\varepsilon(y, 0, t, \omega) = t\overline{L}\left(\frac{\pi_m(y)}{t}\right), \quad \forall\, \omega \in \Omega^{t,y}. \tag{5.14}$$

Next we want to derive the local uniform convergence for $L^\varepsilon(x, y, t, \omega)$ under the constraint $y \in V_x$. The following proof is a simple adaptation of the ideas developed by Souganidis in [37] and later by the same author and co-authors in [4, 5, 6, 7]. The main difference is that we work directly with the functional $L^\varepsilon(x, y, t, \omega)$ and not with the solutions $u^\varepsilon(t, x)$. This will guarantee in once both the uniform convergence in y (essential to pass to the infimum in the limit) and the uniform convergence in x and t (that will allow to apply our approximation argument in Section 6).

Theorem 5.3. *Under assumptions* **(L1)-(L4)** *and* **(L6)** *and the additional constraint* $x \in V_y$*, we have that*

$$\lim_{\varepsilon \to 0^+} L^\varepsilon(x, y, t, \omega) = t\overline{L}\left(\frac{\pi_m(-y \circ x)}{t}\right) = t\overline{L}\left(\frac{\pi_m(x) - \pi_m(y)}{t}\right) \tag{5.15}$$

locally uniformly in x, y, t *and a.s.* ω*, where* $\overline{L}$ *is the effective Lagrangian defined by* (5.13).

Proof. We first show that

$$L^\varepsilon(x, y, t, \omega) = L^\varepsilon\left(-y \circ x, 0, t, \tau_{\delta_{1/\varepsilon}(y)}(\omega)\right). \tag{5.16}$$

Note that $x \in V_y$ if and only if $-y \circ x \in V_0$ (recall that $y^{-1} = -y$ in exponential coordinates). To prove (5.16), for each $\xi \in \mathcal{A}^t_{y,x}$ we define $\eta(s) := -y \circ \xi(s)$. By Lemma 2.3-(i) we have $\alpha^\eta(s) = \alpha^\xi(s)$, $\eta(0) = -y \circ x$, $\eta(t) = 0$, hence

$$\begin{aligned}
L^\varepsilon(x, y, t, \omega) \quad &= \inf_{\mathcal{A}^t_{0,-y\circ x}} \int_0^t L\big(\delta_{1/\varepsilon}(y \circ \eta(s)), \alpha^\eta(s), \omega\big) ds \\
&= \inf_{\mathcal{A}^t_{0,-y\circ x}} \int_0^t L\big(\delta_{1/\varepsilon}(y) \circ \delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \omega\big) ds \\
&= \inf_{\mathcal{A}^t_{0,-y\circ x}} \int_0^t L\big(\delta_{1/\varepsilon}(\eta(s)), \alpha^\eta(s), \tau_{\delta_{1/\varepsilon}(y)}\omega\big) ds \\
&= L^\varepsilon\left(-y \circ x, 0, t, \tau_{\delta_{\frac{1}{\varepsilon}}(y)}(\omega)\right),
\end{aligned}$$

where we have used property **(L4)**.
By combining the estimates found in Section 4 with Egoroff's Theorem and the Ergodic Theorem we can conclude. Since the argument is standard and has

been used already in several papers, then we recall only the main steps.
Since $L^\varepsilon(x,y,t,\omega)$ are equi-uniformly continuous in $t>0$ and $x,y\in\mathbb{R}^N$ (see Theorem 4.1), using the density of $\mathbb{Q}$ in $\mathbb{R}$ we can restrict our attention only to points of the form $t_z\in(0,+\infty)\cap\mathbb{Q}=\mathbb{Q}^+$, $x_z\in\mathbb{Q}^N$ and $y_z\in\mathbb{Q}^N$. We then define the following set:

$$\Omega_0 := \bigcap_{t_z\in\mathbb{Q}^+, x_z\in\mathbb{Q}^N,\, y_z\in\mathbb{Q}^N} \Omega^{t_z}_{x_z,y_z},$$

where $\Omega^{t_z}_{x_z,y_z}$ is a set of full measure such that

$$\lim_{\varepsilon\to 0^+} L^\varepsilon(-y_z\circ x_z,0,t_z,\omega) = t_z\,\overline{L}\left(\frac{\pi_m(-y_z\circ x_z)}{t_z}\right).$$

Note that Ω_0 does not depend anymore on t, x and y and $\mathbb{P}(\Omega_0)=1$.
Using the structure of Carnot group in exponential coordinates that implies $\pi_m(-y\circ x)=\pi_m(x)-\pi_m(y)$, since by (5.14) $-y\circ x\in\mathbb{Q}^N$, we know that,

$$\lim_{\varepsilon\to 0^+} L^\varepsilon(-y_z\circ x_z,0,t_z,\omega) = t_z\,\overline{L}\left(\frac{\pi_m(x_z)-\pi_m(y_z)}{t_z}\right), \quad \text{for all } \omega\in\Omega_0.$$

Applying Egoroff Theorem, we find a "very big" subset of Ω_0 where the convergence is uniform in ω. More precisely for any fixed $\delta>0$, there exists $A_\delta\subset\Omega_0$ such that $\mathbb{P}(\Omega_0\setminus A_\delta)\le\delta$ (i.e. $\mathbb{P}(A_\delta)=1-\delta$) and

$$\lim_{\varepsilon\to 0^+} L^\varepsilon(-y_z\circ x_z,0,t_z,\omega) = t_z\,\overline{L}\left(\frac{\pi_m(x_z)-\pi_m(y_z)}{t_z}\right),$$

uniformly for all t_z, x_z, y_z and all $\omega\in A_\delta$.
To conclude one can use the Ergodic Theorem to show that with very high probability $\tau_{\delta_{1/\varepsilon}(y)}(\omega)\in A_\delta$. The application of the Ergodic Theorem is quite technical, so we refer to Lemma 5.1. in [5] for the detailed argument. We just like to remark that by Lemma 2.2 one can easily replace the Euclidean ball with the homogeneous ball (and the reverse), up to consider a different power for the radius which depends only on the step of the Carnot group.
This argument together with the estimates in Section 4, where we found a uniform modulus of continuity depending only on the assumption on H and on the Carnot group (see Theorem 4.1) and relation (5.16) conclude the proof. □

Adding $g(y)$ on both sides of (5.15) and using the local uniform convergence w.r.t. y and Proposition 4.2, we deduce the following convergence for the infimum.

Corollary 5.1. *Under the assumptions of Theorem 5.3 and assuming that* $g:\mathbb{R}^N\to\mathbb{R}$ *satisfies* (4.16)*, then*

$$\lim_{\varepsilon\to 0^+}\inf_{y\in V_x}\left[g(y)+L^\varepsilon(x,y,t,\omega)\right] = \inf_{y\in V_x}\left[g(y)+t\overline{L}\left(\frac{\pi_m(x)-\pi_m(y)}{t}\right)\right]. \quad (5.17)$$

Note that the right-hand side in (5.17) coincides with the Hopf-Lax formula introduced in [8, Theorem 1.1]. Then, whenever the initial condition satisfies the additional assumption $g(x)\ge g(\pi_m(x))$ for all $x\in\mathbb{R}^N$, the right-hand side is

the unique viscosity solution of the associated Hamilton-Jacobi Cauchy problem (defining $\overline{H} = \overline{L}^*$ and proving convexity for $\overline{L}$ see Section 7) . Unfortunately in general $v_\varepsilon(t,x) = \inf_{y\in V_x}[g(y) + L^\varepsilon(x,y,t,\omega)]$ does not solve the ε-problem (3.1). Then it is crucial to get rid of the additional constraint $y \in V_x$. For this purpose, in Section 6 we will introduce a novel approximation argument, by using a suitable construction by $\mathcal{X}$-lines.

We conclude the section investigating some properties for the effective Lagrangian that will be used later.

Lemma 5.3. *For any $y = (y^1, y^2) \in V_0$, we have*

$$\inf_\xi \int_0^t |\alpha^\xi(s)|ds \geq |y^1|,$$

where the infimum is taken over all the horizontal curves $\xi(s)$ such that $\xi(0) = (y^1, y^2)$ and $\xi(t) = 0$.

Proof. Given any horizontal curve ξ such that $\xi(0) = (y^1, y^2) \in \mathbb{R}^N$, $\xi(t) = 0 \in \mathbb{R}^N$, we define $\eta : [0,t] \to \mathbb{R}^m$ as $\eta(s) := \pi_m(\xi(s))$. Then $\eta(0) = y^1 \in \mathbb{R}^m$, $\eta(t) = 0 \in \mathbb{R}^m$. Moreover, from the structure of σ (see (2.3)), we have $\dot\eta(s) = (\dot\xi_1(s), \dots, \dot\xi_m(s)) = \alpha^\xi(s)$.
Then, since η are curves in the Euclidean $\mathbb{R}^m$ joining y^1 to 0 at time t,

$$\int_0^t |\alpha^\xi(s)|ds \geq \inf_\eta \int_0^t |\dot\eta(s)|ds = |y^1|,$$

and we can conclude taking the infimum on the left-hand side term. □

Proposition 5.1. *$\overline{L}(q)$ is continuous and superlinear in q, i.e.*

$$\overline{L}(q) \geq C_1^{-1}(|q|^\lambda - 1) \tag{5.18}$$

where C_1 and λ are the constants introduced in **(L2)**.

Proof. The continuity follows from the uniform convergence of L^ε in (5.13). For each $q \in \mathbb{R}^m$, take $y^1 = q$, $y = (q, y^2) \in V_0$, $t = 1$ and $x = 0$.

$$L^\varepsilon(0,y,1,\omega) = \inf_{\xi\in\mathcal{A}^t_{y,0}} \int_0^1 L(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega)ds.$$

From assumption **(L2)**, Jensen's inequality and Lemma 5.3, we get

$$\begin{aligned} L^\varepsilon(0,y,1,\omega) &\geq C_1^{-1} \inf_{\xi\in\mathcal{A}^1_{y,0}} \left(\int_0^1 |\alpha^\xi(s)|^\lambda ds\right) - C_1^{-1} \\ &\geq C_1^{-1} \inf_{\xi\in\mathcal{A}^1_{y,0}} \left(\int_0^1 |\alpha^\xi(s)|ds\right)^\lambda - C_1^{-1} \\ &\geq C_1^{-1}|q|^\lambda - C_1^{-1}, \end{aligned}$$

which implies (5.18), passing to the limit as $\varepsilon \to 0^+$. □

6 Approximation by $\mathcal{X}$-lines and convergence of the variational problem.

To remove the constraint $y \in V_x$ the idea is to apply Theorem 5.3 to suitable step-$\mathcal{X}$-lines, i.e. horizontal curves whose horizontal velocity is step-constant w.r.t. the given vector fields. More precisely we want to approximate the horizontal velocity $\alpha(t) \in \mathbb{R}^m$ in L^1 by step-constant functions. (Recall that if two horizontal velocities are close in L^1-norm then the a associated horizontal curves are close in L^∞-norm, see Lemma 2.4.)

We will treat the liminf and the limsup separately. Both are treated in the spirit as one would do for the Γ-liminf and the Γ-limsup for integral functionals. One of the technical difficulties here is how to approximate limits of a sequence of minimizing paths by $\mathcal{X}$-lines. Due to the fast oscillations of our integrands in ξ, this is not straightforward, we refer to the discussion in [24]. As we cannot assume that our limit paths are smooth but only belong to some Sobolev space, we have to work with Lebesgue points of the horizontal velocity. Here this is more subtle than in the Euclidean case.

Lemma 6.1. *Suppose $\dot\xi(s) = \alpha_1(s)X_1(\xi(s)) + \ldots + \alpha_m(s)X_m(\xi(s))$ and $t_0 \in \mathbb{R}$ is a Lebesgue point for $\alpha_1, \ldots, \alpha_m$, that means*

$$\lim_{\delta\to 0} \max_{i=1,\ldots,m} \delta^{-1} \int_{t_0-\delta}^{t_0+\delta} |\alpha_i(s) - \alpha_i(t_0)| ds = 0. \tag{6.1}$$

Consider the $\mathcal{X}$-line $\ell(s) := l^{\alpha(t_0)}(s)$ i.e.

$$\dot\ell(s) = \alpha_1(t_0)X_1(\ell(s)) + \ldots + \alpha_m(t_0)X_m(\ell(s)), \quad \ell(t_0) = \xi(t_0).$$

Then for any $\varepsilon > 0$ there exists $\delta_0 > 0$ such that for all $\delta < \delta_0$

$$\sup_{[t_0-\delta,t_0+\delta]} d_{CC}(\xi(s), \ell(s)) < \varepsilon \left(\delta + \int_{t_0-\delta}^{t_0+\delta} |\alpha(s)|\, d\, s \right). \tag{6.2}$$

Proof. We use the exponential representation of the Carnot group. Moreover, since the Carnot-Carathéodory distance is locally equivalent to the homogeneous distance, we estimate $\| - \ell(s) \circ \xi(s)\|_h$.

Case 1: the Heisenberg group $\mathbb{H}$. We prove the statement in the Heisenberg group $\mathbb{H}$ by explicit computations. W.l.o.g. we assume $t_0 = 0$; for the first two coordinates we have

$$\xi_i(t) = \ell_i(0) + t\alpha_i(0) + t \underbrace{\left(\frac{1}{t} \int_0^t (\alpha_i(s) - \alpha_i(0)) ds \right)}_{:= r_i(t)} = \ell_i(t) + t\, r_i(t).$$

For the third coordinate we have

$$\begin{aligned} \ell_3(t) &= \frac{t}{2} \left(\ell_1(0)\alpha_2(0) - \ell_2(0)\alpha_1(0) \right) + \ell_3(0) \\ \xi_3(t) &= \int_0^t \frac{1}{2} \left(\xi_1(s)\alpha_2(s) - \xi_2(s)\alpha_1(s) \right) ds + \ell_3(0). \end{aligned}$$

Writing $\alpha_2(s) = \alpha_2(s) \pm \alpha_2(0)$, we get

$$\begin{aligned}\int_0^t \xi_1(s)\alpha_2(s)ds &= \frac{t^2}{2}\alpha_1(0)\alpha_2(0) + t\ell_1(0)\alpha_2(0) + \ell_1(0)t\, r_2(t) + \\ &\quad + \int_0^t s\, r_1(s)\alpha_2(s)ds + \alpha_1(0)\int_0^t s(\alpha_2(s) - \alpha_2(0))ds,\end{aligned}$$

and since $\int_0^t \xi_2(s)\alpha_1(s)ds$ can be treated similarly, we have

$$\begin{aligned}\xi_3(t) &= \ell_3(t) + \frac{t}{2}\left(\ell_1(0)r_2(t) - \ell_2(0)r_1(t)\right) \\ &\quad + \int_0^t s\big(r_1(s)\alpha_2(s) - r_2(s)\alpha_1(s)\big)ds \\ &\quad + \int_0^t s\left[\alpha_1(0)\big(\alpha_2(s) - \alpha_2(0)\big) - \alpha_2(0)\big(\alpha_1(s) - \alpha_1(0)\big)\right]ds.\end{aligned}$$

Let us denote the last two lines by $R(t)$. Now from

$$(-\ell(t) \circ \xi(t))_3 = \xi_3(t) - \ell_3(t) + \frac{\ell_2(t)\xi_1(t) - \ell_1(t)\xi_2(t)}{2},$$

we get

$$(-\ell \circ \xi)_3 = R(t) + t^2\frac{\alpha_2(0)r_1(t) - \alpha_1(0)r_2(t)}{2}.$$

All error terms can be estimated by $t^2\|\alpha\|_\infty \sup_{[0,t]} \max_{1,2} |r_i|$ but we need an estimate where the constant depends only on $\|\alpha_i\|_{L^1}$.
Note that $|R(t)|$ can be estimated by

$$\left|\int_0^t s\alpha_1(0)\big(\alpha_2(s) - \alpha_2(0)\big)ds\right| \le \int_0^t t^2\frac{|\alpha_1(0)||\alpha_2(s) - \alpha_2(0)|}{t}\,ds \le |\alpha_1(0)||r_2(t)|t^2$$

$$\left|\int_0^t sr_1(s)\big(\alpha_2(s) - \alpha_2(0) + \alpha_2(0)\big)ds\right| \le |\alpha_2(0)|\sup_{[0,t]}|r_1|t^2 + |r_2(t)|\sup_{[0,t]}|r_1|t^2,$$

and similarly for the remaining terms. Denoting by $r(t)$ a term vanishing with $|r_1(t)| + |r_2(t)|$, we have to show that terms of the form $\sqrt{\alpha_i(0)}\; t\, r(t)$ can be estimated in a way which can be summed over a partition of the unit interval. Since for $t \in [0,1]$

$$\sqrt{|\alpha_i(0)|}\; t \le \frac{1}{2}|\alpha_i(0)|t + \frac{1}{2},$$

and

$$|\alpha_i(0)|t \le \int_0^t |\alpha_i(s) - \alpha_i(0)|ds + \int_0^t |\alpha_i(s)|ds = t|r_i(t)| + \int_0^t |\alpha_i(s)|ds,$$

the claim is shown by applying these estimates on both sub-intervals $(t_0 - \delta, t_0)$ and $(t_0, t_0 + \delta)$.

Case 2: general case.

Step 1. As the left-translation leaves the CC-distance between two points invariant, we may assume w.l.o.g. $t_0 = 0$ and $\xi(0) = \ell(0) = 0$.
For a general path $\eta : [0,T] \to \mathbb{R}^m$ we define the 1-variation norm in the

following way: let $\Delta[0,T]$ be the family of all partitions of the interval $[0,T]$. Then

$$|\eta|_{1-var[0,T]} = \sup_{\Delta[0,T]} \sum_{(t_k,t_{k+1})\in\Delta[0,T]} |\eta(t_{k+1}) - \eta(t_k)|.$$

It is easy to see that

$$|\eta|_{1-var[0,T]} = \sup_{\Delta[0,T]} \sum \left| \int_{t_k}^{t_{k+1}} \dot{\eta}(s) ds \right| \leq \int_0^t |\dot{\eta}(s)| ds. \tag{6.3}$$

We define two paths in $\mathbb{R}^m$ as the projection on the first m components of ξ and ℓ and we denote them respectively by η^ξ and η^ℓ. Using the structure given by (2.3), we have that $\dot{\eta}^\xi(t) = \alpha(t)$ and $\dot{\eta}^\ell(t) = \alpha(0)$, then, for δ sufficiently small, (6.3) implies

$$|\eta^\xi - \eta^\ell|_{1-var[0,\delta]} \leq \delta\varepsilon$$

because by assumption $t_0 = 0$ is a Lebesgue point for α.
Step 2. By [29, Proposition 7.63], we have for the signature of the path (i.e. for the difference of all iterated integrals)

$$\sup_{k=1,\dots n-m} \sup_{0<t_1<\delta} \left| \int_0^{t_1} \int_0^{t_2} \dots \int_0^{t_k} d\eta^\xi(s_1) \otimes \dots \otimes d\eta^\xi(s_k) ds_1 \dots ds_k \right.$$
$$\left. - \int_0^{t_1} \int_0^{t_2} \dots \int_0^{t_k} d\eta^\xi(s_1) \otimes \dots \otimes d\eta^\xi(s_k) ds_1 \dots ds_k \right| \delta^{-k} < C\varepsilon.$$

Step 3. The Chen-Strichartz formula, which is a deep generalization of the Baker-Campbell-Hausdorff formula, allows to compute the solution of flows driven by absolutely continuous paths via multiplying the terms appearing in the signature to the corresponding commutators of the vector fields. Adapting the notation in [11, Ch. 2], to the notation used here we have for a path ξ as in (2.4) starting from the origin (in exponential coordinates)

$$\xi = \sum_{k=1}^{r} \sum_{I=\{i_1,\dots,i_k\}} \Lambda_I(\alpha_1, \dots, \alpha_m) X_I.$$

Here

$$X_I := [X_{i_1}, [X_{i_2}, ..., [X_{i_{k1}}, X_{i_k}]...].$$

Moreover for $t_1 < \dots < t_k < \delta$

$$\Lambda_I := \sum_{\sigma\in S_k} \frac{(-1)^{e(\sigma)}}{k^2 \begin{pmatrix} k-1 \\ e(\sigma) \end{pmatrix}} \int_0^{t_1} \int_0^{t_2} \dots \int_0^{\delta} \alpha_{i_1}(t_{i_1}) \dots \alpha(t_{i_k}) dt_{i_1} \dots dt_{i_k},$$

where S_k is the symmetric group of k elements and $e(\sigma)$ is a nonnegative integer depending only on the permutation $\sigma \in S_k$ (see [11, Ch.1]).

Note that all sums are finite as the Lie algebra is nilpotent, and that the projection of the solution on the k-th layer of the graded algebra is a multiple of a k-times iterated integral of the α_i, $i = 1, \dots, m$.

Combining step 2 and step 3, the desired estimate in the homogeneous (and hence Carnot-Caratheodory) distance follows. □

In the following lemma we build a partition using sub-intervals where we can apply the previous lemma up to a set of Lebesgue-measure arbitrarily small.

Lemma 6.2. *For any $\rho > 0$ and $\delta > 0$ there exist $N_1, N_2 \in \mathbb{N}$ and a partition of $[0,1)$ formed by the union of the intervals $I_k = [t_k - \ell_k, t_k + \ell_k)$ for $k = 1, \dots, N_1$ and the intervals $J_k = [t'_k - \ell'_k, t'_k + \ell'_k)$ for $k = 1, \dots, N_2$ such that*

- *$0 < \ell_k < \rho$ for $k = 1, \dots, N_1$,*
- *for $k = 1, \dots, N_1$ and $i = 1, \dots, m$ we have*

$$\int_{t_k - r}^{t_k + r} |\alpha_i(s) - \alpha_i(0)| ds < r\delta, \quad \text{for } 0 < r < \ell_k,$$

- *$\sum_{k=1}^{N_2} |J_k| < \delta$.*

Proof. By the Lebesgue point theorem, there exits a set $\mathcal{N}$ of zero Lebesgue measure such that any $\tau \in [0,1] \setminus \mathcal{N}$ is a joint Lebesgue point of $\alpha_1, \dots, \alpha_m$. By definition of the Lebesgue measure, $\mathcal{N}$ can be covered by a countable union of intervals with total length smaller than δ. For each $\tau \in [0,1] \setminus \mathcal{N}$ there exists a $\rho_\tau > 0$ such that

$$\int_{\tau - r}^{\tau + r} |\alpha_i(s) - \alpha_i(0)| ds < r\delta, \quad \text{for } i = 1, \dots, m \text{ and } 0 < r < \rho_\tau.$$

In this way we obtain an open cover of the compact unit intervals and extract a finite subcover. These finitely many intervals can be ordered according to their center and made into a partition by shortening them, starting from the leftmost center, until the desired partition is obtained. □

In the following lemma we prove the main lower bound for the liminf of L^ε.

Lemma 6.3. *Let us assume that $L(x, q, \omega)$ satisfies assumptions* **(L1)-(L4)** *and* **(L6)**. *Then, locally uniformly in $t > 0$, $x, y \in \mathbb{R}^N$ and a.s. in $\omega \in \Omega$*

$$\liminf_{\varepsilon \to 0^+} L^\varepsilon(x, y, t, \omega) \geq t \inf_{\alpha \in \mathcal{F}^t_{y,x}} \int_0^t \overline{L}(\alpha(s)) ds, \tag{6.4}$$

where $\mathcal{F}^t_{y,x}$ is the set of all the measurable functions $\alpha : [0,t] \to \mathbb{R}^m$ such that the corresponding horizontal curve $\xi^\alpha(s)$ joins y to x in a time t, and $\overline{L}(q)$ is the effective Lagrangian defined by limit (5.13).

Proof. Step 1: For sake of simplicity we assume that $t = 1$ and that for each ε there exists a minimizing curve ξ^ε for $L^\varepsilon(x, y, 1, \omega)$. We observe that, by Corollary 4.1, the sequence $\{\xi^\varepsilon\}_\varepsilon$ is equibounded in $\mathbb{L}^\infty(0,1)$ and in $W^{1,\lambda}(0,1)$. In particular, it is equi-bounded in the Hölder norm with Hölder exponent $\gamma < 1 - \frac{1}{\lambda}$. Hence, by Ascoli theorem and by Sobolev embedding theorem, up to a subsequence, ξ^ε uniformly converges to some Hölder curve $\bar{\xi}$. We claim that $\bar{\xi}$ is horizontal. Actually, by Proposition 4.1, there holds $\|\alpha^{\xi^\varepsilon}\|_{L^\lambda} < C_2$; hence, possibly passing to a subsequence, $\{\alpha^{\xi^\varepsilon}\}$ weakly converges to some $\bar{\alpha}$ in L^λ. Moreover, there holds

$$\xi^\varepsilon(t) = \sum_{i=1}^m \int_0^t \alpha_i^{\xi^\varepsilon}(s) X_i(\xi^\varepsilon(s))\, ds;$$

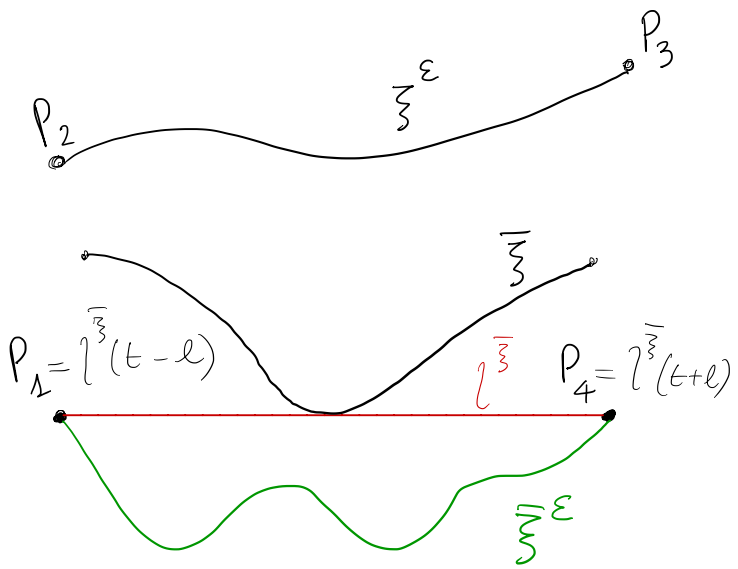

Figure 2: this picture illustrates Steps 3 and 4 of Lemma 6.3.

so, taking into account that ξ^ε uniformly converge to $\bar\xi$, we infer

$$\bar\xi(t) = \sum_{i=1}^{m} \int_0^t \alpha_i^{\bar\xi}(s) X_i(\bar\xi(s))\, ds$$

which means that $\bar\xi$ is horizontal. Finally, smoothing the horizontal velocity $\alpha^{\bar\xi}$, we obtain a family of smooth and horizontal curves uniformly approximating $\bar\xi$. Therefore, since now on, w.l.o.g. we assume that $\bar\xi$ is admissible.

Step 2: Choose a partition of $[0,1]$ in intervals $I_k = [t_k - l_k, t_k + l_k)$ and $J_k = [t'_k - l'_k, t'_k + l'_k)$ as in Lemma 6.2 and error δ for the velocity $\alpha^{\bar\xi}$. Denote by

$$\mathcal{B} := \bigcup_{k=1}^{N_2} J_k$$

the bad set of total length δ. By the a-priori bounds on $\|\alpha\|_{L^1}$ (see Corollary 4.1), assumption **(L2)** and the continuity of $\bar L$ (see Proposition 5.1), there exists $r(\delta) \to 0$ as $\delta \to 0$ such that

$$\begin{aligned} L^\varepsilon(x,y,t,\omega) &= \int_0^1 L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)), \alpha^{\xi^\varepsilon}(s), \omega\right) ds \\ &\geq \int_{[0,1]\setminus\mathcal{B}} L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)), \alpha^{\xi^\varepsilon}(s), \omega\right) ds - r(\delta), \qquad (6.5) \\ \int_0^1 \overline{L}(\alpha^{\bar\xi}(s)) ds &\geq \int_{[0,1]\setminus\mathcal{B}} \overline{L}(\alpha^{\bar\xi}(s)) ds - r(\delta). \end{aligned}$$

Hence we can ignore the bad intervals.
As the number of good intervals N_1 is fixed and finite and ξ^ε uniformly converges

to $\overline{\xi}$, we can choose ε sufficiently small such that

$$\max_{k=1,\dots N_1} \sup_{I_k} |I_k|^{-1} d_{CC}(\xi^\varepsilon(s), \overline{\xi}(s)) < \delta. \tag{6.6}$$

In the interval I_k, consider the constant velocity $\alpha^{\overline{\xi}}(t_k)$. By the continuity of $\overline{L}$, the integral on the right hand side of (6.4) with $\alpha = \alpha^{\overline{\xi}}$, can be approximated by any Riemann sum, and the bad intervals can be ignored:

$$\sum_{k=1}^{N_1} |I_k| \overline{L}\left(\alpha^{\overline{\xi}}(t_k)\right) \to \int_0^1 \overline{L}(\alpha^{\overline{\xi}}(s))ds, \quad \text{as } N_1 \to +\infty. \tag{6.7}$$

Step 3: Let us consider now one "good" interval denoted by $I = [t-\ell, t+\ell]$ for simplicity, and consider the $\mathcal{X}$-line through $\overline{\xi}(t)$ with velocity $\alpha^{\overline{\xi}}(t)$ which we denote by $l^{\overline{\xi}}$ (see Figure 2).

Let us consider a curve $\overline{\xi}^\varepsilon$ with $\overline{\xi}^\varepsilon(t-\ell) = l^{\overline{\xi}}(t-\ell)$ and $\overline{\xi}^\varepsilon(t+\ell) = l^{\overline{\xi}}(t+\ell)$ which is the minimizer of

$$L^\varepsilon\left(l^{\overline{\xi}}(t+\ell), l^{\overline{\xi}}(t-\ell), I, \omega\right) = \int_{t-\ell}^{t+\ell} L\left(\delta_{1/\varepsilon}(\overline{\xi}^\varepsilon(s)), \alpha^{\overline{\xi}^\varepsilon}(s), \omega\right) ds.$$

where $L^\varepsilon(x, y, I, \omega)$ is defined as in (3.4) with the infimum is over the admissible curves with $\xi(t-\ell) = y$ and $\xi(t+\ell) = x$.
From Theorem 5.3, we can choose ε sufficiently small such that

$$\int_{t-\ell}^{t+\ell} L\left(\delta_{1/\varepsilon}(\overline{\xi}^\varepsilon(s)), \alpha^{\overline{\xi}^\varepsilon}(s), \omega\right) ds \quad \geq \quad |I|\overline{L}(\alpha^{l^{\overline{\xi}}}) - |I|\delta. \tag{6.8}$$

This can be done uniformly for all good intervals I_k as their number is already fixed.

Step 4: We now claim that

$$\int_I L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)), \alpha^{\xi^\varepsilon}(s), \omega\right) ds \geq \int_I L\left(\delta_{1/\varepsilon}(\overline{\xi}^\varepsilon(s)), \alpha^{\overline{\xi}^\varepsilon}(s), \omega\right) ds - C\left(|I| + \|\alpha^{\overline{\xi}}\|_{L^1(I)}\right)\delta. \tag{6.9}$$

Let us now prove the claim (6.9): by Lemma 6.1 we know that

$$\sup_I d_{CC}(l^{\overline{\xi}}, \overline{\xi}) < C\left(|I| + \|\alpha^{\overline{\xi}}\|_{L^1(I)}\right)\delta. \tag{6.10}$$

Consider the points

$$P_1 := \overline{\xi}^\varepsilon(t-\ell) = l^{\overline{\xi}}(t-\ell), \quad P_2 := \xi^\varepsilon(t-\ell), \quad P_3 := \xi^\varepsilon(t+\ell) \quad P_4 := \overline{\xi}^\varepsilon(t+\ell) = l^{\overline{\xi}}(t+\ell)$$

(see Figure 2). Then by (6.6), and (6.10)

$$d_{CC}(P_1, P_2) \leq d_{CC}(l^{\overline{\xi}}(t-\ell), \overline{\xi}(t-\ell)) + d_{CC}(\overline{\xi}(t-\ell), \xi^\varepsilon(t-\ell)) \leq C\delta\left(|I| + \|\alpha^{\overline{\xi}}\|_{L^1(I)}\right),$$

and analogously for $d_{CC}(P_3, P_4)$. By Lemmas 4.6 and 4.1 we have

$$L^\varepsilon(P_3, P_2, I, \omega) \geq L^\varepsilon(P_4, P_1, I, \omega) - \delta C\left(|I| + \|\alpha^{\overline{\xi}}\|_{L^1(I)}\right).$$

Since ξ^ε is admissible for $L^\varepsilon(P_3,P_2,I,\omega)$ then

$$L^\varepsilon(P_3,P_2,I,\omega)\le\int_I L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)),\alpha^{\xi^\varepsilon}(s),\omega\right)ds.$$

Combining the last two inequalities, (6.9) is shown.

Step 5:

Since the claim (6.9) is true for any of the good intervals I_k, we can easily conclude. Indeed, using respectively that ξ^ε are minimizer of $L^\varepsilon(x,y,1,\omega)$, (6.5), Definition 2.4 and (6.8)

$$\begin{aligned}
L^\varepsilon(x,y,1,\omega)&=\int_0^1 L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)),\alpha^{\xi^\varepsilon}(s),\omega\right)ds\\
&\ge\sum_{k=1}^{N_1}\int_{I_k} L\left(\delta_{1/\varepsilon}(\xi^\varepsilon(s)),\alpha^{\xi^\varepsilon}(s),\omega\right)ds-r(\delta)\\
&\ge\sum_{k=1}^{N_1}\int_{I_k} L\left(\delta_{1/\varepsilon}(\overline{\xi}_k^\varepsilon(s)),\alpha^{\overline{\xi}_k^\varepsilon}(s),\omega\right)ds-\sum_{k=1}^{N_1}C\delta\left(|I_k|+\|\alpha^{\overline{\xi}}\|_{L^1(I_k)}\right)-r(\delta)\\
&\ge\sum_{k=1}^{N_1}\int_{I_k} L\left(\delta_{1/\varepsilon}(\overline{\xi}_k^\varepsilon(s)),\alpha^{\overline{\xi}_k^\varepsilon}(s),\omega\right)ds-C\delta(1+d_{CC}(x,y))-r(\delta)\\
&\ge\sum_{k=1}^{N_1}|I_k|\overline{L}\left(\alpha^{l^{\overline{\xi}}}(t_k)\right)-C\delta(1+d_{CC}(x,y))-r(\delta)\\
&\ge\int_0^1\overline{L}(\alpha^{\overline{\xi}}(s))ds-r(\delta)-C\delta(1+d_{CC}(x,y))\\
&\longrightarrow\int_0^1\overline{L}(\alpha^{\overline{\xi}}(s))ds,\qquad\text{as }\delta\to0.
\end{aligned}$$

□

In the following lemma we prove the upper bound for the limsup of L^ε.

Lemma 6.4. *Let us assume that $L(x,q,\omega)$ satisfies assumptions* **(L1)-(L4)** *and* **(L6)**. *Then locally uniformly in $t>0$, $x,y\in\mathbb{R}^N$ and a.s. in $\omega\in\Omega$*

$$\limsup_{\varepsilon\to0^+}L^\varepsilon(x,y,t,\omega)\le t\inf_{\alpha\in\mathcal{F}^t_{y,x}}\int_0^t\overline{L}(\alpha(s))ds,\tag{6.11}$$

where $\mathcal{F}^t_{y,x}$ is the set of all the measurable functions $\alpha:[0,t]\to\mathbb{R}^m$ such that the corresponding horizontal curve $\xi^\alpha(s)$ joins y to x in time t and $\overline{L}(q)$ is the effective Lagrangian defined by limit (5.13).

Proof. W.l.o.g. we show the result for $t=1$. Let us choose $\overline{\alpha}\in\mathcal{F}^1_{y,x}$ which realizes the infimum on the right-hand side of (6.11). We assume $\overline{\alpha}$ smooth; in fact we can uniformly approximate $\overline{\alpha}$ by smooth horizontal velocities and use the continuity of $\overline{L}$.

We fix a partition π of the interval $[0,1]$ in n equal length intervals $\left(\frac{i-1}{n},\frac{i}{n}\right)$ and we set

$$\overline{\alpha}^i:=\overline{\alpha}\left(\frac{2i-1}{2n}\right),\quad i=1,\dots,n.$$

We define the step-function

$$\overline{\alpha}^{\pi}(s) := \overline{\alpha}^{i} = \text{constant}, \quad \forall\, s \in \left(\frac{i-1}{n}, \frac{i}{n}\right), \text{ for } i = 1, \dots, n.$$

By Taylor expansion, we know that

$$\|\overline{\alpha} - \alpha^{\pi}\|_{L^1(0,1)} = O\left(\frac{1}{n}\right). \tag{6.12}$$

Note that the constants in the Taylor expansion depend on higher derivatives of $\overline{\alpha}$, but this is fixed throughout the proof, in particular it does not depend on ε. We define the following sequence of points in $\mathbb{R}^N$:

$$z^0 := y, \quad z^i := \zeta^i(1/n),$$

where $\zeta^i : \left[0, \frac{1}{n}\right] \to \mathbb{R}^N$ is the unique $\mathcal{X}$-line starting from z^{i-1} with constant horizontal velocity $\overline{\alpha}^i$ (i.e. $\dot{\zeta}^i(s) = \sum_{j=1}^m \overline{\alpha}^i_j X_j(\zeta^i(s))$, for $s \in \left[0, \frac{1}{n}\right]$, with $\zeta^i(0) = z^{i-1}$). Note that $z^i \in V_{z^{i-1}}$, for all $i = 1, \dots, n$.

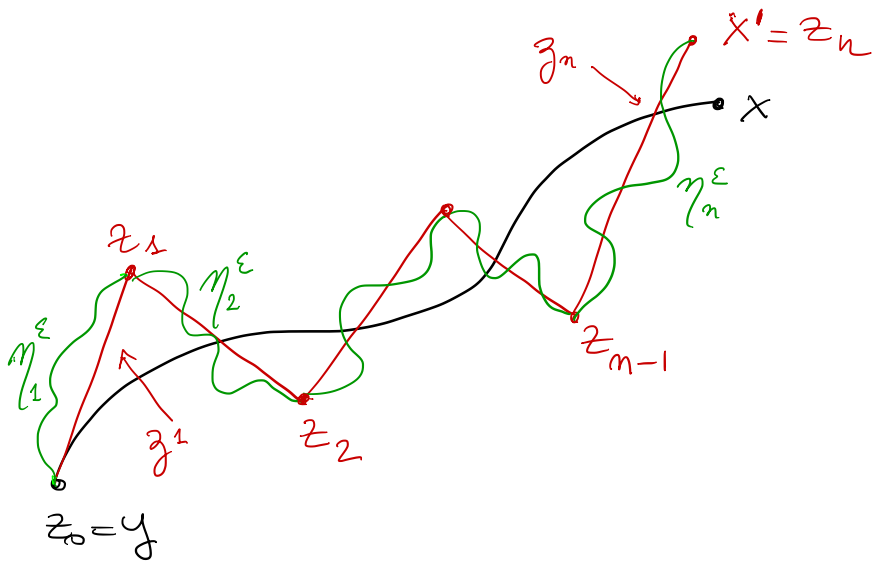

Figure 3: this picture illustrates the arguments of the proof of Lemma 6.4.

Since in general $\mathcal{X}$-lines do not minimize the integral functional between the two points z^{i-1} and z^i, we consider the curves $\eta_i^\varepsilon(s)$ which are minimizers of $L^\varepsilon\left(z^i, z^{i-1}, \frac{1}{n}, \omega\right)$. We look at the two curves (see Picture 3):

$\overline{\xi} \in \mathcal{A}^1_{y,x}$ horizontal curve associated to the horizontal velocity $\overline{\alpha}(s)$,

$\overline{\eta}^\varepsilon \in \mathcal{A}^1_{y,x'}$ horizontal curve defined as union of η_i^ε, for $i = 1, \dots, n$

where $x' := \zeta^n(1/n)$ satisfies, by (6.12) and Lemma 2.4,

$$|x - x'| = O\left(\frac{1}{n}\right). \tag{6.13}$$

Note that $\overline{\eta}^\varepsilon$ is an admissible curve, i.e. $\overline{\eta}^\varepsilon \in \mathcal{A}^1_{y,x'}$, but it may be not a minimizer for $L^\varepsilon(x', y, 1, \omega)$. Moreover the curve $\overline{\eta}^\varepsilon$ depends on $\overline{\alpha}$, on the partition π (i.e. on n) and on ε; nevertheless these dependences do not influence our final estimate.
From (6.13) and the continuity of L^ε in y, uniformly w.r.t. ε (see Theorem 4.1), denoting by $o(1)$ a function which goes to zero if $n \to +\infty$, we get

$$\begin{aligned} L^\varepsilon(x, y, 1, \omega) &= L^\varepsilon(x', y, 1, \omega) + o(1) \\ &\leq \int_0^1 L\left(\delta_{1/\varepsilon}(\overline{\eta}^\varepsilon(s)), \alpha^{\overline{\eta}^\varepsilon}(s), \omega\right) ds + o(1) \\ &= \sum_{i=1}^n \int_0^{1/n} L\left(\delta_{1/\varepsilon}(\eta_i^\varepsilon(s)), \alpha^{\eta_i^\varepsilon}(s), \omega\right) ds + o(1) \\ &= \sum_{i=1}^n L^\varepsilon\left(z^i, z^{i-1}, \frac{1}{n}, \omega\right) + o(1) \end{aligned} \tag{6.14}$$

where we have used the definition of $\overline{\eta}^\varepsilon$ as union of minimisers for each interval of the partition. Now we first choose n big enough that the $o(1)$ term in the last line is smaller than the desired error. In the next step we then choose ε depending on n.
Let us assume for the moment the following claim

$$L^\varepsilon\left(z^i, z^{i-1}, \frac{1}{n}, \omega\right) = \frac{1}{n}\left(\overline{L}(\overline{\alpha}^i) + r(\varepsilon)\right), \tag{6.15}$$

where $r(\varepsilon)$ is a function which goes to zero if $\varepsilon \to 0$. Hence, by (6.14) and (6.15), we get

$$\begin{aligned} L^\varepsilon(x, y, 1, \omega) &\leq \sum_{i=1}^n \frac{1}{n}\left(\overline{L}(\overline{\alpha}^i) + r(\varepsilon)\right) + o(1) \\ &= \sum_{i=1}^n \frac{1}{n}\overline{L}(\overline{\alpha}^i) + o\left(\frac{1}{n}\right) + o(1) \to \int_0^1 \overline{L}(\overline{\alpha}(s))ds \quad \text{as } n \to +\infty. \end{aligned}$$

It remains only to prove (6.15). By stationarity (see also (5.16)) we have

$$L^\varepsilon\left(z^i, z^{i-1}, \frac{1}{n}, \omega\right) = L^\varepsilon\left(-z^{i-1} \circ z^i, 0, \frac{1}{n}, \tau_{\delta_{1/\varepsilon}(z^{i-1})}(\omega)\right).$$

Using the relation between the Euclidean distance and the Carnot-Carathéodory distance and the fact that $\mathcal{X}$-lines are horizontal curves, we get

$$|-z^{i-1} \circ z^i| \leq C d_{CC}(z^{i-1}, z^i) \leq C \int_0^{\frac{1}{n}} |\overline{\alpha}^i| ds = C_i \frac{1}{n},$$

where $C_i = C|\overline{\alpha}^i|$.
Then, up to a constant, we can write $-z^{i-1} \circ z^i = \frac{\overline{z}}{n}$ where $|\overline{z}| \leq 1$ (Euclidean

norm in $\mathbb{R}^N$). Setting $\overline{z}^1 = \pi_m(\overline{z})$ and using identity (5.16)

$$\begin{aligned} L^\varepsilon\left(z^i, z^{i-1}, \frac{1}{n}, \omega\right) &= L^\varepsilon\left(\frac{\overline{z}}{n}, 0, \frac{1}{n}, \tau_{\delta_{1/\varepsilon}(z^{i-1})}(\omega)\right) \\ &= \varepsilon\mu_{(-\overline{z}^1)}\left(\left[0, \frac{1}{\varepsilon n}\right), \tau_{\delta_{1/\varepsilon}(z^{i-1})}(\omega)\right) \\ &= \frac{1}{n}(\varepsilon n)\mu_{\overline{z}^1}\left(\left[0, \frac{1}{\varepsilon n}\right), \tau_{\delta_{1/\varepsilon}(z^{i-1})}(\omega)\right) \\ &= \frac{1}{n}\big(\overline{\mu}(\overline{z}^1) + r(\varepsilon)\big) = \frac{1}{n}\big(\overline{L}(\alpha^i) + r(\varepsilon)\big), \end{aligned}$$

where one can use the same argument as in the proof of Theorem 5.3 to show that $\frac{1}{n}(\varepsilon n)\mu_{\overline{z}^1}([0, \frac{1}{\varepsilon n}), \tau_{\delta_{1/\varepsilon}(z^{i-1})}(\omega)) \approx \frac{1}{n}(\varepsilon n)\mu_{\overline{z}^1}([0, \frac{1}{\varepsilon n}), \omega)$, as $\varepsilon \to 0^+$. □

Combining Lemmas 6.3 and 6.4 and using Proposition 4.2 we are finally able to prove our main convergence result.

Theorem 6.1. *Let us assume that $L(x, q, \omega)$ satisfies assumptions* **(L1)-(L4)** *and* **(L6)**.

1. *Then*

$$\lim_{\varepsilon \to 0^+} L^\varepsilon(x, y, t, \omega) = t \inf_{\alpha \in \mathcal{F}^t_{y,x}} \int_0^t \overline{L}(\alpha(s))ds, \tag{6.16}$$

locally uniformly in $t > 0$ and $x, y \in \mathbb{R}^N$ and almost surely $\omega \in \Omega$, where $\mathcal{F}^t_{y,x}$ is the set of all the measurable functions $\alpha : [0, t] \to \mathbb{R}^m$ such that the corresponding horizontal curve $\xi^\alpha(s)$ joins y to x in time t.

2. *Given any $g : \mathbb{R}^N \to \mathbb{R}$ satisfying* (4.16) *we have*

$$\lim_{\varepsilon \to 0^+} \inf_{y \in \mathbb{R}^N} [g(y) + L^\varepsilon(x, y, t, \omega)] = \inf_{y \in \mathbb{R}^N} \left[g(y) + t \inf_{\alpha \in \mathcal{F}^t_{y,x}} \int_0^t \overline{L}(\alpha(s))ds\right], \tag{6.17}$$

locally uniformly in $t > 0$ and $x \in \mathbb{R}^N$ and almost surely $\omega \in \Omega$.

7 Homogenization for the Hamilton-Jacobi problem

We want to use Theorem 6.1 to derive the convergence of the viscosity solutions of problem (3.1) to the unique solution of the deterministic problem (1.3). Our strategy is to use the Hopf-Lax variational formula from [8].
The key point is the convexity of the effective Lagrangian $\overline{L}(q)$ defined in (5.13). In the Euclidean case this is an easy consequence of the Dynamical Programming Principle but in our degenerate case this strategy fails since it is not possible to find three points related to a convex combination satisfying simultaneously the associated constraints.

Proposition 7.1. *Let us suppose that $L(x, q, \omega)$ satisfies assumptions* **(L1)-(L4)** *and* **(L6)**. *Then $\overline{L}(p)$ defined in* (5.13) *is convex in $\mathbb{R}^m$.*

Proof. For sake of simplicity, we prove the midpoint convexity (which is equivalent to the convexity), i.e. we want to prove

$$\overline{L}\left(\frac{p+q}{2}\right) \leq \frac{1}{2}\overline{L}(p) + \frac{1}{2}\overline{L}(q) \qquad \forall p, q \in \mathbb{R}^m. \tag{7.1}$$

By definition of $\overline{L}$ we have that

$$\begin{aligned} \overline{L}\left(\frac{p+q}{2}\right) &= \lim_{\varepsilon \to 0^+} L^\varepsilon\left(y^{\frac{p+q}{2}}, 0, 1, \omega\right) \\ &= \lim_{\varepsilon \to 0^+} \inf_{\xi \in \mathcal{A}^1_{0, y^{(p+q)/2}}} \int_0^1 L\left(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega\right) ds \end{aligned} \tag{7.2}$$

where $y^{\frac{p+q}{2}} = l^{(p+q)/2}(1)$ and $l^{(p+q)/2}$ is the $\mathcal{X}$-line starting from 0 with horizontal velocity $\frac{p+q}{2}$. We define the curve ξ_n as the horizontal curve with $\xi_n(0) = 0$ and horizontal velocity

$$\alpha^{\xi_n}(s) = \begin{cases} p, & if\ s \in \left[\frac{i-1}{2n}, \frac{i}{2n}\right] \text{ and } i \text{ even}, \\ q, & if\ s \in \left[\frac{i-1}{2n}, \frac{i}{2n}\right] \text{ and } i \text{ odd} \end{cases}$$

for $i = 1, \ldots, 2n$. We call $x_k = \xi_n(\frac{k}{2n})$, $k = 0, \ldots, 2n$ (see Figure 4). We observe that

$$x_i \in V_{x_{i+1}}, \quad \forall\, i = 1, \ldots, 2n. \tag{7.3}$$

We claim that

$$|\xi_n(1) - y^{(p+q)/2}| = O\left(\frac{1}{n}\right), \tag{7.4}$$

where $|\cdot|$ denotes the Euclidean norm.
Assume for the moment that claim (7.4) is true. Then by the uniform continuity of L^ε (see Theorem 4.1) we can deduce

$$L^\varepsilon\big(y^{(p+q)/2}, 0, 1, \omega\big) = L^\varepsilon\big(\xi_n(1), 0, 1, \omega\big) + O\left(\frac{1}{n}\right). \tag{7.5}$$

We consider now the curve ξ_n^ε which is the union of the curves $\xi_{i,n}^\varepsilon$ defined in $\left[\frac{i-1}{2n}, \frac{i}{2n}\right]$ that are the minimizers for $L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right)$. Observe that ξ_n^ε is an admissible curve between 0 and $\xi_n(1)$. Hence

$$\begin{aligned} &L^\varepsilon(\xi_n(1), 0, 1, \omega) \leq \int_0^1 L\left(\delta_{1/\varepsilon}(\xi_n^\varepsilon(s)), \alpha^{\xi_n^\varepsilon}(s), \omega\right) ds \\ &= \sum_{i \text{ odd}} \int_{\frac{i-1}{2n}}^{\frac{i}{2n}} L\left(\delta_{1/\varepsilon}(\xi_{i,n}^\varepsilon(s)), \alpha^{\xi_{i,n}^\varepsilon}(s), \omega\right) ds \\ &+ \sum_{i \text{ even}} \int_{\frac{i-1}{2n}}^{\frac{i}{2n}} L\left(\delta_{1/\varepsilon}(\xi_{i,n}^\varepsilon(s)), \alpha^{\xi_{i,n}^\varepsilon}(s), \omega\right) ds \\ &= \sum_{i \text{ odd}} L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right) + \sum_{i \text{ even}} L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right), \end{aligned} \tag{7.6}$$

where the last identity comes from the definition of $\xi_{i,n}^\varepsilon$.

From (7.3), we can apply (6.15) obtaining

$$L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right) = \begin{cases} \frac{1}{2n}\big(\overline{L}(p) + r(\varepsilon)\big), & \text{if } i \text{ is even,} \\ \frac{1}{2n}\big(\overline{L}(q) + r(\varepsilon)\big), & \text{if } i \text{ is odd} \end{cases} \tag{7.7}$$

where $r(\varepsilon) \to 0$ as $\varepsilon \to 0^+$. By applying (7.2), (7.5),(7.6), (7.7) we have

$$\begin{aligned} \overline{L}\left(\frac{p+q}{2}\right) &= \lim_{\varepsilon\to 0^+} L^\varepsilon\left(y^{\frac{p+q}{2}}, 0, 1, \omega\right) = \lim_{\varepsilon\to 0^+} L^\varepsilon\big(\xi_n(1), 0, 1, \omega\big) + O\left(\frac{1}{n}\right) \\ &\le \lim_{\varepsilon\to 0^+} \left(\sum_{i \text{ odd}} L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right) + \sum_{i \text{ even}} L^\varepsilon\left(x_{i+1}, x_i, \frac{1}{2n}, \omega\right)\right) + O\left(\frac{1}{n}\right) \\ &= \lim_{\varepsilon\to 0^+} \left(\frac{1}{2}\overline{L}(p) + \frac{1}{2}\overline{L}(q) + r(\varepsilon)\right) + O\left(\frac{1}{n}\right). \end{aligned}$$

Passing to the limits, we get (7.1).

Now it remains to prove claim (7.4). First of all we estimate the distance between x_2 and $l^{(p+q)/2}(1/n)$ (recall that $x_2 = \xi_n(2/2n) = \xi_n(1/n)$). For $a = (a_1, \cdots a_m) \in \mathbb{R}^m$, we set $X_a := a_1X_1 + \cdots a_mX_m$. Using the exponential coordinates, we can write

$$x_2 = \exp\left(\frac{1}{2n}X_p\right) \circ \exp\left(\frac{1}{2n}X_q\right)$$

and

$$l^{(p+q)/2}\left(\frac{1}{n}\right) = \exp\left(\frac{1}{n}X_{\frac{p+q}{2}}\right).$$

The Baker-Campbell-Hausdorff formula ([16]) allows to write:

$$x_2 = \exp\left(\frac{1}{2n}X_p\right) \circ \exp\left(\frac{1}{2n}X_q\right) = \sum_{k_1,k_2} \frac{1}{k_1!k_2!}\left(\frac{1}{2n}X_p\right)^{k_1}\left(\frac{1}{2n}X_q\right)^{k_2} \tag{7.8}$$

Moreover

$$l^{(p+q)/2}\left(\frac{1}{n}\right) = \exp\left(\frac{1}{n}X_{\frac{p+q}{2}}\right) = \sum_k \frac{1}{k!}\left(\frac{1}{2n}X_{p+q}\right)^k. \tag{7.9}$$

Hence considering the first three terms of expansions (7.8) and (7.9) we obtain that $\left|x_2 - l^{(p+q)/2}(1/n)\right| = O(1/n^2)$. Iteratively, we get $\left|x_{2i} - l^{(p+q)/2}(i/n)\right| = O(1/n^2) + iO(1/n^2)$; in particular, since $x_{2n} = \xi(1)$ and $y^{\frac{p+q}{2}} = l^{(p+q)/2}(1)$, for $i = n$ we obtain the claim (7.4). □

We can now prove Theorem 3.2, i.e. the homogenization result for non-coercive Hamilton-Jacobi problem.

Proof of Theorem 3.2. Note that by Lemma 3.1, assumptions **(H1)-(H4)** implies **(L1)-(L4)**. By Theorem 3.1, the function $u^\varepsilon(t, x, \omega)$ in the left-hand

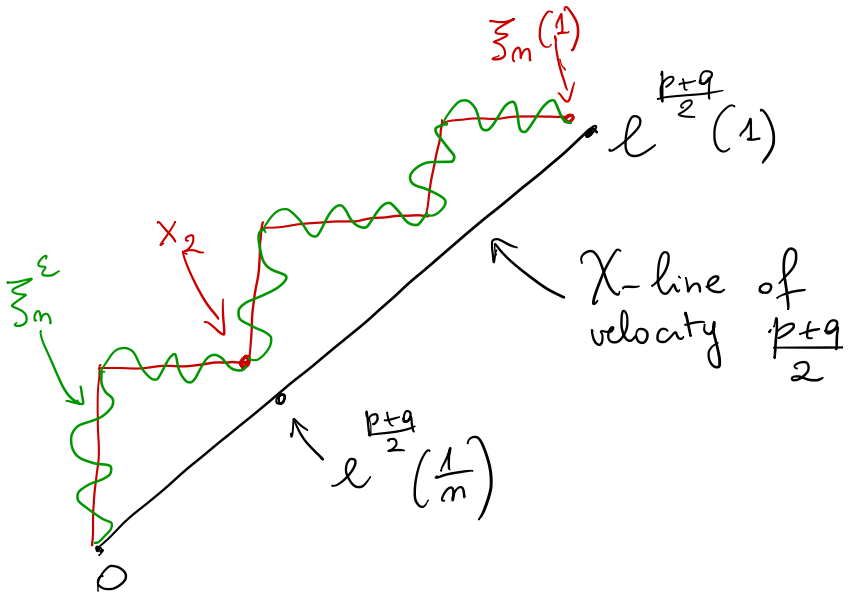

Figure 4: this picture illustrates the arguments of the proof of Proposition 7.1.

side of (6.17) is the unique viscosity solution of (3.1). We denote by $\bar{u}$ the right-hand side of (6.17), i.e.

$$\bar{u}(t,x) := \inf_{y\in\mathbb{R}^N}\left[g(y) + t\inf_{\alpha\in\mathcal{F}^t_{y,x}}\int_0^t \overline{L}(\alpha(s))ds\right],$$

We define the effective Hamiltonian $\overline{H}(q) := \overline{L}^*(q)$. The convexity and the superlinearity of $\bar{L}$ (see Theorem 7.1 and Proposition 5.1) imply $\overline{L}(q) = (\overline{L}^*)^*(q) = \overline{H}^*(q)$. By the Hopf-Lax formula in [8, Theorem 3.4], the function $\bar{u}(t,x)$ is the unique viscosity solution of (1.3). The convergence easily follows from (6.17). □

8 Appendix

In this appendix we prove Theorem 3.1 on the well posedness of problem (3.1).

Proposition 8.1. *Under the assumptions of Theorem 3.1, let u^ε be the function defined in* (3.3)*. Then u^ε is uniformly continuous in $[0,T]\times\mathbb{R}^N$.*

Proof. We want to prove that for any $\overline{\eta} > 0$ there exists $\overline{\delta} > 0$ such that

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(s,y,\omega)| < \overline{\eta}, \qquad \text{if } |t-s| + \|-y\circ x\|_{CC} < \overline{\delta}. \tag{8.1}$$

Step 1. We claim that for any $\eta > 0$ there exists $\delta > 0$ such that

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(0,y,\omega)| < \eta + m(d_{CC}(x,y)),\ \forall t\in[0,\delta], \forall x,y\in\mathbb{R}^N. \tag{8.2}$$

Indeed, by definition of u^ε and **(L2)** we have

$$u^\varepsilon(t,x,\omega) - u^\varepsilon(0,x,\omega) \le L^\varepsilon(x,x,t,\omega) \le \int_0^t L(\delta_{1/\varepsilon}(x),0,\omega)ds \le C_1 t.$$

On the other hand, by the assumption on g, for any $\overline{\eta} > 0$ there exists $\overline{y} \in \mathbb{R}^N$ such that

$$u^\varepsilon(t,x,\omega) - u^\varepsilon(0,x,\omega) \ge -m(d_{CC}(\overline{y},x)) + L^\varepsilon(x,\overline{y},t,\omega) - \overline{\eta}.$$

Moreover, by **(L2)**, for any $\overline{\eta}$ there exists $\xi \in \mathcal{A}^t_{\overline{y},x}$ such that

$$L^\varepsilon(x,\overline{y},t,\omega) \ge \int_0^t C_1^{-1}\left(|\alpha^\xi(s)|^\lambda - 1\right) ds - \overline{\eta} \ge -C_1^{-1}t - \overline{\eta}.$$

By the last two inequalities and by (**??**), we get

$$\begin{aligned} &u^\varepsilon(t,x,\omega) - u^\varepsilon(0,x,\omega) \ge \\ &-m\left(C_1^{1/\lambda}(\|u^\varepsilon\|_\infty + \|g\|_\infty + C_1^{-1}t + 1)^{1/\lambda} t^{1/\lambda - 1}\right) - C_1^{-1}t - 2\overline{\eta}, \end{aligned}$$

where the bound of $\|u^\varepsilon\|_\infty$ is an easy consequence of Lemma 4.1. Hence, for t sufficiently small and by the assumption on g, we get (8.2).

Step 2. We claim that, for any $\delta_1 > 0$ there exists m_{δ_1} such that

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(s,x,\omega)| < m_{\delta_1}(|t-s|),\ \forall t,s \ge \delta_1, \forall x \in \mathbb{R}^N. \tag{8.3}$$

Indeed, for any $\eta > 0$ there exists $\overline{y} \in \mathbb{R}^N$ such that

$$u^\varepsilon(t,x,\omega) - u^\varepsilon(s,x,\omega) \ge L^\varepsilon(x,\overline{y},t,\omega) - L^\varepsilon(x,\overline{y},s,\omega) - \eta.$$

The other inequality is similar. Using Lemma 4.4, we get (8.3).

Step 3. We claim that for any $\delta_2 > 0$ there exists m_{δ_2} such that

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(t,y,\omega)| < m_{\delta_2}(\|y^{-1}\circ x\|_{CC}),\ \forall t \ge \delta_2, \forall x,y \in \mathbb{R}^N. \tag{8.4}$$

Indeed, arguing as in Step 2, it is enough to prove

$$|L^\varepsilon(x,z,t,\omega) - L^\varepsilon(y,z,t,\omega)| \le m_{\delta_2}(\|y^{-1}\circ x\|_{CC}).$$

Actually, adding and subtracting $L^\varepsilon(y,z,t+\|-y\circ x\|_{CC},\omega))$, and using Lemma 4.2, we can conclude (8.4).

Step 4 W.l.o.g. assume $s \ge t$ and δ sufficiently small. For $0 \le t \le s \le \delta$, from step 1, we have

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(s,y,\omega)| \le 2\eta + m_\delta(d_{CC}(x,y)).$$

For $s > \delta$ and $|t-s| < \bar{\delta}$ (with $\bar{\delta} < \delta/2$), by step 2 and step 3, we get

$$|u^\varepsilon(t,x,\omega) - u^\varepsilon(s,y,\omega)| \le m_{\bar{\delta}}(\| - y\circ x\|_{CC} + |t-s|).$$

To conclude it suffices to choose $\bar{\delta}$ sufficiently small. □

Here we state the following result that will play a crucial role for the proof of Theorem 3.1 .

Lemma 8.1. *For every $R, T > 0$ there exists $\mu = \mu(T, R) > 0$ such that for every $(t, x) \in (0, T) \times B_R(0)$ there holds*

$$u^\varepsilon(t, x, \omega) = \inf \left\{ g(\xi(0)) + \int_0^t H^* \left(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega \right) ds \right\}$$

where the infimum is over all the $\alpha \in \mathcal{F}_x^t$ with $|\alpha^\xi| \leq \mu(R, T)$ where $\mathcal{F}_x^t$ is the set of all the m-valued measurable functions α such that the corresponding horizontal curve is $\xi^\alpha(s)$ with $\xi^\alpha(t) = x$.

Proof. The proof is the same as [8, Theorem 2.1] using [17, Theorem 7.4.6]. □

Finally we can prove Theorem 3.1.

Proof of Theorem 3.1.

We only sketch the proof; for the detailed calculations we refer the reader to [25, Section 10.3.3] and to [8] (see also [22, 33] for similar results). First of all we prove that u^ε is a solution of (3.1). We observe that by Lemma 8.1 u^ε satisfies the following optimality condition: for any $0 \leq h \leq t$ we have

$$u^\varepsilon(t, x, \omega) = \inf \left\{ \int_{t-h}^t L \left(\delta_{1/\varepsilon}(\xi(s)), \alpha^\xi(s), \omega \right) ds + u^\varepsilon(t - h, \xi(t - h), \omega) \right\},$$

where the infimum is over all the $\alpha \in \mathcal{F}_x^t$ with $|\alpha^\xi| \leq \mu(R, T)$.
From assumption **(L2)**, Proposition 4.1 and [25, Lemma 10.3.3], we get

$$\|u^\varepsilon\|_\infty \leq C, \qquad \text{for any compact } K \subset \mathbb{R}^N, \quad \|u^\varepsilon\|_{W^{1,\infty}([0,T]\times K)} \leq C_K.$$

Following the same arguments of [25, Theorem 2, Section 10.3.3] and [8], we get that u^ε fulfills $u^\varepsilon(0, x) = g(x)$ and it is a viscosity solution of

$$u_t + \mathcal{H}(x, Du) = 0,$$

where $\mathcal{H}(x, Du) = \max_{a \in \mathbb{R}^m, |a| \leq \mu(R,T)} \{p \cdot \sigma(x)a - L(x, a)\}$.
Arguing as in [8, equation (45) and proof of Theorem 3.2] we get that, if u is differentiable, then $\mathcal{H}(x, Du) = H(x, \sigma Du)$. Applying this property to a smooth test function, we conclude that u^ε is a viscosity solution of problem (3.1).

The uniqueness of the solution follows from the uniform continuity of u^ε (see Proposition 8.1) and applying the result of Biroli [14, Theorem 4.4]. □

Acknowledgments: The first author was partially supported by the Leverhulme Trust via grant RPG-2013-261 and by EPSRC via grant EP/M028607/1. The second author was partially supported by the EPSRC Grant "Random Perturbations of ultra-parabolic PDEs under rescaling". The third and forth authors are members of GNAMPA-INdAM and were partially supported also by the research project of the University of Padova "Mean-Field Games and Nonlinear PDEs" and by the Fondazione CaRiPaRo Project "Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games".
The authors would also like to thank P.E. Souganidis and R. Monti for the many interesting conversations.
The author are also grateful to the anonymous referee for his/her useful comments.
Conflict of Interest: The authors declare that they have no conflict of interest.

References

[1] M.A. Akcoglu, U. Krengel, *Ergodic Theorems for Superadditive Processes*, J. Reine Angew. Math. **323** (1981), 53–67.

[2] O. Alvarez, M. Bardi, *Ergodicity, stabilization, and singular perturbations for Bellman-Isaacs equations.*, Mem. Amer. Math. Soc. **204** (2010), no. 960, vi+77 pp.

[3] O. Alvarez, M. Bardi, C. Marchi, *Multiscale problems and homogenisation for second-order Hamilton-Jacobi equations*, J. Differential Equations **243** (2007), no. 2, 349–387.

[4] S.N. Armstrong, P. Cardaliaguet, *Quantitative stochastic homogenization of viscous Hamilton-Jacobi equations*, Comm. Partial Differential Equations **40** (2015), no. 3, 540–600.

[5] S.N. Armstrong, P. Cardaliaguet, P.E. Souganidis, *Error estimates and convergence rates for the stochastic homogenization of Hamilton-Jacobi equations*, J. Amer. Math. Soc. **27** (2014), no. 2, 479–540.

[6] S.N. Armstrong, P.E. Souganidis, *Stochastic homogenization of Hamilton-Jacobi and degenerate Bellman equations in unbounded environments*, J. Math. Pures Appl. (9) **97** (2012), no. 5, 460–504.

[7] S.N. Armstrong, P.E. Souganidis, *Stochastic homogenization of level-set convex Hamilton-Jacobi equations.*, Int. Math. Res. Not. IMRN 2013, no. 15, 3420–3449.

[8] Z.M. Balogh, A. Calogero, R. Pini, *The Hopf-Lax formula in Carnot groups: a control theoretic approach*, Calc. Var. Partial Differential Equations **49** (2014), no. 3-4, 1379–1414.

[9] M. Bardi, F. Dragoni, *Convexity and semiconvexity along vector fields*, Calc. Var. Partial Differential Equations **42** (2011), no.3-4, 405–427.

[10] M. Bardi, F. Dragoni; *Subdifferential and Properties of Convex Functions with respect to Vector Fields*, J. Convex Analysis **21** (2014), no. 3, 785–810.

[11] F. Baudoin; *An introduction to the geometry of stochastic flows*, Imperial College press, London, 2004.

[12] C. Bellettini, E. Le Donne, *Regularity of sets with constant horizontal normal in the Engel group*, Comm. Anal. Geom. **21** (2013), no. 3, 469–507.

[13] I. Birindelli, J. Wigniolle, *Homogenisation of Hamilton-Jacobi equations in the Heisenberg group*, Commun. Pure Appl. Anal. **2** (2003), no. 4, 461–479.

[14] M. Biroli, *Subelliptic Hamilton-Jacobi equations: the coercive evolution case*, Appl. Anal. **92** (2013), no. 1, 1–14.

[15] M. Biroli, U. Mosco, N.A. Tchou, *Homogenisation for degenerate operators with periodical coefficients with respect to the Heisenberg group.* C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), no. 5, 439–444.

[16] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, *Stratified Lie Groups and Potential Theory for their Sub-Laplacians*, Springer, Berlin 2007.

[17] P. Cannarsa, C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and their Applications 58, Birkhäuser, Boston MA, 2004.

[18] P. Cardaliaguet, P.E. Souganidis. *Homogenization and Enhancement of the G -Equation in Random Environments*, Comm. Pure Appl. Math. **60** (2013), no 10, 1582—1628.

[19] A. Ciomaga, P.E. Souganidis, H. Tran. *Stochastic homogenization of interfaces moving with changing sign normal velocity*, J. Diff. Equations, **245** (2015), 1025–1057.

[20] G. Dal Maso, L. Modica *Nonlinear stochastic homogenisation and ergodic theory*, J. Reine Angew. Math. **368** (1986), 28–42.

[21] A.I. Danilenko, *Mixing actions of the Heisenberg group*, Ergodic Theory Dynam. Systems **34** (2014), no. 4, 1142–1167.

[22] F. Dragoni, *Metric Hopf-Lax formula with semicontinuous data*, Discrete Contin. Dyn. Syst. **17** (2007), no. 4, 713–729.

[23] M. Duerinckx, A. Gloria *Stochastic homogenization of nonconvex unbounded integral functionals with convex growth*, Arch. Ration. Mech. Anal. **221** (2016), no. 3, 1511–1584.

[24] W. E, *A Class of homogenisation problems in the calculus of variations*, Comm. Pure Appl. Math. **44** (1991), no. 7, 733–759.

[25] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics 19, American Mathematical Society, Providence RI, 2010.

[26] W.H. Feldman, P.E. Souganidis, Homogenization and Non-Homogenization of certain Non-Convex Hamilton-Jacobi Equations J. Math. Pures Appl. (9) **108** (2017), no. 5, 751–782.

[27] B. Franchi, C.E. Gutierrez, T. van Nguyen, *Homogenisation and convergence of correctors in Carnot groups*, Comm. Partial Differential Equations **30** (2005), no. 10-12, 1817–1841.

[28] B. Franchi, M.C. Tesi, *Two-scale homogenisation in the Heisenberg group*, J. Math. Pures Appl. (9) **81** (2002), no. 6, 495–532.

[29] P.K. Friz, N.B. Victoir *Multidimensional Stochastic Processes as Rough Paths* Cambridge Studies in Advanced Mathematics 120, Cambridge University Press, Cambridge, 2010.

[30] E. Kosygina, F. Rezakhanlou, S.R.S. Varadhan, *Stochastic homogenization of Hamilton-Jacobi-Bellman equations*, Comm. Pure Appl. Math. **59**, no. 10, 1489–1521.

[31] P.-L. Lions, P.E. Souganidis, *Stochastic homogenization of Hamilton-Jacobi and "viscous"-Hamilton-Jacobi equations with convex nonlinearities–revisited*, Commun. Math. Sci. **8** (2010), no. 2, 627–637.

[32] P.-L. Lions, P.E. Souganidis, *Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media*, Comm. Partial Differential Equations **30** (2005), no. 1-3, 335–375.

[33] J.J. Manfredi, B. Stroffolini, *A version of the Hopf-Lax formula in the Heisenberg group*, Comm. Partial Differential Equations **27** (2002), no. 5-6, 1139–1159.

[34] P. Mannucci, B. Stroffolini, *Periodic homogenisation under a hypoellipticity condition*, NoDEA Nonlinear Differential Equations Appl. **22** (2015), no. 4, 579–600.

[35] R. Montgomery, *A tour of subriemannian geometries, their geodesics and application*, Mathematical Surveys and Monographs 91, American Mathematical Society, Providence RI, 2002.

[36] F. Rezakhanlou, J.E. Tarver, *Homogenisation for stochastic Hamilton-Jacobi equations*, Arch. Ration. Mech. Anal. **151** (2000), no. 4, 277–309.

[37] P. Souganidis, *Stochastic homogenisation of Hamilton-Jacobi equations and some applications*, Asymptot. Anal. **20** (1999), no. 1, 1–11.

[38] B. Stroffolini, *Homogenisation of Hamilton-Jacobi equations in Carnot groups*, ESAIM Control Optim. Calc. Var. **13** (2007), no. 1, 107–119.

[39] B. Ziliotto *Stochastic homogenization of nonconvex Hamilton-Jacobi equations: a counterexample*, Comm. Pure Appl. Math. **70** (2017), no. 9, 1798–1809.