Calc. Var. (2025) 64:107
https://doi.org/10.1007/s00526-025-02959-x

Calculus of Variations

Regularity for almost minimizers of a one-phase Bernoulli-type functional in Carnot groups of step two

Fausto Ferrari[1] · **Nicolò Forcillo**[2] · **Enzo Maria Merlino**[1]

Received: 11 July 2023 / Accepted: 10 February 2025

Abstract

We prove that nonnegative almost minimizers of the horizontal Bernoulli-type functional

$$J(u, \Omega) := \int_{\Omega} \left(|\nabla_{\mathbb{G}} u(x)|^2 + \chi_{\{u>0\}}(x) \right) dx$$

are Lipschitz continuous with respect to the Carnot-Carathéodory distance.

Mathematics Subject Classification 35R35 · 35R03

Contents

1 Introduction and main result
2 Notation and preliminary results
2.1 Carnot groups
2.2 Intrinsic distance and gauge pseudo-distance

Dedicated to Bruno Franchi: a master, a friend.

Communicated by L. Szekelyhidi.

The authors are members of Gruppo Nazionale per l'Analisi Matematica, la Probabilitá e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM): F.F. has been partially supported by the GNAMPA research project 2023 "Equazioni completamente non lineari locali e non locali", N.F. has been supported by the GNAMPA research project 2023 "Problemi variazionali/nonvariazionali: interazione tra metodi integrali e principi del massimo" and E.M.M. has partially supported by the GNAMPA research project 2023 "Equazioni nonlocali di tipo misto e geometrico". F.F. and E.M.M. have been supported by the PRIN research project 2022 7HX33Z - CUP J53D23003610006, "Pattern formation in nonlinear phenomena". The authors wish to thank Bruno Franchi and Francesca Corni for some useful discussions.

✉ Fausto Ferrari
fausto.ferrari@unibo.it

Nicolò Forcillo
forcill1@msu.edu

Enzo Maria Merlino
enzomaria.merlino2@unibo.it

1 Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, 40126 Bologna, Italy

2 Department of Mathematics, Michigan State University, 619 Red Cedar Road, East Lansing, MI 48824, USA

Published online: 24 March 2025

2.3 Folland-Stein and horizontal Sobolev classes . . . .
3 Dichotomy results . . . .
4 Lipschitz continuity of almost minimizers and proof of theorem 1.2 . . . .
References . . . .

1 Introduction and main result

In this paper, we study the regularity of almost minimizers of Bernoulli-type energy functionals in Carnot groups.

In the Euclidean setting, the regularity of minimizers for the classical one-phase Bernoulli energy functional

$$\mathscr{F}(u,\Omega) := \int_\Omega \Big(|\nabla u(x)|^2 + \chi_{\{u>0\}}(x) \Big)\, dx \tag{1.1}$$

has been deeply studied since the pioneering work of Alt and Caffarelli, see [3, 4]. We refer the reader to [52] and the references therein for a comprehensive description of the subject.

More recently, the regularity of almost minimizers associated with Bernoulli-type functionals was investigated as well. We recall that the notion of *almost minimality* is a relaxed notion of minimality that arises from variational problems with constraints, such as problems with prescribed volume or curvature and obstacle problems. Almgren introduced the notion of almost minimizers in geometric measure theory [2]. In the nonparametric setting, this notion was introduced by Anzellotti in [5]. Roughly speaking, almost minimizers may be considered as local perturbations of minimizers, which have smaller contribution at small scales. Hence, one may expect that similar regularity results for minimizers hold also for almost minimizers. Nevertheless, the condition satisfied by almost minimizers does not allow to obtain an Euler-Lagrange equation as in the case of minimizers.

Concerning Bernoulli-type functionals, in [16], David and Toro proved that almost minimizers of (1.1) are locally Lipschitz continuous, which is the optimal regularity for minimizers. Results about the free boundary regularity can be found in [15]. These works employ variational techniques involving tools coming from potential theory and geometric measure theory. We refer to [48] for a comprehensive overview on the topic. More recently, De Silva and Savin, in [18], introduced a more direct approach to obtain Lipschitz regularity of almost minimizers of (1.1) based on a dichotomy argument. Since their techniques involve, mainly, only metric properties, we point out that they may be extended to different more general frameworks. In this direction, see, for instance, [17] and [19], dedicated to the *thin* Bernoulli functional and a one-phase problem driven by the p-Laplace operator respectively.

Free boundary problems arising from the minimization of Bernoulli-type functionals may be formulated even in the noncommutative setting of Carnot groups. See, for instance, [20] for the case of the Heisenberg group. Carnot groups are an intensively-studied noncommutative structure playing a relevant role in many applications. For instance, they represent the tangent model of a general sub-Riemannian manifold and are the natural framework to describe some systems with nonholonomic constraints. On Carnot groups, there exists a comprehensive literature involving geometric measure theory [28, 29, 43], subelliptic partial differential equations [9, 24, 25], and differential geometry [11, 35, 36]. Nevertheless, very little is known on Bernoulli free boundary problems in this setting. For the one-phase case, in [23], only in the Heisenberg group, the authors proved the existence of minimizers, which turn out to be intrinsic harmonic away from the free boundary and present linear growth close to the free boundary. Besides, they proved some density estimates. Furthermore, we recall that, concerning the two-phase case, classical tools seem to fail in this setting. For instance, as recently shown in [21, 22], an intrinsic Alt-Caffarelli-Friedman monotonicity formula, written as the natural counterpart to the classical Euclidean one, does not hold. In addition, the viscosity approach occurs some difficulties due to the characteristic points. For instance, this is the case of the comparison principle, see [20, Section 8]. Moreover, we point out that, for these functionals, the regularity results do not follow from the classical theory in the calculus of variations, since the integrand of the functional is discontinuous (in the classical case, the integrand is required to be, for example, Lipschitz continuous, see e.g., [31, condition (8.48)] in the Euclidean context or, for example, under stronger regularity assumption in [12, 45], in Carnot groups).

The setting in which we are working on may be summarized as follows. Let $\mathbb{G}$ be a Carnot group of step two. Assume that $\Omega \subset \mathbb{G}$ is a measurable domain. We define the following energy functional

$$J(u, \Omega) := \int_{\Omega} \Big(|\nabla_{\mathbb{G}} u(x)|^2 + \chi_{\{u>0\}}(x) \Big)\, dx, \tag{1.2}$$

where u belongs to the *horizontal Sobolev space* $HW^{1,2}(\Omega)$, $u \geqslant 0$ a.e. in Ω, $\nabla_{\mathbb{G}} u$ is the so-called *horizontal gradient* of u and dx denotes the Haar measure of the group. We refer to Sect. 2 for the main notation and definitions.

In the present paper, we focus on some regularity properties of almost minimizers for J.

Definition 1.1 Let $\kappa \geqslant 0$ and $\beta > 0$, and $\Omega \subseteq \mathbb{G}$ be an open set. We say that $u \in HW^{1,2}(\Omega)$ is an almost minimizer of J in Ω, with constant κ and exponent β, if $u \geqslant 0$ $\mathcal{L}^n$-almost everywhere in Ω and

$$J(u, B_\varrho(x)) \leqslant (1 + \kappa \varrho^\beta) J(v, B_\varrho(x)), \tag{1.3}$$

for every metric ball $B_\varrho(x)$, of radius $\varrho \leqslant 0$ and center $x \in \mathbb{G}$, such that $\overline{B_\varrho(x)} \subset \Omega$ and for every $v \in HW^{1,2}(B_\varrho(x))$ such that $u - v \in HW_0^{1,2}(B_\varrho(x))$.

Our main result is the following one.

Theorem 1.2 *Let u be an almost minimizer for J in B_1 with constant κ and exponent β. Then,*

$$\|\nabla_{\mathbb{G}} u\|_{L^\infty(B_{1/2})} \leqslant C \Big(\|\nabla_{\mathbb{G}} u\|_{L^2(B_1)} + 1 \Big)$$

where $C > 0$ is a constant only depending on the homogeneous dimension Q, κ, and β.

In addition, u is uniformly Lipschitz continuous in a neighborhood of $\{u = 0\}$, namely if $u(0) = 0$, then

$$|\nabla_{\mathbb{G}} u| \leqslant C \quad \text{in } B_{r_0},$$

for some $C > 0$, only depending on Q, κ and β, and $r_0 \in (0, 1)$, depending on Q, κ, β and $\|\nabla_{\mathbb{G}} u\|_{L^2(B_1)}$.

In particular, the main contribution of our paper concerns the Lipschitz intrinsic regularity of almost minimizers. We remark that our result is also relevant for the case of minimizers (when $\kappa = 0$), since it is new also in this direction. Our approach is a generalization of the method introduced in [18]. Nevertheless, some tools in [18], mainly based on comparison arguments with harmonic replacements, cannot be applied on Carnot groups. Indeed, if the step of the group is strictly grater then one, the vector fields $X_1, \dots, X_{m_1}$, generating the first layer of the Lie algebra associated to $\mathbb{G}$, have no trivial commutators, i.e.

$$\left[X_i, X_j\right] \neq 0 \quad \Leftrightarrow \quad X_i X_j \neq X_j X_i, \quad \text{for every } i, j \in \{1, \dots, m_1\}, \text{ if } i \neq j.$$

Consequently, denoting by $\mathcal{L} = -\sum_{j=1}^{m_1} X_j^2$ the positive sub-Laplacian on $\mathbb{G}$, it results

$$X_i \mathcal{L} \neq \mathcal{L} X_i \tag{1.4}$$

for all $i = 1, \dots, m_1$. The lack of commutativity, encoded in (1.4), implies that functions $X_i u$, for $i \in \{1, \dots, m_1\}$, are not $\mathbb{G}$-harmonic even if u is. Therefore, the mean value characterization of $|\nabla_{\mathbb{G}} u|^2$ on which the argument proposed in [18] is based, is no longer applicable in our case. To face this aspect, we apply the regularity estimates for subelliptic equations, proved in [13], see also [44] in the case of the Heisenberg group. In addition, following the original idea of Anzellotti in [5], we exploit well-known results concerning the Morrey-Campanato-type estimates, developed in [39, 40] in homogeneous spaces. Moreover, we employ some classical tools coming from geometric analysis in Carnot groups, such as Poincaré and Sobolev-Poincaré inequalities, see, for instance, [26, 27, 30, 33, 37, 38]. Finally, it is worth recalling that our Lipschitz intrinsic regularity result only implies Hölder continuous regularity from a Euclidean point of view, since the Carnot-Charathéodory distance is not equivalent to the Euclidean one, see, e.g., [9, Proposition 5.15.1].

The classical motivations to study Bernoulli free boundary problems come from flows with jets and cavities, see, for instance, [10, Section 1.1], heat flows [1], electrochemical machining [34], combustion theory [7], electrical impedance tomography [6], and phase transitions, see, e.g., [8, 46, 47, 50].

The paper is organized as follows. In Sect. 2, we recall basic notions concerning Carnot groups and some useful tools. Next, in Sect. 3, we deal with a dichotomy property, a key step to prove our main result. Finally, Sect. 4 contains the proof of Theorem 1.2. We remind, even if it is not explicitly stated, that all the constants appearing in the sequel may change from line to line when they are universal, i.e. when they depend only on the group $\mathbb{G}$.

2 Notation and preliminary results

2.1 Carnot groups

A connected and simply connected Lie group $(\mathbb{G}, \circ)$ (in general non-abelian) is said to be a *Carnot group of step k* if its Lie algebra $\mathfrak{g}$ admits a *k-step stratification*, i.e. there exist linear subspaces $V_1, \dots, V_k$ such that

$$\mathfrak{g} = V_1 \oplus \ldots \oplus V_k, \quad [V_1, V_i] = V_{i+1}, \quad V_k \neq \{0\}, \quad V_i = \{0\} \text{ if } i > k, \tag{2.1}$$

where $[V_1, V_i]$ is the subspace of $\mathfrak{g}$ generated by the commutators $[X, Y]$ with $X \in V_1$ and $Y \in V_i$. The first layer V_1, the so-called *horizontal layer*, plays a key role in the theory since it generates all the algebra $\mathfrak{g}$ by commutation. We point out that a stratified Lie algebra can admit more than one stratification. However, the stratification turns out to be unique up to isomorphisms, thus, the related Carnot group structure is essentially unique (see [36, Proposition 1.17]). Note that when $k = 1$ the group $\mathbb{G}$ is abelian, so we return to the Euclidean situation.

Setting $m_i = \dim(V_i)$ and $h_i = m_1 + \cdots + m_i$, with $h_0 = 0$, for $i = 1, \dots, k$ (so that $h_k = n = \dim \mathfrak{g} = \dim \mathbb{G}$), we choose a basis $e_1, \dots, e_n$ of $\mathfrak{g}$ adapted to the stratification, that is such that

$$e_{h_{j-1}+1}, \dots, e_{h_j} \text{ is a basis of } V_j \text{ for each } j = 1, \dots, k.$$

Let $X = \{X_1, \dots, X_n\}$ be the family of left-invariant vector fields such that $X_i(e) = e_i$, $i = 1, \dots, n$, where e is the identity of $(\mathbb{G}, \circ)$. Thanks to (2.1), the subfamily $\{X_1, \dots, X_{m_1}\}$ generates by commutations all the other vector fields, we will refer to $X_1, \dots, X_{m_1}$ as *generating vector fields* of $(\mathbb{G}, \circ)$.

The map $X \mapsto X(e)$, associating a left-invariant vector field X to its value in e, defines an isomorphism from $\mathfrak{g}$ to $T\mathbb{G}_e$ (in turn identified with $\mathbb{R}^n$). We systematically use these identifications. Furthermore, by the assumption that $\mathbb{G}$ is simply connected, the exponential map is a global real-analytic diffeomorphism from $\mathfrak{g}$ onto $\mathbb{G}$ (see, e.g., [14, 51]), so each $x \in \mathbb{G}$ can be written in a unique way as $x = \exp(x_1 X_1 + \dots + x_n X_n)$. Using these *exponential coordinates*, we identify x with the n-tuple $(x_1, \dots, x_n) \in \mathbb{R}^n$ and $\mathbb{G}$ with $(\mathbb{R}^n, *)$ where the explicit expression of the group operation $*$ is determined by the Campbell-Hausdorff formula (see, for example, [9, 25]). In these coordinates, $e = (0, \dots, 0)$ and $(x_1, \dots, x_n)^{-1} = (-x_1, \dots, -x_n)$, and the adjoint operator in $L^2(\mathbb{G})$ of X_j, X_j^*, for $j = 1, \dots, n$, turns out to be $-X_j$ (see, for instance, [29, Proposition 2.2]). Moreover, if $x \in \mathbb{G}$ and $i = 1, \dots, k$, we set $x^{(i)} := (x_{h_{i-1}+1}, \dots, x_{h_i}) \in \mathbb{R}^{m_i}$, so that we can also identify x with $[x^{(1)}, \dots, x^{(k)}] \in \mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_k} = \mathbb{R}^n$.

Two important families of automorphism of $\mathbb{G}$ are the so-called *intrinsic translations* and *dilations*. For any $x \in \mathbb{G}$, the *(left) translation* $\tau_x : \mathbb{G} \to \mathbb{G}$ is defined as

$$z \mapsto \tau_x z := x \circ z.$$

For any $\lambda > 0$ we call *dilation*, associated with λ the linear automorphism of $\mathfrak{g}$ such that

$$\delta_\lambda(v) = \lambda^i v, \quad \text{if } v \in V_i, \quad \text{for } i = 1, \dots, k.$$

Through the exponential map, one can transfer the notion of dilation on the group $\mathbb{G}$. We denote again by δ_λ the map

$$\exp \circ \delta_\lambda \circ \exp^{-1} : \mathbb{G} \to \mathbb{G}.$$

Using the identification given by exponential coordinates, one can read the map $\delta_\lambda : \mathbb{G} \to \mathbb{G}$ as

$$\delta_\lambda(x_1, \dots, x_n) = (\lambda^{\alpha_1} x_1, \dots, \lambda^{\alpha_n} x_n), \tag{2.2}$$

where $\alpha_i \in \mathbb{N}$ is called *homogeneity of the variable* x_i in $\mathbb{G}$ (see [25, Chapter 1]) and is defined as

$$\alpha_j = i \quad \text{whenever } h_{i-1} + 1 \leqslant j \leqslant h_i,$$

hence $1 = \alpha_1 = \dots = \alpha_m < \alpha_{m+1} = 2 \leqslant \dots \leqslant \alpha_n = k$.

From Definition (2.2), one can easily verify the following properties of intrinsic dilations.

Lemma 2.1 *For all* $\lambda, \mu > 0$*, one has:*

(1) $\delta_1 = \mathrm{id}_{\mathbb{G}}$;
(2) $\delta_\lambda^{-1} = \delta_{\lambda^{-1}}$;
(3) $\delta_\lambda \circ \delta_\mu = \delta_{\lambda\mu}$;
(4) for every $p, p' \in \mathbb{G}$*, it holds* $\delta_\lambda(p) \cdot \delta_\lambda(p') = \delta_\lambda(p \cdot p')$.

By left translation, the horizontal layer determines a subbundle of the tangent bundle $T\mathbb{G}$ over $\mathbb{G}$. This subbundle, spanned by the family of vector fields $X = (X_1, \dots, X_{m_1})$, plays a crucial role in the theory and is called the *horizontal bundle* $H\mathbb{G}$. The fibers of $H\mathbb{G}$ are

$$H\mathbb{G}_x = \operatorname{span}\{X_1(x), \dots, X_{m_1}(x)\}, \qquad x \in \mathbb{G}.$$

A sub-Riemannian structure is induced on $\mathbb{G}$ by endowing each fiber of $H\mathbb{G}$ with a scalar product $\langle \cdot, \cdot \rangle_x$, that makes the basis $X_1(x), \dots, X_{m_1}(x)$ orthonormal, and the related norm $|\cdot|_x$. These are defined as follows: if $v = \sum_{i=1}^{m_1} v_i X_i(x) = (v_1, \dots, v_{m_1})$ and $w = \sum_{i=1}^{m_1} w_i X_i(x) = (w_1, \dots, w_{m_1})$ are in $H\mathbb{G}_x$, then $\langle v, w \rangle_x := \sum_{j=1}^{m_1} v_j w_j$ and $|v|_x^2 := \langle v, v \rangle_x$.

The sections of $H\mathbb{G}$ are called *horizontal sections*, and for any $x \in \mathbb{G}$, a vector of $H\mathbb{G}_x$ is a *horizontal vector*, while a vector in $T\mathbb{G}_x$ that is not horizontal is called a vertical vector. Each horizontal section is identified by its canonical coordinates with respect to this moving frame $X_1, \dots, X_{m_1}$. In this way, a horizontal section φ is represented with a function $\varphi : \mathbb{G} \to \mathbb{R}^{m_1}$. To simplify the notation, when dealing with two horizontal sections ϕ and ψ, we drop the index x in the scalar product and norm.

We collect some properties of the group operation and canonical vector fields, see [9].

Proposition 2.2 *The group operation has the form*

$$x \circ y = x + y + \mathcal{Q}(x, y), \quad \textit{for all } x, y \in \mathbb{G},$$

where $\mathcal{Q} = (\mathcal{Q}_1, \dots, \mathcal{Q}_n) : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ *and every* $\mathcal{Q}_i$, $i = 1, \dots, n$, *is a homogeneous polynomial of degree* α_i *with respect to the intrinsic dilations of* $\mathbb{G}$, *that is,*

$$\mathcal{Q}_i(\delta_\lambda(x), \delta_\lambda(y)) = \lambda^{\alpha_i} \mathcal{Q}_i(x, y) \quad \textit{for all } x, y \in \mathbb{G} \textit{ and } \lambda > 0.$$

Moreover, for all $x, y \in \mathbb{G}$, *we have:*

(i) $\mathcal{Q}$ *is antisymmetric, namely,*

$$\mathcal{Q}_i(x, y) = -\mathcal{Q}_i(-y, -x), \quad \textit{for } i = 1, \dots, n;$$

(ii)

$$\begin{aligned} &\mathcal{Q}_1(x, y) = \dots = \mathcal{Q}_{m_1}(x, y) = 0, \\ &\mathcal{Q}_i(x, 0) = \mathcal{Q}_i(0, y) = 0 \textit{ and } \mathcal{Q}_i(x, x) = \mathcal{Q}_i(x, -x) = 0, \quad \textit{for } m_1 < i \leqslant n, \\ &\mathcal{Q}_i(x, y) = \mathcal{Q}_i(x_1, , x_{h_j - 1}, y_1, \dots, y_{h_j - 1}), \quad \textit{if } 1 < j \leqslant k \textit{ and } i \leqslant h_j; \end{aligned}$$

(iii)

$$\mathcal{Q}_i(x, y) = \sum_{k,h} \mathcal{R}^i_{h,k}(x, y)(x_k y_h - x_h y_k),$$

where the functions $\mathcal{R}^i_{h,k}$ *are homogeneous polynomials of degree* $\alpha_i - \alpha_k - \alpha_h$ *with respect to the intrinsic dilations and the sum is extended to all h and k such that* $\alpha_h + \alpha_k \leqslant \alpha_i$.

The following result is contained in [29, Proposition 2.2].

Proposition 2.3 *The vector fields* X_j *have polynomial coefficients and are of the form*

$$X_j(x) = \partial_j + \sum_{i > h_l}^{n} q_{i,j}(x) \partial_i, \quad j = 1, \dots, n \quad \textit{and} \quad j \leqslant h_l,$$

where $q_{i,j}(x) = \frac{\partial \mathcal{Q}_i}{\partial y_j}(x, y)|_{y=0}$ *and the* $\mathcal{Q}_i$*'s are those ones defined in Proposition 2.2 for* $h_{l-1} < j \leqslant h_l$ *and* $1 \leqslant l \leqslant k$. *So, if* $h_{l-1} < j \leqslant h_l$, *then* $q_{i,j}(x) = q_{i,j}(x_1, \dots, x_{h_{l-1}})$ *and* $q_{i,j}(0) = 0$.

2.2 Intrinsic distance and gauge pseudo-distance

An absolutely continuous curve $\gamma : [0, T] \to \mathbb{G}$ is called *sub-unitary* with respect to $X_1, \dots, X_{m_1}$ if it is a *horizontal curve*, that is, if there exist real measurable functions $c_1, \dots, c_{m_1} : [0, T] \to \mathbb{R}$ such that

$$\dot{\gamma}(s) = \sum_{j=1}^{m_1} c_j(s) X_j(\gamma(s)) \quad \text{for } \mathcal{L}^1\text{-a.e. } s \in [0, T],$$

with $\sum_{j=1}^{m_1} c_j(s)^2 \leqslant 1$ for $\mathcal{L}^1$-a.e. $s \in [0, T]$.

Definition 2.4 (Carnot-Carathéodory distance) Let $\mathbb{G}$ be a Carnot group. For $p, q \in \mathbb{G}$, we define their *Carnot-Carathéodory distance* $d_c(p, q)$ as

$$d_c(p, q) := \inf\{T > 0 : \text{ there exists a sub-unitary curve } \gamma \text{ with } \gamma(0) = p,\ \gamma(T) = q\}.$$

We remark that this distance is well defined: the set of sub-unitary curves connecting p and q is nonempty by Chow's theorem [9, Theorem 19.1.3], since by (2.1), the rank of the Lie algebra generated by $X_1, \dots, X_{m_1}$ is n. The Carnot-Carathéodory distance d_c induces on $\mathbb{G}$ the same topology as the standard Euclidean one. We shall denote by $B_r(p)$ the open ball, centered in p and of radius $r > 0$, associated with d_c. For the sake of simplicity, if $p = e$, we will use the notation $B_r(e) = B_r$.

The distance d_c is equivalent to a more explicit pseudo-distance, called the *gauge pseudo-distance*, defined as follows. Let $|| \cdot ||$ denote the Euclidean distance to the origin in the Lie algebra $\mathfrak{g}$. For $u = u_1 + \cdots + u_k \in \mathfrak{g}$, with $u_i \in V_i$, one defines

$$|u|_{\mathfrak{g}} := \left(\sum_{i=1}^{k} ||u_i||^{2k!/i} \right)^{\frac{1}{2k!}}. \tag{2.3}$$

The *nonisotropic gauge* in $\mathbb{G}$ is

$$|p|_{\mathbb{G}} := |\exp^{-1} p|_{\mathfrak{g}}, \qquad p \in \mathbb{G},$$

see [24] and [25]. Since the exponential map $\exp : \mathfrak{g} \to \mathbb{G}$ is a C^∞-diffeomorphism (actually, analytic), the map $p \to |p|_{\mathbb{G}}$ is $C^\infty(\mathbb{G} \setminus \{e\})$. Notice that, from (2.3) and (2.2), we have, for any $\lambda > 0$ and $p \in \mathbb{G}$,

$$|\delta_\lambda(p)|_{\mathbb{G}} = \lambda |p|_{\mathbb{G}}. \tag{2.4}$$

The *gauge pseudo-distance* in $\mathbb{G}$ is defined by

$$d(p, q) := |p^{-1} \circ q|_{\mathbb{G}}, \qquad p, q \in \mathbb{G}. \tag{2.5}$$

The function d has all the properties of a distance, except the triangle inequality, which is satisfied with a universal constant, usually different from one, on the right-hand side, see [9, 25]. Since the dilations are group automorphisms, from (2.5) and (2.4), it follows that d is homogeneous of degree one with respect to the group dilations, that is, for any $\lambda > 0$ and all $p, p' \in \mathbb{G}$

$$d(\delta_\lambda(p), \delta_\lambda(p')) = \lambda d(p, p').$$

It is well known, see, for instance, Proposition 5.1.4 in [9], that there exists a constant $\Lambda_{\mathbb{G}}$, only depending on the group $\mathbb{G}$, such that, for $p \in \mathbb{G}$,

$$\Lambda_{\mathbb{G}}^{-1} |p|_{\mathbb{G}} \leqslant d_c(e, p) \leqslant \Lambda_{\mathbb{G}} |p|_{\mathbb{G}}. \tag{2.6}$$

The integer

$$Q := \sum_{i=1}^{k} i \dim(V_i) \tag{2.7}$$

is the *homogeneous dimension* of $\mathbb{G}$. It is the Hausdorff dimension of $\mathbb{G} \cong \mathbb{R}^n$ with respect to the distance d_c, see [41, 42].

The n-dimensional Lebesgue measure $\mathcal{L}^n$ is the Haar measure of the group $\mathbb{G}$, i.e., for every $\mathcal{L}^n$-measurable set $E \subset \mathbb{G}$ and $x \in \mathbb{G}$, it results $\mathcal{L}^n(x \circ E) = \mathcal{L}^n(E)$. Moreover, if $\lambda > 0$, then $\mathcal{L}^n(\delta_\lambda(E)) = \lambda^Q \mathcal{L}^n(E)$, where Q is the homogeneous dimension, see (2.7). In particular, for any $r > 0$ and $p \in \mathbb{G}$, it holds

$$\mathcal{L}^n(B_r(p)) = r^Q \mathcal{L}^n(B_1(p)) = r^Q \mathcal{L}^n(B_1).$$

Let us point out that all the spaces $L^p(\mathbb{G})$ we will deal with are defined with respect to the Lebesgue measure $\mathcal{L}^n$. If $A \subset \mathbb{G}$ is $\mathcal{L}^n$-measurable, we write $|A| = \mathcal{L}^n(A)$. Hereafter, unless differently specified, all considered domains are at least $\mathcal{L}^n$-measurable.

A map $L : \mathbb{G} \to \mathbb{R}$ is $\mathbb{G}$-*linear* if it is a group homeomorphism from $(\mathbb{G}, \circ)$ to $(\mathbb{R}, +)$ and is homogeneous of degree 1 with respect to the intrinsic dilations of $\mathbb{G}$, that is, $L(\delta_\lambda(x)) = \lambda L(x)$ for $\lambda > 0$ and $x \in \mathbb{G}$. Similarly, we say that a map $\ell : \mathbb{G} \to \mathbb{R}$ is $\mathbb{G}$-affine if there exists a $\mathbb{G}$-linear map L and $c \in \mathbb{R}$ such that $\ell(x) = L(x) + c$ for every $x \in \mathbb{G}$. Given a basis $(X_1, \dots, X_n)$, all $\mathbb{G}$-linear maps are represented as follows.

Proposition 2.5 *A map $L : \mathbb{G} \to \mathbb{R}$ is $\mathbb{G}$-linear if and only if there is $a = (a_1, \dots, a_{m_1}) \in \mathbb{R}^{m_1}$ such that, if $x = (x_1, \dots, x_n) \in \mathbb{G}$, then*

$$L(x) = \sum_{i=1}^{m_1} a_i x_i.$$

Moreover, if $x = (x_1, \dots, x_n) \in \mathbb{G}$ and $x_0 \in \mathbb{G}$ are given, we set

$$\pi_{x_0}(x) := \sum_{j=1}^{m_1} x_j X_j(x_0). \tag{2.8}$$

2.3 Folland-Stein and horizontal Sobolev classes

Fixed $\Omega \subseteq \mathbb{G}$, the action of vector fields X_j, with $j = 1, \dots, m_1$, on a function $f : \Omega \to \mathbb{R}$ is specified by the Lie derivative. We say that a function f is differentiable along the direction X_j at $x_0 \in \mathbb{G}$, when $j \in \{1, \dots, m_1\}$, if the map $\lambda \mapsto f(\tau_{x_0}(\delta_\lambda(e_j)))$ is differentiable at $\lambda = 0$, where e_j is the j-th vector of the canonical basis of $\mathbb{R}^{m_1}$. In this case, we will write

$$X_j f(x_0) = \left. \frac{d}{d\lambda} f(\tau_{x_0}(\delta_\lambda(e_j))) \right|_{\lambda=0}.$$

If, instead, $f \in L^1_{\mathrm{loc}}(\Omega)$, $X_j f$ exists in a weak sense if

$$\int_\Omega f X_j \varphi \, d\mathcal{L}^n = -\int_\Omega \varphi X_j f \, d\mathcal{L}^n$$

for each $\varphi \in C_0^\infty(\Omega)$.

Let us fix a basis $X_1, \ldots, X_{m_1}$ of the horizontal layer. For any function $f : \Omega \to \mathbb{R}$ for which the partial derivatives $X_j f$, $j = 1, \ldots, m_1$, exist, we define the horizontal gradient of f, denoted by $\nabla_{\mathbb{G}} f$, as the horizontal section

$$\nabla_{\mathbb{G}} f := \sum_{i=1}^{m_1} (X_i f) X_i, \tag{2.9}$$

whose coordinates are $(X_1 f, \ldots, X_{m_1} f)$. Moreover, if $\phi = (\phi_1, \ldots, \phi_{m_1})$ is a horizontal section such that $X_j \phi_j \in L^1_{\mathrm{loc}}(\mathbb{G})$ for $j = 1, \ldots, m_1$, we define $\mathrm{div}_{\mathbb{G}} \phi$ as the real-valued function

$$\mathrm{div}_{\mathbb{G}} \phi := -\sum_{j=1}^{m_1} X_j^* \phi_j = \sum_{j=1}^{m_1} X_j \phi_j.$$

The positive *sub-Laplacian* operator on $\mathbb{G}$ is the second-order differential operator given by

$$\mathcal{L} := \sum_{j=1}^{m_1} X_j^* X_j = -\sum_{j=1}^{m_1} X_j^2.$$

It is easy to see that

$$\mathcal{L} u = -\mathrm{div}_{\mathbb{G}}(\nabla_{\mathbb{G}} u). \tag{2.10}$$

The operator $\mathcal{L}$ is left-invariant with respect to group translations and homogeneous of degree two with respect to group dilations, i.e., for any $x \in \mathbb{G}$ and $\lambda > 0$, we have

$$\mathcal{L}(u \circ \tau_x) = (\mathcal{L} u) \circ \tau_x, \qquad \mathcal{L}(\delta_\lambda(u)) = \lambda^2 \delta_\lambda(\mathcal{L} u).$$

Furthermore, by the assumption (2.1), the system $\{X_1, \ldots, X_{m_1}\}$ satisfies the finite rank condition

$$\mathrm{rank}\ \mathrm{Lie}[X_1, \ldots, X_{m_1}] = n,$$

therefore by Hörmander's theorem [32] the operator $\mathcal{L}$ is hypoelliptic. However, when the step k of $\mathbb{G}$ is greater than one, the operator $\mathcal{L}$ fails to be elliptic at every point of $\mathbb{G}$.

The Hölder classes $C^{k,\alpha}$ had been introduced by Folland and Stein, see [24, 25]. The functions in these classes are Hölder continuous with respect to the metric d_c.

Definition 2.6 (Folland-Stein classes) Let $\Omega \subseteq \mathbb{G}$ be an open set, $0 < \alpha \leqslant 1$ and $u : \Omega \to \mathbb{R}$. We say that $u \in C^{0,\alpha}(\Omega)$ if there exists a constant $M > 0$ such that

$$|u(x) - u(y)| \leqslant M \, d_c(x, y)^\alpha \qquad \text{for every } x, y \in \Omega.$$

Defining the Hölder seminorm of $u \in C^{0,\alpha}(\Omega)$ as

$$[u]_{C^{0,\alpha}(\Omega)} := \sup_{\substack{x,y\in\Omega \\ x \neq y}} \frac{|u(x) - u(y)|}{d_c(x, y)^\alpha},$$

the space $C^{0,\alpha}(\Omega)$ is a Banach space with the norm

$$\|u\|_{C^{0,\alpha}(\Omega)} := \|u\|_{L^\infty(\Omega)} + [u]_{C^{0,\alpha}(\Omega)}.$$

For any $k \in \mathbb{N}$, we say that $u \in C^{k,\alpha}(\Omega)$ if $X_i u \in C^{k-1,\alpha}(\Omega)$ for every $i = 1, \ldots, m_1$.

Note that $C^{0,1}(\Omega)$ coincides with the class of Lipschitz continuous functions with respect to the metric d_c on Ω. Also, we denote by $C^{k,\lambda}(\Omega, H\mathbb{G})$ the space of all the horizontal sections $\phi : \Omega \to H\mathbb{G}$, $\phi := (\phi_1, \ldots, \phi_{m_1})$, such that $\phi_j \in C^{k,\lambda}(\Omega)$ for $j = 1, \ldots, m_1$.

Definition 2.7 (Horizontal Sobolev spaces) Given an open set $\Omega \subseteq \mathbb{G}$, $u : \Omega \to \mathbb{R}$ and $1 \leqslant p < \infty$, the horizontal Sobolev spaces are defined as

$$HW^{1,p}(\Omega) := \left\{ u \in L^p(\Omega) : \ X_j u \in L^p(\Omega) \text{ for all } \ j = 1, \ldots, m_1 \right\},$$

which is a Banach space with the norm

$$\|u\|_{HW^{1,p}(\Omega)} := \|u\|_{L^p(\Omega)} + \|\nabla_{\mathbb{G}} u\|_{L^p(\Omega, H\mathbb{G})}.$$

The subspace $HW_0^{1,p}(\Omega)$ of $HW^{1,p}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ with respect to the norm $\| \cdot \|_{HW^{1,p}(\Omega)}$.

3 Dichotomy results

The first step for reaching Theorem 1.2 is to prove a *dichotomy* result. Roughly speaking, two situations can occur: either the average of the energy of an almost minimizer decreases in a smaller ball, or the distance between its horizontal gradient and a suitable constant horizontal section becomes as small as we wish (that is, $\mathbb{G}$-linear functions are the "only ones for which the average does not improve in small balls"). Before passing to the proof of this result, let us state the following remark.

Remark 3.1 Let be $\ell : \mathbb{G} \to \mathbb{R}$ a $\mathbb{G}$-affine map. By Proposition 2.5, there exist a suitable $q = (q_1, \ldots, q_{m_1}) \in \mathbb{R}^{m_1}$ and $c \in \mathbb{R}$ such that

$$\ell(x) = \sum_{j=1}^{m_1} q_i x_i + c \qquad \text{for any } x \in \mathbb{G}.$$

Recalling (2.9), by Proposition 2.3, we directly get

$$\nabla_{\mathbb{G}} \ell(x) = \sum_{j=1}^{m_1} q_j X_j(x) =: \mathbf{q}(x), \qquad \text{for any } x \in \mathbb{G}. \tag{3.1}$$

Thus, the horizontal gradient of a $\mathbb{G}$-affine map ℓ is a horizontal section $\mathbf{q}$, which has constant components with respect to the moving frame given by generating vector fields. From now on, we call such a horizontal section a *constant* horizontal section. Note, however, that a constant horizontal section still depends on the point $x \in \mathbb{G}$ due to the moving frame. We also observe that all $\mathbb{G}$-affine functions ℓ as before are $\mathbb{G}$-harmonic in $\mathbb{G}$, in the sense that $\mathcal{L}\ell = 0$ in $\mathbb{G}$. Indeed, by (2.10) and (3.1), it immediately holds $\mathcal{L}\ell = 0$ in $\mathbb{G}$. We refer to [9] for further characterizations of $\mathbb{G}$-harmonicity within Carnot groups.

Our first result is the following.

Proposition 3.2 *Let $u \in HW^{1,2}(B_1)$ be such that*

$$J(u, B_1) \leqslant (1+\sigma) J(v, B_1) \tag{3.2}$$

for all $v \in HW^{1,2}(B_1)$ such that $u - v \in HW_0^{1,2}(B_1)$. Denoting by

$$a := \left(\fint_{B_1} |\nabla_{\mathbb{G}} u(x)|^2 dx \right)^{1/2}, \tag{3.3}$$

there exists $\varepsilon_0 \in (0,1)$ such that, for every $\varepsilon \in (0, \varepsilon_0)$ there exist $\eta \in (0,1)$, $M \geqslant 1$, and $\sigma_0 \in (0,1)$, depending on ε and Q, such that, if $\sigma \in [0, \sigma_0]$ and $a \geqslant M$, then the following dichotomy holds. Either

$$\left(\fint_{B_\eta} |\nabla_{\mathbb{G}} u(x)|^2 dx \right)^{1/2} \leqslant \frac{a}{2}, \tag{3.4}$$

or

$$\left(\fint_{B_\eta} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2 dx \right)^{1/2} \leqslant \varepsilon a, \tag{3.5}$$

where $\mathbf{q} : \mathbb{G} \to H\mathbb{G}$ is a constant horizontal section, see (3.1), with

$$\frac{a}{4} < |\mathbf{q}| \leqslant C_0 a, \tag{3.6}$$

for some universal constant $C_0 > 0$.

Proof We split the proof into several steps.

Step 1: pointwise estimates. Let $v : B_1 \to \mathbb{R}$ be the $\mathbb{G}$-harmonic replacement of u in B_1, that is, the unique solution of Dirichlet problem

$$\begin{cases} \mathcal{L} v = 0 & \text{in } B_1, \\ u - v \in HW_0^{1,2}(B_1), \end{cases}$$

or, equivalently, the minimizer of Dirichlet energy among all competitors $w \in HW^{1,2}(B_1)$ such that $u - w \in HW_0^{1,2}(B_1)$, i.e.,

$$\int_{B_1} |\nabla_{\mathbb{G}} v(x)|^2 \, dx = \min_{u - w \in HW_0^{1,2}(B_1)} \int_{B_1} |\nabla_{\mathbb{G}} w(x)|^2 \, dx. \tag{3.7}$$

By Theorem 1.1 in [13], there exists a constant $C_0 > 0$, only depending on Q, such that

$$\sup_{B_{1/2}} |\nabla_{\mathbb{G}} v| \leqslant C_0 \left(\fint_{B_1} |\nabla_{\mathbb{G}} v(x)|^2 dx \right)^{1/2}.$$

Then, recalling (3.3), by the minimality (3.7) of v for the Dirichlet energy, we conclude

$$|\nabla_{\mathbb{G}} v| \leqslant C_0 a \quad \text{in } B_{1/2}. \tag{3.8}$$

Step 2: oscillation estimates. By Theorem 1.3 in [13], we have the following oscillation estimates: for all $\eta \in (0, 1/2]$, it holds

$$\max_{1 \leqslant i \leqslant m_1} \operatorname{osc}_{B_\eta} X_i v \leqslant C \left(\frac{\eta}{1/2} \right)^{\alpha} \left(\fint_{B_1} |\nabla_{\mathbb{G}} v(y)|^2 dy \right)^{1/2},$$

for some $\alpha = \alpha(Q) \in (0, 1]$ and universal constant $C > 0$. Therefore, we have

$$\max_{1 \leqslant i \leqslant m_1} |X_i v(x) - X_i v(0)| \leqslant \max_{1 \leqslant i \leqslant m_1} \operatorname{osc}_{B_\eta} X_i v \leqslant C \left(\frac{\eta}{1/2}\right)^{\alpha} \left(\fint_{B_1} |\nabla_{\mathbb{G}} v(y)|^2 dy \right)^{1/2} \quad \text{for } x \in B_\eta. \tag{3.9}$$

Thus, for any $i = 1, \ldots, m_1$, denoting by $q_i = X_i v(0)$ and $\mathbf{q} : \mathbb{G} \to HG$ the constant horizontal section defined as in (3.1), it results

$$|\nabla_{\mathbb{G}} v - \mathbf{q}| \leqslant C_1 \left(\frac{\eta}{1/2}\right)^{\alpha} \left(\fint_{B_1} |\nabla_{\mathbb{G}} v(y)|^2 dy \right)^{1/2} \quad \text{in } B_\eta, \tag{3.10}$$

for some universal constant $C_1 > 0$. This gives that, for all $\eta \in (0, 1/2]$,

$$\begin{aligned} \fint_{B_\eta} |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)|^2 \, dx &\leqslant \fint_{B_\eta} \left(C_1^2 \left(\frac{\eta}{1/2}\right)^{2\alpha} \fint_{B_1} |\nabla_{\mathbb{G}} v(y)|^2 \, dy \right) dx \\ &= C_2 \, \eta^{2\alpha} \fint_{B_1} |\nabla_{\mathbb{G}} v(y)|^2 dy \\ &\leqslant C_2 \, \eta^{2\alpha} \, a^2, \end{aligned} \tag{3.11}$$

for some $\alpha \in (0, 1)$ and $C_2 > 0$ universal.

Step 3: proximity to the $\mathbb{G}$ *-harmonic replacement.* We know that

$$\begin{aligned} \int_{B_1} |\nabla_{\mathbb{G}} u - \nabla_{\mathbb{G}} v|^2 \, dx &\leqslant J(u, B_1) + \int_{B_1} (|\nabla_{\mathbb{G}} v|^2 - 2 \langle \nabla_{\mathbb{G}} u, \nabla_{\mathbb{G}} v \rangle) \, dx \\ &= J(u, B_1) - \int_{B_1} (|\nabla_{\mathbb{G}} v|^2 + 2 \langle \nabla_{\mathbb{G}} (u - v), \nabla_{\mathbb{G}} v \rangle) \, dx. \end{aligned} \tag{3.12}$$

On the other hand, since $X_j^* = -X_j$ for $j = 1, \ldots, m_1$ and $u - v \in HW_0^{1,2}(B_1)$, it occurs

$$\begin{aligned} \int_{B_1} \langle \nabla_{\mathbb{G}} (u - v), \nabla_{\mathbb{G}} v \rangle \, dx &= \sum_{j=1}^{m_1} \int_{B_1} X_j (u - v) X_j v \, dx \\ &= - \sum_{j=1}^{m_1} \int_{B_1} (u - v) X_j^2 v \, dx = \int_{B_1} (u - v) \mathcal{L} v \, dx = 0, \end{aligned} \tag{3.13}$$

which together with (3.12) gives

$$\int_{B_1} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \, dx \leqslant J(u, B_1) - \int_{B_1} |\nabla_{\mathbb{G}} v(x)|^2 dx. \tag{3.14}$$

Then, using the hypothesis (3.2) and the minimality of v in (3.7), for $\sigma > 0$ to be made precise later, we obtain

$$\begin{aligned} \int_{B_1} \left(|\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \right) dx &\leqslant (1 + \sigma) J(v, B_1) - \textstyle\int_{B_1} |\nabla_{\mathbb{G}} v(x)|^2 \, dx \\ &\leqslant C \left(\sigma \textstyle\int_{B_1} |\nabla_{\mathbb{G}} v(x)|^2 \, dx + 1 \right) \\ &\leqslant C \left(\sigma \textstyle\int_{B_1} |\nabla_{\mathbb{G}} u(x)|^2 \, dx + 1 \right) \end{aligned}$$

for some suitable positive constant C only depending on Q. Consequently, taking the average over B_1, we achieve

$$\fint_{B_1} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \, dx \leqslant C(\sigma a^2 + 1).$$

Using this and (3.11), we get

$$\begin{aligned}
\fint_{B_\eta} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2 \, dx &\leqslant 2 \fint_{B_\eta} \left(|\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 + |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)|^2 \right) dx \\
&\leqslant 2 \left(\frac{|B_1|}{|B_\eta|} \fint_{B_1} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \, dx + C_2 \, \eta^{2\alpha} \, a^2 \right) \quad (3.15) \\
&\leqslant 2 \Big(C \eta^{-Q} (\sigma a^2 + 1) + C_2 a^2 \eta^{2\alpha} \Big), \\
&\leqslant 2C\eta^{-Q} \sigma a^2 + 2C\eta^{-Q} + 2C_2 a^2 \eta^{2\alpha},
\end{aligned}$$

and hence

$$\fint_{B_\eta} |\nabla_{\mathbb{G}} u(x)|^2 \, dx \leqslant 4C\eta^{-Q} \sigma a^2 + 4C\eta^{-Q} + 4C_2 a^2 \eta^{2\alpha} + 2 \, |\mathbf{q}|^2 . \tag{3.16}$$

Step 4: perturbative estimates. Now, given $\varepsilon_0 \in (0, 1/4]$, we claim that for every $\varepsilon \in (0, \varepsilon_0)$, there exists η small enough (depending on ε) such that if σ is sufficiently small and a sufficiently large (depending on η, and thus on ε), then

$$4C\eta^{-Q} \sigma a^2 + 4C\eta^{-Q} + 4C_2 a^2 \eta^{2\alpha} \leqslant 2\varepsilon^2 a^2 \leqslant \frac{a^2}{8}. \tag{3.17}$$

Let $\eta > 0$ sufficiently small such that $\varepsilon^2 - 2C\eta > 2C_2 \eta^{2\alpha}$. So, defining

$$M := \left(\frac{2C\eta^{-Q}}{\varepsilon^2 - 2C\eta - 2C_2\eta^{2\alpha}} \right)^{1/2},$$

we can suppose $M \geqslant 1$, by taking η small enough. Therefore, for every $a \geqslant M$ and $0 \leqslant \sigma \leqslant \eta^{Q+1} =: \sigma_0$, we have, exploiting the expression of M,

$$\begin{aligned}
&4C\eta^{-Q} \sigma a^2 + 4C\eta^{-Q} + 4C_2 a^2 \eta^{2\alpha} \\
&\leqslant a^2 \Big(4C\eta + 4C_2 \eta^{2\alpha} \Big) + 4C\eta^{-Q} \\
&= a^2 \Big(4C\eta + 4C_2 \eta^{2\alpha} \Big) + 2 \, M^2 \Big(\varepsilon^2 - 2C\eta - 2C_2 \eta^{2\alpha} \Big) \\
&\leqslant a^2 \Big(4C\eta + 4C_2 \eta^{2\alpha} \Big) + 2 \, a^2 \Big(\varepsilon^2 - 2C\eta - 2C_2 \eta^{2\alpha} \Big) \\
&= 2\varepsilon^2 a^2,
\end{aligned}$$

which proves (3.17), choosing ε small enough. Note that, by Step 2, $\eta \leqslant 1/2$, thus $\sigma_0 < 1$.
Step 5: conclusion of the proof. In order to complete the proof of Proposition 3.2, we distinguish two cases. First, we suppose that

$$|\mathbf{q}| \leqslant \frac{a}{4}.$$

Then, by (3.16) and (3.17), we conclude that

$$\fint_{B_\eta} |\nabla_{\mathbb{G}} u(x)|^2 \, dx \leqslant \frac{a^2}{8} + 2\frac{a^2}{16} = \frac{a^2}{4},$$

from which we immediately deduce the first alternative (3.4).

Otherwise, we assume, recalling (3.8),

$$\frac{a}{4} < |\mathbf{q}| \leqslant C_0\, a,$$

and therefore, by (3.15) and (3.17), the second alternative (3.5) directly follows. □

Now, we show that the alternative (3.5) in Proposition 3.2 can be "improved" when ε and σ are sufficiently small. This result is the counterpart of Lemma 2.3 in [18] in the more general setting of Carnot groups of step two.

Lemma 3.3 *Let $u \in HW^{1,2}(B_1)$ be such that $u \geqslant 0$ a.e. in B_1 and*

$$J(u, B_1) \leqslant (1+\sigma) J(v, B_1) \tag{3.18}$$

for all $v \in HW^{1,2}(B_1)$ such that $u - v \in HW_0^{1,2}(B_1)$.

Let

$$a := \left(\fint_{B_1} |\nabla_{\mathbb{G}} u(x)|^2 \, dx \right)^{1/2}$$

and suppose that

$$a_1 \geqslant a \geqslant a_0 > 0, \tag{3.19}$$

for some suitable $a_1 \geqslant a_0$. Assume also

$$\left(\fint_{B_1} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2 \, dx \right)^{1/2} \leqslant \varepsilon a, \tag{3.20}$$

for some constant horizontal section $\mathbf{q} : \mathbb{G} \to H\mathbb{G}$ such that

$$\frac{a}{8} < |\mathbf{q}| \leqslant 2C_0 a, \tag{3.21}$$

where $C_0 > 0$ is the universal constant in Proposition 3.2.

There exists $\alpha_0 \in (0, 1]$, only depending on Q, such that, given $0 < \alpha < \alpha_0$, there exist $\rho = \rho(\alpha, Q) > 0$, $\varepsilon_0 = \varepsilon_0(Q, \alpha, a_0, a_1) > 0$, and $c_0 = c_0(Q, \alpha, a_0, a_1) > 0$, such that if

$$\varepsilon \leqslant \varepsilon_0 \quad \text{and} \quad \sigma \leqslant c_0 \varepsilon^2,$$

then

$$\left(\fint_{B_\rho} |\nabla_{\mathbb{G}} u(x) - \widetilde{\mathbf{q}}(x)|^2 \, dx \right)^{1/2} \leqslant \rho^\alpha \varepsilon a, \tag{3.22}$$

where $\widetilde{\mathbf{q}} : \mathbb{G} \to H\mathbb{G}$ is a constant horizontal section, see (3.1), for some suitable $\widetilde{q} = (\widetilde{q}_1, \ldots, \widetilde{q}_{m_1}) \in \mathbb{R}^{m_1}$ such that

$$|\mathbf{q} - \widetilde{\mathbf{q}}| \leqslant \widetilde{C} \varepsilon a. \tag{3.23}$$

for some universal constant $\widetilde{C} > 0$.

Proof of Lemma 3.3 We divide the proof into several steps.

Step 1: energy estimates for the $\mathbb{G}$ -harmonic replacement and comparison of energies. Let us set $\tau := \frac{1}{10\Lambda_{\mathbb{G}}^2} < 1$, where $\Lambda_{\mathbb{G}} \geqslant 1$ is the structural constant given by (2.6) only depending on $\mathbb{G}$. Let $\bar{v}$ denote the $\mathbb{G}$-harmonic replacement of u in B_τ and v be defined as

$$v := \begin{cases} \bar{v} & \text{in } B_\tau, \\ u & \text{in } B_1 \setminus B_\tau. \end{cases} \tag{3.24}$$

We first note that $v \in HW^{1,2}(B_1)$ and $u - v \in HW_0^{1,2}(B_1)$, thus, by hypothesis (3.18), we have

$$J(u, B_1) \leqslant (1+\sigma) J(v, B_1),$$

which leads to

$$\begin{aligned} J(u, B_\tau) + J(u, B_1 \setminus B_\tau) &= J(u, B_1) \\ &\leqslant (1+\sigma)\, J(v, B_1) \\ &= J(v, B_1) + \sigma J(v, B_1) \\ &= J(v, B_\tau) + J(v, B_1 \setminus B_\tau) + \sigma J(v, B_1) \\ &= J(v, B_\tau) + J(u, B_1 \setminus B_\tau) + \sigma J(v, B_1), \end{aligned}$$

namely,

$$J(u, B_\tau) \leqslant J(v, B_\tau) + \sigma J(v, B_1).$$

By definition of J in (1.2), this reads as

$$\begin{aligned} \int_{B_\tau} |\nabla_{\mathbb{G}} u(x)|^2 \, dx + |\{u > 0\} \cap B_\tau| &\leqslant \int_{B_\tau} |\nabla_{\mathbb{G}} v(x)|^2 \, dx + |\{u > 0\} \cap B_\tau| + \sigma J(v, B_1) \\ &\leqslant \int_{B_\tau} |\nabla_{\mathbb{G}} v(x)|^2 \, dx + |B_\tau| + \sigma J(v, B_1), \end{aligned}$$

which, since $u \geqslant 0$ a.e. in B_1 by assumption, yields

$$\int_{B_\tau} \left(|\nabla_{\mathbb{G}} u(x)|^2 - |\nabla_{\mathbb{G}} v(x)|^2 \right) dx \leqslant |\{u = 0\} \cap B_\tau| + \sigma J(v, B_1). \tag{3.25}$$

Moreover, by definition of v in (3.24), we have

$$\begin{aligned} J(v, B_1) &= \int_{B_1} \left(|\nabla_{\mathbb{G}} v(x)|^2 + \chi_{\{v>0\}}(x) \right) dx \\ &\leqslant \int_{B_\tau} |\nabla_{\mathbb{G}} v(x)|^2 \, dx + \int_{B_1 \setminus B_\tau} |\nabla_{\mathbb{G}} v(x)|^2 \, dx + |B_1| \\ &= \int_{B_\tau} |\nabla_{\mathbb{G}} \bar{v}(x)|^2 \, dx + \int_{B_1 \setminus B_\tau} |\nabla_{\mathbb{G}} u(x)|^2 \, dx + |B_1| \\ &\leqslant \int_{B_1} |\nabla_{\mathbb{G}} u|^2 \, dx + |B_1| \leqslant |B_1| \, (a^2 + 1), \end{aligned} \tag{3.26}$$

where in the second inequality we use the fact that $\bar{v}$ is the $\mathbb{G}$-harmonic replacement of u in B_τ and therefore $\bar{v}$ minimizes the Dirichlet energy on B_τ.

Finally, recalling (3.13), by (3.25) and (3.26) we obtain that

$$\begin{aligned}\int_{B_\tau} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \, dx &= \int_{B_\tau} \left(|\nabla_{\mathbb{G}} u(x)|^2 - 2\langle \nabla_{\mathbb{G}} u, \nabla_{\mathbb{G}} v\rangle + |\nabla_{\mathbb{G}} v(x)|^2 \right) dx \\ &= \int_{B_\tau} \left(|\nabla_{\mathbb{G}} u(x)|^2 - 2\langle \nabla_{\mathbb{G}} (u-v), \nabla_{\mathbb{G}} v\rangle - |\nabla_{\mathbb{G}} v(x)|^2 \right) dx \\ &= \int_{B_\tau} \left(|\nabla_{\mathbb{G}} u(x)|^2 - |\nabla_{\mathbb{G}} v(x)|^2 \right) dx \\ &\leqslant |\{u=0\} \cap B_\tau| + \sigma J(v, B_1) \\ &\leqslant |\{u=0\} \cap B_\tau| + |B_1| \sigma (a^2+1). \end{aligned} \tag{3.27}$$

Step 2: measure estimates for the zero level set. Now, we claim that

$$|B_\tau \cap \{u=0\}| \leqslant C_1 \varepsilon^{2+\delta}, \tag{3.28}$$

for some $C_1 > 0$ and $\delta > 0$.

Step 2.1: comparison with a $\mathbb{G}$-affine function. To prove (3.28), we consider a function $\ell : \mathbb{G} \to \mathbb{R}$ defined by

$$\ell(x) := b + \langle \mathbf{q}(x), \pi_x(x) \rangle = b + \sum_{j=1}^{m_1} q_j x_j, \qquad b := \fint_{B_1} u(x)\, dx. \tag{3.29}$$

We remark that

$$(u-\ell)_{B_1} := \fint_{B_1} \big(u(x) - \ell(x)\big)\, dx = b - \left(b + \fint_{B_1} \sum_{j=1}^{m_1} q_j x_j \, dx \right) = 0,$$

where the last equality is a consequence of the symmetry with respect to the identity element of the Carnot-Carathéodory ball B_1. Then, by the Poincaré inequality (see, e.g., [33]) we have

$$\left\| u - \ell - (u-\ell)_{B_1} \right\|_{L^2(B_1)} = \|u-\ell\|_{L^2(B_1)} \leqslant C \left\| \nabla_{\mathbb{G}} (u-\ell) \right\|_{L^2(B_1)},$$

for some $C > 0$ universal.

Since by Proposition 2.3, $\nabla_{\mathbb{G}} \ell = \mathbf{q}$, this, together with hypothesis (3.20), leads to

$$\fint_{B_1} |u(x) - \ell(x)|^2 \, dx \leqslant C \varepsilon^2 a^2. \tag{3.30}$$

Finally, we point out that, since by assumption $u \geqslant 0$, it holds $\ell^- \leqslant |u-\ell|$, so by (3.30), we obtain

$$\fint_{B_1} (\ell^-(x))^2 \, dx \leqslant C \varepsilon^2 a^2, \tag{3.31}$$

for some $C > 0$ universal.

Step 2.2: lower bounds on the $\mathbb{G}$-affine function. We want to show that if ε is sufficiently small, then

$$\ell \geqslant c_1 a \quad \text{in } B_\tau, \tag{3.32}$$

for some $c_1 > 0$. To prove this, we argue by contradiction assuming that

$$\min_{x \in \overline{B_\tau}} \ell(x) < ca$$

for any $c > 0$. We notice that, for every $x \in B_\tau$, recalling $\tau = \frac{1}{10\Lambda_{\mathbb{G}}^2}$ and (2.6), we get

$$|\ell(x) - b| = |\langle \mathbf{q}(x), \pi_x(x)\rangle| \leqslant |\mathbf{q}| \left|x^{(1)}\right| \leqslant |\mathbf{q}|\, |x|_{\mathbb{G}} \leqslant |\mathbf{q}|\, \Lambda_{\mathbb{G}}\, d_c(e,x) \leqslant \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}},$$

where $|x^{(1)}|$ denotes the Euclidean norm on $\mathbb{R}^{m_1}$. Consequently,

$$-\frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}} \leqslant \ell(x) - b \leqslant \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}} \qquad \text{for any } x \in B_\tau,$$

and therefore

$$ca > \min_{x \in \overline{B_\tau}} \ell(x) \geqslant b - \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}},$$

which gives

$$b \leqslant ca + \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}}, \tag{3.33}$$

for any $c > 0$. Now, taking into account the usual identifications given by exponential coordinates, let us consider

$$\begin{aligned}\mathcal{B} := \Big\{ x = \left[x^{(1)}, x^{(2)}\right] \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \equiv \mathbb{G} :\ & x_j^{(1)} = -\frac{tq_j}{|\mathbf{q}|} + \eta_j,\ \text{for } j = 1, \dots, m_1;\\ & x_i^{(2)} = \xi_i,\ \text{for } i = m_1 + 1, \dots, m_2,\ \text{for some } (t, \eta, \xi) \in \mathcal{A} \Big\},\end{aligned}$$

where we set

$$\begin{aligned}\mathcal{A} := \Big\{ (t, \eta, \xi) \in \mathbb{R} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} : t \in \left[\frac{1}{4\Lambda_{\mathbb{G}}}, \frac{3}{10\Lambda_{\mathbb{G}}}\right],\\ \sum_{j=1}^{m_1} \eta_j^2 \leqslant \frac{1}{100\Lambda_{\mathbb{G}}^2} \ \text{and} \sum_{i=m_1+1}^{m_2} \xi_i^2 \leqslant \frac{1}{100\Lambda_{\mathbb{G}}^4} \Big\}.\end{aligned}$$

We observe that if $x \in \mathcal{B}$, then

$$\begin{aligned} d_c(e,x)^4 &\leqslant \Lambda_{\mathbb{G}}^4 |x|_{\mathbb{G}}^4 = \Lambda_{\mathbb{G}}^4 \Big(\sum_{j=1}^{m_1} \big(-\frac{tq_j}{|\mathbf{q}|} + \eta_j\big)^2 \Big)^2 + \Lambda_{\mathbb{G}}^4 \sum_{i=m_1+1}^{m_2} \xi_i^2 \\ &\leqslant 4\Lambda_{\mathbb{G}}^4 \Big(t^2 + \sum_{j=1}^{m_1} \eta_j^2 \Big)^2 + \Lambda_{\mathbb{G}}^4 \sum_{i=m_1+1}^{m_2} \xi_i^2 \leqslant 4\Lambda_{\mathbb{G}}^4 \Big(\frac{9}{100\Lambda_{\mathbb{G}}^2} + \frac{1}{100\Lambda_{\mathbb{G}}^2} \Big)^2 + \frac{1}{100} < 1, \end{aligned}$$

so we have

$$\mathcal{B} \subseteq B_1. \tag{3.34}$$

Furthermore, by (3.33), we have, for $x \in \mathcal{B}$,

$$\ell(x) = b - t|\mathbf{q}| + \sum_{j=1}^{m_1} q_j \eta_j \leqslant ca + \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}} - \frac{|\mathbf{q}|}{4\Lambda_{\mathbb{G}}} + \frac{|\mathbf{q}|}{10\Lambda_{\mathbb{G}}} = ca - \frac{|\mathbf{q}|}{20\Lambda_{\mathbb{G}}}.$$

which implies, using hypothesis (3.21),

$$\ell(x) \leqslant ca - \frac{a}{160\Lambda_{\mathbb{G}}}.$$

Then, taking $c \in \left(0, \frac{1}{320\Lambda_{\mathbb{G}}}\right)$, we infer that

$$\ell(x) \leqslant -\frac{a}{320\Lambda_{\mathbb{G}}}.$$

Accordingly, exploiting (3.34) into (3.31), we obtain

$$C\,|B_1|\,\varepsilon^2 a^2 \geqslant \int_{B_1} (\ell^-(x))^2\,dx \geqslant \int_{\mathcal{B}} (\ell^-(x))^2\,dx \geqslant \int_{\mathcal{B}} \left(\frac{a}{320\Lambda_{\mathbb{G}}}\right)^2\,dx \geqslant \overline{c}a^2,$$

for some positive universal constant $\overline{c}$. This establishes the desired contradiction if ε is sufficiently small, and thus the proof of (3.32) is complete.

Step 2.3: conclusion of the proof of (3.28). We can now address the completion of the proof of the measure estimate in (3.28).

Recalling the Poincaré-Sobolev inequality (see, e.g., [26, 38]), it holds

$$\left(\int_{B_1} |u(x) - \ell(x)|^{2^*}\,dx\right)^{1/2^*} \leqslant C\left(\int_{B_1} \left|\nabla_{\mathbb{G}}\big(u(x) - \ell(x)\big)\right|^2\,dx\right)^{1/2}$$

for some $C > 0$ universal, where

$$2^* := \frac{2Q}{Q-2}.$$

Then, by virtue of (3.20), (3.29), we get

$$\begin{aligned}\left(\int_{B_1} |u(x) - \ell(x)|^{2^*}\,dx\right)^{1/2^*} &\leqslant C\left(\int_{B_1} \left|\nabla_{\mathbb{G}}\big(u(x) - \ell(x)\big)\right|^2\,dx\right)^{1/2}\\ &= C\left(\int_{B_1} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2\,dx\right)^{1/2} \leqslant C\varepsilon a.\end{aligned}$$

This and (3.32) entail

$$C\varepsilon a \geqslant \left(\int_{B_\tau \cap \{u=0\}} |\ell(x)|^{2^*}\,dx\right)^{1/2^*} \geqslant c_1 a\,|B_\tau \cap \{u=0\}|^{1/2^*},$$

which means, up to renaming constants,

$$|B_\tau \cap \{u=0\}| \leqslant C\varepsilon^{2^*}. \tag{3.35}$$

We notice that

$$2^* = \frac{2Q+4-4}{Q-2} = 2 + \frac{4}{Q-2}.$$

Therefore, setting

$$\delta := \frac{4}{Q-2} > 0, \tag{3.36}$$

(3.28) immediately follows from (3.35).

Step 3: energy comparison. By (3.27) and (3.28), we have

$$\int_{B_\tau} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2\,dx \leqslant C_1\varepsilon^{2+\delta} + |B_1|\sigma(a^2+1). \tag{3.37}$$

Consequently, using (3.20) and (3.37), it holds

$$\int_{B_\tau} |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)|^2 \, dx \leqslant 2\left(\int_{B_\tau} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2 \, dx + \int_{B_\tau} |\nabla_{\mathbb{G}} v(x) - \nabla_{\mathbb{G}} u(x)|^2 \, dx\right) \leqslant C_2\varepsilon^2 a^2 + C_1\varepsilon^{2+\delta} + C_3\sigma(a^2+1), \tag{3.38}$$

up to renaming constants. Let us suppose $\sigma \leqslant c_0\varepsilon^2$, with c_0 to be made precise later. Recalling (3.38), we infer

$$\int_{B_\tau} |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)|^2 \, dx \leqslant C_2\varepsilon^2 a^2 + C_1\varepsilon^{2+\delta} + C_3 c_0\varepsilon^2(a^2+1) \leqslant C\varepsilon^2 a^2, \tag{3.39}$$

for some $C > 0$ universal.

Step 4: conclusion of the proof of Lemma 3.3. Since, by Remark 3.1, $v - \langle \mathbf{q}, \pi.\rangle$ is $\mathbb{G}$-harmonic in B_τ and

$$\big(u - \langle \mathbf{q}, \pi.\rangle\big) - \big(v - \langle \mathbf{q}, \pi.\rangle\big) = u - v \in HW_0^{1,2}(B_\tau),$$

we have that $v - \langle \mathbf{q}, \pi.\rangle$ is $\mathbb{G}$-harmonic replacement of $u - \langle \mathbf{q}, \pi.\rangle$ in B_τ, being v the $\mathbb{G}$-harmonic replacement of u in B_τ. Thus, by Theorem 1.1 in [13], it exists a constant $C_0 > 0$, depending only on Q, such that, for every $x \in B_\tau$,

$$|\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)|^2 \leqslant \sup_{B_\tau} |\nabla_{\mathbb{G}} v - \mathbf{q}|^2 \leqslant C_0 \fint_{B_\tau} |\nabla_{\mathbb{G}} v(y) - \mathbf{q}(x)|^2 \, dy \leqslant C_0\varepsilon^2 a^2, \tag{3.40}$$

where the last inequality is a consequence of hypothesis (3.20). Consequently,

$$|\nabla_{\mathbb{G}} v - \mathbf{q}| \leqslant C\varepsilon a \quad \text{in } B_\tau \tag{3.41}$$

for some constant $C > 0$ universal.

Denoting by $\bar{q}_j := X_j v(0) - q_j$ for $j = 1, \ldots, m_1$, let us define the constant horizontal section

$$\bar{\mathbf{q}}(x) := \sum_{j=1}^{m_1} \bar{q}_j X_j(x), \quad \text{for } x \in \mathbb{G}.$$

By (3.41), for all $x \in B_\tau$, we have

$$|\bar{\mathbf{q}}(x)| = |\bar{\mathbf{q}}(0)| = |\nabla_{\mathbb{G}} v(0) - \mathbf{q}(0)| \leqslant C\varepsilon a. \tag{3.42}$$

Hence, combining (3.41) and (3.42), we deduce that, for $x \in B_\tau$,

$$|\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x) - \bar{\mathbf{q}}(x)| \leqslant |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x)| + |\bar{\mathbf{q}}(x)| \leqslant C\varepsilon a, \tag{3.43}$$

up to renaming $C > 0$ universal.

By Theorem 1.3 in [13], it exists a constant $C_0 > 0$, depending only on Q, such that, for each $\rho \in (0, \tau/2)$, it holds

$$\fint_{B_\rho} |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x) - \bar{\mathbf{q}}(x)|^2 \, dx \leqslant \fint_{B_\rho} \left(C_0 \left(\frac{\rho}{\tau}\right)^\mu \|\nabla_{\mathbb{G}} v - \mathbf{q} - \bar{\mathbf{q}}\|_{L^\infty(B_\tau)}\right)^2 dx \leqslant C_2\, \rho^{2\mu}\, \varepsilon^2 a^2, \tag{3.44}$$

for some $\mu = \mu(Q) \in (0, 1]$ and $C_2 > 0$ universal, where in the last inequality we used estimate (3.43). Putting together (3.37) and (3.44), we get

$$\begin{aligned} &\fint_{B_\rho} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x) - \bar{\mathbf{q}}(x)|^2 \, dx \\ &\leqslant 2\left(\fint_{B_\rho} |\nabla_{\mathbb{G}} u(x) - \nabla_{\mathbb{G}} v(x)|^2 \, dx + \fint_{B_\rho} |\nabla_{\mathbb{G}} v(x) - \mathbf{q}(x) - \bar{\mathbf{q}}(x)|^2 \, dx\right) \\ &\leqslant 2C_1 \varepsilon^{2+\delta} \rho^{-Q} + 2C\sigma(a^2+1)\rho^{-Q} + 2C_2 \rho^{2\mu} \varepsilon^2 a^2, \end{aligned} \tag{3.45}$$

up to renaming constants.

Now, setting $\alpha_0 := \mu$, for every $\alpha \in (0, \alpha_0)$, we choose

$$\rho := \min\left\{(8C_2)^{\frac{1}{2(\alpha-\alpha_0)}}, \frac{\tau}{2}\right\}, \qquad \varepsilon_0 := \left(\frac{\rho^{2\alpha+Q} a_0^2}{8C_1}\right)^{\frac{1}{\delta}}, \quad \text{and} \quad c_0 := \frac{\rho^{2\alpha+Q} a_0^2}{8C(a_1^2+1)}. \tag{3.46}$$

With these choices, it holds, for every $\varepsilon \leqslant \varepsilon_0$ and $\sigma \leqslant c_0 \varepsilon^2$,

$$\begin{aligned} 2C_2 \rho^{2\mu} &\leqslant \frac{1}{4}\rho^{2\alpha}, \\ 2C_1 \varepsilon^{2+\delta} \rho^{-Q} &\leqslant \frac{1}{4}\rho^{2\alpha}\varepsilon^2 a^2 \\ \text{and} \quad 2C\sigma(a^2+1)\rho^{-Q} &\leqslant \frac{1}{4}\rho^{2\alpha}\varepsilon^2 a^2. \end{aligned}$$

Consequently, using (3.45), we reach

$$\fint_{B_\rho} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x) - \bar{\mathbf{q}}(x)|^2 \, dx \leqslant \frac{1}{4}\rho^{2\alpha}\varepsilon^2 a^2 + \frac{1}{4}\rho^{2\alpha}\varepsilon^2 a^2 + \frac{1}{4}\rho^{2\alpha}\varepsilon^2 a^2 \leqslant \rho^{2\alpha}\varepsilon^2 a^2,$$

which gives the desired result in (3.22) by calling $\widetilde{q}_i := q_i + \bar{q}_i$, for all $i = 1, \ldots, m_1$.

Moreover, from (3.42) we have

$$|\mathbf{q} - \widetilde{\mathbf{q}}| = |\bar{\mathbf{q}}| \leqslant C\varepsilon a,$$

which establishes (3.23). This completes the proof of Lemma 3.3. □

Next, by iterating Lemma 3.3 we obtain the following estimates.

Corollary 3.4 *Let u be an almost minimizer for J in B_1 (with constant κ and exponent β) and*

$$a := \left(\fint_{B_1} |\nabla_{\mathbb{G}} u(x)|^2 \, dx\right)^{1/2}.$$

Suppose that there exist $a_1 > a_0 > 0$ such that

$$a \in [a_0, a_1] \tag{3.47}$$

and u satisfies (3.6) and (3.20).

Then, there exist ε_0, κ_0, and $\gamma \in (0, 1)$, depending on Q, β, a_0, and a_1, such that, for every $\varepsilon \leqslant \varepsilon_0$ and $\kappa \leqslant \kappa_0 \varepsilon^2$, it holds

$$\|u - \ell\|_{C^{1,\gamma}(B_{1/2})} \leqslant C\varepsilon a \tag{3.48}$$

for a positive constant C only depending on Q and a $\mathbb{G}$-affine function ℓ of slope $\mathbf{q}$.

Moreover,

$$\|\nabla_{\mathbb{G}} u\|_{L^\infty(B_{1/2})} \leqslant \widetilde{C} a, \tag{3.49}$$

with $\widetilde{C} > 0$ *only depending on* Q.

Remark 3.5 We point out that a consequence of (3.48) in Corollary 3.4 is that, if ε is sufficiently small,

$$\nabla_{\mathbb{G}} u \neq \mathbf{0} \quad \text{in } B_{1/2}, \tag{3.50}$$

where $\mathbf{0}$ is the null horizontal section. Indeed, by (3.48), we have

$$|\nabla_{\mathbb{G}} u - \mathbf{q}| \leqslant C\varepsilon a \quad \text{in } B_{1/2},$$

which gives

$$|\nabla_{\mathbb{G}} u| \geqslant |\mathbf{q}| - C\varepsilon a \quad \text{in } B_{1/2}.$$

Finally, by (3.6), we get

$$|\nabla_{\mathbb{G}} u| \geqslant \frac{a}{4} - C\varepsilon a > 0 \quad \text{in } B_{1/2},$$

as long as ε is small enough.

Furthermore, we can conclude that

$$u > 0 \quad \text{in } B_{1/2}. \tag{3.51}$$

To check this, we suppose by contradiction that there exists a point $x_0 \in B_{1/2}$ such that $u(x_0) = 0$ (by assumption $u \geqslant 0$). Therefore, since $u \in C^{1,\gamma}(B_{1/2})$, we know that $\nabla_{\mathbb{G}} u(x_0) = 0$, which contradicts (3.50), and so (3.51) holds.

In order to prove Corollary 3.4, we recall the notion of Morrey-Campanato spaces in the setting of Carnot Groups, see [40].

Definition 3.6 (Morrey-Campanato spaces) Let $\Omega \subseteq \mathbb{G}$. For every $1 \leqslant p < +\infty$ and $\lambda \in (0, +\infty)$, a horizontal section $f = (f_1, \dots, f_{m_1}) \in L^p_{\text{loc}}(\Omega, H\mathbb{G})$ is said to be in the Morrey-Campanato space $\mathcal{E}^{\lambda,\, p}(\Omega, H\mathbb{G})$ if

$$[f]_{\mathcal{E}^{\lambda,\, p}(\Omega, H\mathbb{G})} := \sup_{B \subset \Omega} \left(\frac{1}{|B|^{1+p\lambda}} \int_B |f(y) - \mathbf{f_B}(y)|^p \, dy \right)^{1/p} < +\infty,$$

where the supremum is taken over all metric balls $B \subset \Omega$ such that $\overline{B} \subset \Omega$ and $\mathbf{f_B}$ is the constant horizontal section given by

$$\mathbf{f_B}(y) := \sum_{i=1}^{m_1} \left(\fint_B f_i(z)\, dz \right) X_i(y), \quad \text{for } y \in \Omega.$$

Remark 3.7 For every $u = (u_1, \dots, u_{m_1}) \in \mathcal{E}^{\lambda,p}(\Omega, H\mathbb{G})$, the quantity $[u]_{\mathcal{E}^{p,\lambda}(\Omega, H\mathbb{G})}$ is a seminorm in $\mathcal{E}^{\lambda,p}(\Omega, H\mathbb{G})$ and

$$\text{it is equivalent to the quantity} \quad \sup_{B \subset \Omega} \left(\frac{1}{|B|^{1+p\lambda}} \inf_{\boldsymbol{\xi}} \int_B |u(x) - \boldsymbol{\xi}(x)|^p \, dx \right)^{1/p}, \tag{3.52}$$

where the supremum is taken over all metric balls B such that $\overline{B} \subset \Omega$ and the infimum is taken over all the constant horizontal sections such that

$$\boldsymbol{\xi}(x) := \sum_{i=1}^{m_1} \xi_i X_i(x), \quad \text{for } x \in \mathbb{G} \text{ and some } \xi := (\xi_1, \dots, \xi_{m_1}) \in \mathbb{R}^{m_1}. \tag{3.53}$$

To prove (3.52), we observe that for every ball $B \subset \Omega$ such that $\overline{B} \subset \Omega$ and constant horizontal section $\boldsymbol{\xi}$ as in (3.53), by applying Jensen's Inequality, we have

$$\begin{aligned} |\mathbf{u_B} - \boldsymbol{\xi}|^p &\leqslant \left| \sum_{i=1}^{m_1} \left(\fint_B |u_i(y) - \xi_i| \, dy \right)^2 \right|^{\frac{p}{2}} \leqslant \left| \sum_{i=1}^{m_1} \fint_B |u_i(y) - \xi_i| \, dy \right|^p \\ &\leqslant C \left(\fint_B |u(y) - \boldsymbol{\xi}(y)| \, dy \right)^p \leqslant C \fint_B |u(y) - \boldsymbol{\xi}(y)|^p \, dy, \end{aligned}$$

for some constant $C > 0$ only depending on p and $\mathbb{G}$. Consequently, up to renaming $C > 0$, it occurs,

$$\begin{aligned} \int_B |u(y) - \mathbf{u_B}(y)|^p \, dy &\leqslant 2^p \int_B |u(y) - \boldsymbol{\xi}(y)|^p \, dy + 2^p |B| \, |\mathbf{u_B} - \boldsymbol{\xi}|^p \\ &\leqslant C \int_B |u(y) - \boldsymbol{\xi}(x)|^p \, dy \end{aligned}$$

for every constant horizontal section $\boldsymbol{\xi}$ as in (3.53) and for every ball $B \subset \Omega$ such that $\overline{B} \subset \Omega$, which leads to

$$\int_B |u(y) - \mathbf{u_B}(y)|^p \, dy \leqslant C \inf_{\boldsymbol{\xi}} \int_B |u(y) - \boldsymbol{\xi}(y)|^p \, dy.$$

On the other hand, for every ball $B \subset \Omega$ such that $\overline{B} \subset \Omega$, it is clear that

$$\int_B |u(y) - \mathbf{u_B}(y)|^p \, dy \geqslant \inf_{\boldsymbol{\xi}} \int_B |u(y) - \boldsymbol{\xi}(y)|^p \, dy.$$

Now, we recall a result, proved in [40], which we will use in the proof of Corollary 3.4.

Theorem 3.8 ([40] Theorem 5) *For every metric ball $B \subset \mathbb{G}$, $\lambda \in (0, 1)$, and $p \in [1, +\infty)$, one has*

$$\mathcal{E}^{\lambda,p}(B, H\mathbb{G}) = C^{0,\lambda}(B, H\mathbb{G}).$$

More precisely, it results that a function ϕ belongs to $\mathcal{E}^{\lambda,p}(B)$ if and only if ϕ is equal, $\mathcal{L}^n$-almost everywhere, to a function ψ in the Hölder class $C^{0,\lambda}(B)$. Moreover, the seminorms $[\phi]_{\mathcal{E}^{\lambda,p}(B,H\mathbb{G})}$ and $[\psi]_{C^{0,\lambda}(B,H\mathbb{G})}$ are equivalent.

We also point out the following scaling property of almost minimizers.

Lemma 3.9 *Let u be an almost minimizer for J in B_1 with constant κ and exponent β. For any $r \in (0, 1)$, let*

$$u_r(x) := \frac{u(\delta_r(x))}{r}. \tag{3.54}$$

Then, u_r is an almost minimizer for J in $B_{1/r}$ with constant κr^β and exponent β, namely,

$$J(u_r, B_\varrho(x_0)) \leqslant (1 + \kappa r^\beta \varrho^\beta) J(v, B_\varrho(x_0)), \tag{3.55}$$

for every ball $B_\varrho(x_0)$ *such that* $\overline{B_\varrho(x_0)} \subset B_{1/r}$ *and* $v \in HW^{1,2}(B_\varrho(x_0))$ *such that* $u_r - v \in HW_0^{1,2}(B_\varrho(x_0))$.

Proof Given $x_0 \in B_{1/r}$, we take ϱ and v as in the statement of Lemma 3.9 and we define

$$w(x) := rv\left(\delta_{1/r}(x)\right). \tag{3.56}$$

Then, with the notation $y_0 := \delta_r(x_0)$ and $\vartheta := r\varrho$, for all $x \in \partial B_\vartheta(y_0)$, we have $\delta_{1/r}(x) \in \partial B_\varrho(x_0)$ and therefore,

$$w(x) - u(x) = rv\left(\delta_{1/r}(x)\right) - u(x) = rv\left(\delta_{1/r}(x)\right) - ru_r\left(\delta_{1/r}(x)\right) = 0.$$

Accordingly, we can exploit w as a competitor for u, obtaining, see Definition 1.1,

$$\int_{B_{r\varrho}(y_0)} \left(|\nabla_{\mathbb{G}} u(y)|^2 + \chi_{\{u>0\}}(y)\right) dy \leqslant (1 + \kappa r^\beta \varrho^\beta) \int_{B_{r\varrho}(y_0)} \left(|\nabla_{\mathbb{G}} w(y)|^2 + \chi_{\{w>0\}}(y)\right) dy. \tag{3.57}$$

Furthermore, using, consistently with (3.54), the notation $w_r(x) := \frac{w(\delta_r(x))}{r}$, with the change of variable $x := \delta_{1/r}(y)$, we get, recalling that $y_0 := \delta_r(x_0)$,

$$\begin{aligned}\int_{B_{r\varrho}(y_0)} \left(|\nabla_{\mathbb{G}} w(y)|^2 + \chi_{\{w>0\}}(y)\right) dy &= r^Q \int_{B_\varrho(x_0)} \left(|\nabla_{\mathbb{G}} w(\delta_r(x))|^2 + \chi_{\{w>0\}}(\delta_r(x))\right) dx \\ &= r^Q \int_{B_\varrho(x_0)} \left(|\nabla_{\mathbb{G}} w_r(x)|^2 + \chi_{\{w_r>0\}}(x)\right) dx,\end{aligned}$$

and a similar identity holds true replacing w and w_r with u and u_r. Plugging this information into (3.57), we reach the desired result in (3.55) by observing that $v = w_r$ from (3.56). □

Proof of Corollary 3.4 Thanks to Lemma 3.9, without loss of generality we can suppose that

$$u \text{ is an almost minimizer for } J \text{ in } B_2 \text{ (with constant } \kappa \text{ and exponent } \beta). \tag{3.58}$$

We divide the proof into separate steps.
Step 1: iteration of Lemma 3.3. We prove that we can iterate Lemma 3.3 indefinitely with $\alpha := \beta/2$. More precisely, we show that, for all $k \geqslant 0$, there exists a constant horizontal section $\mathbf{q}_k : \mathbb{G} \to H\mathbb{G}$,

$$\mathbf{q}_k(x) := \sum_{j=1}^{m_1} q_{k,j} X_j(x), \qquad \text{for some } q_k =: (q_{k,1}, \ldots, q_{k,m_1}) \in \mathbb{R}^{m_1}, \tag{3.59}$$

such that

$$\begin{aligned} &|\mathbf{q}_k| \in \left[\frac{a}{4} - \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right)}{1-\rho^\alpha},\ C_0 a + \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right)}{1-\rho^\alpha}\right] \quad \text{for all } x \in \mathbb{G}, \\ &\left(\fint_{B_{\rho^k}} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}_k(x)|^2\, dx\right)^{1/2} \leqslant \rho^{k\alpha}\varepsilon a, \\ \text{and}\quad &\left(\fint_{B_{\rho^k}} |\nabla_{\mathbb{G}} u(x)|^2\, dx\right)^{1/2} \in \left[\frac{|\mathbf{q}|}{2}, 2|\mathbf{q}|\right], \end{aligned} \tag{3.60}$$

where $\rho \in (0, 1)$, $C_0 > 0$, and $\widetilde{C} > 0$ are universal constants.

We prove it by induction. When $k = 0$, we choose $\mathbf{q}_0 := \mathbf{q}$. Then, the desired claims follow from (3.6) and (3.20).

Now, we assume that (3.60) holds for k and we show it is true for $k+1$. Setting $r := \rho^k$ and $u_r(\cdot) := \frac{u(\delta_r(\cdot))}{r}$, by the inductive assumption we have

$$\left(\fint_{B_1} |\nabla_{\mathbb{G}} u_r(x) - \mathbf{q}_k(x)|^2\, dx\right)^{1/2} \leqslant r^\alpha \varepsilon a = \varepsilon_k a,$$

with $\varepsilon_k := r^\alpha \varepsilon = \rho^{k\alpha}\varepsilon$. Notice that the inductive assumption also yields (3.21), as long as ε is chosen conveniently small. Therefore, in view of Lemma 3.9, we can apply Lemma 3.3 on the function u_r with $\sigma := \kappa r^\beta$. We remark that the condition $\sigma \leqslant c_0\varepsilon^2$ in Lemma 3.3 translates here into $\kappa \leqslant c_0\varepsilon^2$, which is the requirement in the statement of Corollary 3.4 (by taking κ_0 there smaller or equal to c_0). Hence, we get from (3.22) and (3.23) that there exists a constant horizontal section $\mathbf{q}_{k+1}$ such that

$$\left(\fint_{B_\rho} |\nabla_{\mathbb{G}} u_r(x) - \mathbf{q}_{k+1}(x)|^2\, dx\right)^{1/2} \leqslant \rho^\alpha \varepsilon_k a \qquad \text{and} \qquad |\mathbf{q}_k - \mathbf{q}_{k+1}| \leqslant \widetilde{C}\varepsilon_k a. \tag{3.61}$$

Since $\nabla_{\mathbb{G}} u_r = \nabla_{\mathbb{G}} u(\delta_r(x))$, scaling back, we find, being $\varepsilon_k = \rho^{k\alpha}\varepsilon$,

$$\left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}_{k+1}(x)|^2\, dx\right)^{1/2} \leqslant \rho^\alpha \rho^{k\alpha}\varepsilon a = \rho^{(k+1)\alpha}\varepsilon a. \tag{3.62}$$

Moreover, using again the inductive assumption and (3.61),

$$\begin{aligned} |\mathbf{q}_{k+1}| &\leqslant |\mathbf{q}_k| + |\mathbf{q}_k - \mathbf{q}_{k+1}| \leqslant C_0 a + \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right)}{1-\rho^\alpha} + \widetilde{C}\rho^{k\alpha}\varepsilon a \\ &= C_0 a + \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right) + \widetilde{C}\rho^{k\alpha}\varepsilon a(1-\rho^\alpha)}{1-\rho^\alpha} = C_0 a + \frac{\widetilde{C}\varepsilon a\left(1-\rho^{(k+1)\alpha}\right)}{1-\rho^\alpha} \end{aligned}$$

and, similarly,

$$\begin{aligned} |\mathbf{q}_{k+1}| &\geqslant |\mathbf{q}_k| - |\mathbf{q}_k - \mathbf{q}_{k+1}| \geqslant \frac{a}{4} - \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right)}{1-\rho^\alpha} - \widetilde{C}\rho^{k\alpha}\varepsilon a \\ &= \frac{a}{4} - \frac{\widetilde{C}\varepsilon a\left(1-\rho^{k\alpha}\right) + \widetilde{C}\rho^{k\alpha}\varepsilon a(1-\rho^\alpha)}{1-\rho^\alpha} = \frac{a}{4} - \frac{\widetilde{C}\varepsilon a\left(1-\rho^{(k+1)\alpha}\right)}{1-\rho^\alpha}. \end{aligned}$$

Lastly, from (3.62), we also have

$$\begin{aligned} &\left|\left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x)|^2\, dx\right)^{1/2} - |\mathbf{q}_{k+1}|\right| \\ &= \frac{1}{|B_{\rho^{k+1}}|^{1/2}} \left|\left(\int_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x)|^2\, dx\right)^{1/2} - |\mathbf{q}_{k+1}||B_{\rho^{k+1}}|^{1/2}\right| \\ &= \frac{1}{|B_{\rho^{k+1}}|^{1/2}} \left| \|\nabla_{\mathbb{G}} u\|_{L^2(B_{\rho^{k+1}})} - \|\mathbf{q}_{k+1}\|_{L^2(B_{\rho^{k+1}})} \right| \leqslant \frac{1}{|B_{\rho^{k+1}}|^{1/2}} \|\nabla_{\mathbb{G}} u - \mathbf{q}_{k+1}\|_{L^2(B_{\rho^{k+1}})} \\ &= \left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}_{k+1}(x)|^2\, dx\right)^{1/2} \leqslant \varepsilon a, \end{aligned}$$

which yields, recalling that $\mathbf{q}_0 = \mathbf{q}$ and (3.6),

$$\begin{aligned}\left|\left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x)|^2\,dx\right)^{1/2} - |\mathbf{q}|\right| &= \left|\left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x)|^2\,dx\right)^{1/2} - |\mathbf{q}_0|\right| \\ &\leqslant \left|\left(\fint_{B_{\rho^{k+1}}} |\nabla_{\mathbb{G}} u(x)|^2\,dx\right)^{1/2} - |\mathbf{q}_{k+1}|\right| + \sum_{j=0}^{k} \big||\mathbf{q}_{j+1}| - |\mathbf{q}_j|\big| \\ &\leqslant \varepsilon a + \widetilde{C} a \sum_{j=0}^{k} \varepsilon_j \leqslant \varepsilon a + \widetilde{C}\varepsilon a \sum_{j=0}^{+\infty} \rho^{j\alpha} = \left(1 + \frac{\widetilde{C}}{1-\rho^\alpha}\right)\varepsilon a \\ &\leqslant \left(1 + \frac{\widetilde{C}}{1-\rho^\alpha}\right)\varepsilon |\mathbf{q}| \leqslant \frac{|\mathbf{q}|}{2}.\end{aligned}$$

These observations conclude the proof of the inductive step and establish (3.60).

Step 2: Morrey-Campanato estimates. We now want to exploit the Morrey-Campanato estimates of Theorem 3.8, here applied to the function $\nabla_{\mathbb{G}} u - \mathbf{q}$, with the following choices

$$p = 2, \quad B = B_{1/2}, \quad \text{and} \quad \lambda = \alpha/Q. \tag{3.63}$$

To this end, we claim that, for every $B \subset B_{1/2}$ such that $\overline{B} \subset B_{1/2}$,

$$\frac{1}{|B|^{1+2\lambda}} \inf_{\boldsymbol{\xi}} \int_B |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x) - \boldsymbol{\xi}(x)|^2\,dx \leqslant C\varepsilon^2 a^2, \tag{3.64}$$

where the infimum is taken over all the constant horizontal sections $\boldsymbol{\xi} : \mathbb{G} \to H\mathbb{G}$ as in (3.53) and $C > 0$ is a positive constant. In what follows, C may change from line to line. To prove it, we distinguish two cases, either $|B| \geqslant 1$ or $|B| \in (0, 1)$. If $|B| \geqslant 1$, we use (3.20) to get

$$|B|^{-(1+2\lambda)} \inf_{\boldsymbol{\xi}} \int_B |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x) - \boldsymbol{\xi}(x)|^2\,dx \leqslant \int_{B_{1/2}} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2\,dx \leqslant |B_1|\,\varepsilon^2 a^2,$$

which gives (3.64).

If instead $|B| \in (0, 1)$, let $k_0 = k_0(B) \in \mathbb{N}$ be such that $B \subseteq B_{\rho^{k_0}}$ and $B_{\rho^k} \subseteq B$ for all $k > k_0$. Then, by (3.60) we have, recalling (3.63), choosing $\boldsymbol{\xi} = \mathbf{q}_{k_0} - \mathbf{q}$,

$$\begin{aligned}|B|^{-(1+2\lambda)} \inf_{\boldsymbol{\xi}} \int_B |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x) - \boldsymbol{\xi}(x)|^2\,dx &\leqslant |B_{\rho^{k_0+1}}|^{-(1+2\lambda)} \int_{B_{\rho^{k_0}}} \left|\nabla_{\mathbb{G}} u(x) - \mathbf{q}_{k_0}(x)\right|^2 dx \\ &= |B_1|\,\rho^{-Q(k_0+1)(1+2\lambda)}\,|B_{\rho^{k_0}}| \fint_{B_{\rho^{k_0}}} \left|\nabla_{\mathbb{G}} u(x) - \mathbf{q}_{k_0}(x)\right|^2 dx \\ &\leqslant C|B_1|\,\rho^{-Q(k_0+1)(1+2\lambda)+Qk_0+2k_0\alpha}\,\varepsilon^2 a^2 = C|B_1|\,\rho^{-Q-2\alpha}\,\varepsilon^2 a^2,\end{aligned}$$

which implies (3.64).

Step 4: conclusion of the proof. Since, by (3.20),

$$\|\nabla_{\mathbb{G}} u - \mathbf{q}\|_{L^2(B_{1/2})} \leqslant C\varepsilon a,$$

and by (3.52) and (3.64), we have

$$[\nabla_{\mathbb{G}} u - \mathbf{q}]_{\mathcal{E}^{2,\lambda}(B_{1/2}, H\mathbb{G})} \leqslant C\varepsilon a,$$

we can apply Theorem 3.8, from which it follows that

$$\left[\nabla_{\mathbb{G}} u - \mathbf{q}\right]_{C^{0,\lambda}(B_{1/2}, H\mathbb{G})} \leqslant C\varepsilon a, \tag{3.65}$$

with $\lambda = \alpha/Q \in (0, 1)$.

Now, we define the $\mathbb{G}$-affine function $\ell : \mathbb{G} \to \mathbb{R}$ as $\ell(x) := u(e) + \langle \mathbf{q}(x), \pi_x(x)\rangle$. First, by Proposition 2.3, we have $\nabla_{\mathbb{G}} \ell = \mathbf{q}$. Also, for all $x \in B_{1/2}$, let $\delta_x := d_c(0, x)$ and $\gamma_x : [0, \delta_x] \to B_{1/2}$ a sub-unitary curve such that

$$\gamma_x(0) = e, \quad \gamma_x(\delta_x) = x, \quad \text{and} \quad \dot{\gamma}_x(t) = \sum_{i=1}^{m_1} h_i(t) X_i(\gamma_x(t)) \quad \text{for a.e. } t \in [0, \delta_x],$$

with $\sum_{j=1}^{m_1} (h_j(t))^2 \leqslant 1$ for a.e. $t \in [0, \delta_x]$ (such a γ exists thanks to Chow's Theorem, [9, Theorem 19.1.3]). Thus, we can write, according to (3.65),

$$\begin{aligned}
|u(x) - \ell(x)| &= |u(x) - u(e) - \langle \mathbf{q}(x), \pi_x(x)\rangle - \langle \mathbf{q}(e), \pi_e(e)\rangle| \\
&= \left| \int_0^{\delta_x} \frac{d}{dt} \left(u(\gamma_x(t)) - \left\langle \mathbf{q}(\gamma_x(t)), \pi_{\gamma_x(t)}(\gamma_x(t)) \right\rangle \right) dt \right| \\
&\leqslant C \int_0^{\delta_x} |\nabla_{\mathbb{G}} u(\gamma_x(t)) - \mathbf{q}(\gamma_x(t))|\, dt \leqslant C\varepsilon a,
\end{aligned}$$

which yields, by arbitrariness of $x \in B_{1/2}$,

$$\|u - \ell\|_{L^\infty(B_{1/2})} \leqslant C\varepsilon a.$$

This and (3.65) establish (3.48). □

4 Lipschitz continuity of almost minimizers and proof of theorem 1.2

We now provide the proof of Theorem 1.2.

Proof of Theorem 1.2 By virtue of Lemma 3.9, up to rescaling, we can assume that u is an almost minimizer with constant

$$\widetilde{\kappa} := \kappa s^\beta, \tag{4.1}$$

which can be made as small as we wish by a suitable choice of $s > 0$.

Let $\alpha_0 \in (0, 1]$ be the structural constant in Lemma 3.3 and define

$$\alpha := \frac{1}{2} \min\left\{\alpha_0, \frac{\beta}{2}\right\}.$$

We also consider ε_0 of Proposition 3.2 and fix $\eta \in (0, 1)$ and $M \geqslant 1$ as in Proposition 3.2 (corresponding here to the choice $\varepsilon := \varepsilon_0/2$). Let us denote by

$$a(\tau) := \left(\fint_{B_\tau} |\nabla_{\mathbb{G}} u(x)|^2\, dx \right)^{1/2}. \tag{4.2}$$

We divide the argument into separate steps.

Step 1: estimating the average. We claim that

$$a(r) \leqslant C(M, \eta)(1 + a(1)) \quad \text{for every } r \in (0, \eta], \tag{4.3}$$

for some $C(M,\eta)>0$, possibly depending on Q.

To prove it, let us consider the set $\mathcal{K}\subseteq\mathbb{N}=\{0,1,2,\dots\}$ containing all k's for which

$$a(\eta^k)\leqslant C(\eta)M+2^{-k}a(1), \tag{4.4}$$

with

$$C(\eta):=2\eta^{-Q/2}. \tag{4.5}$$

We point out that, for $k=0$, (4.4) is obvious, hence

$$0\in\mathcal{K}\neq\varnothing. \tag{4.6}$$

We then distinguish two cases, namely, whether (4.4) holds for every k (that is, $\mathcal{K}=\mathbb{N}$) or not (i.e., $\mathcal{K}\subsetneqq\mathbb{N}$).

Step 1.1: the case $\mathcal{K}=\mathbb{N}$. For every $r\in(0,\eta]$, we take $k_0\in\mathbb{N}=\mathcal{K}$ such that $\eta^{k_0+1}<r\leqslant\eta^{k_0}$. Thus, according to (4.2) and (4.4), we get

$$\begin{aligned}a(r)&\leqslant\left(\frac{1}{|B_1|\,\eta^{(k_0+1)Q}}\int_{B_{\eta^{k_0}}}|\nabla_{\mathbb{G}}u(x)|^2\,dx\right)^{1/2}=\eta^{-Q/2}a(\eta^{k_0})\leqslant\eta^{-Q/2}(C(\eta)M+a(1))\\&\leqslant\eta^{-Q/2}\max\{C(\eta)M,1\}\,(1+a(1))\leqslant C(M,\eta)(1+a(1)),\end{aligned}$$

provided $C(M,\eta)\geqslant\eta^{-Q/2}\max\{C(\eta)M,1\}$. The proof of (4.3) is thereby complete in this case.

Step 1.2: the case $\mathcal{K}\subsetneqq\mathbb{N}$. By (4.6), there exists $k_0\in\mathbb{N}$ such that $\{0,\dots,k_0\}\in\mathcal{K}$ and

$$k_0+1\notin\mathcal{K}. \tag{4.7}$$

We notice that this implies, recalling (4.2) and (4.5),

$$\eta^{-Q/2}M<C(\eta)M\leqslant C(\eta)M+2^{-(k_0+1)}a(1)<a(\eta^{k_0+1})\leqslant\eta^{-Q/2}a(\eta^{k_0}),$$

and therefore

$$a(\eta^{k_0})>M. \tag{4.8}$$

Furthermore, (4.7) also gives

$$a(\eta^{k_0+1})>C(\eta)M+2^{-(k_0+1)}a(1)\geqslant\frac{C(\eta)M+2^{-k_0}a(1)}{2}\geqslant\frac{a(\eta^{k_0})}{2}. \tag{4.9}$$

We claim that

$$\left(\fint_{B_{\eta^{k_0+1}}}|\nabla_{\mathbb{G}}u(x)-\mathbf{q}(x)|^2\,dx\right)^{1/2}\leqslant\varepsilon a(\eta^{k_0}), \tag{4.10}$$

for some constant horizontal section $\mathbf{q}:\mathbb{G}\to H\mathbb{G}$ such that

$$\frac{a(\eta^{k_0})}{4}<|\mathbf{q}|\leqslant C_0a(\eta^{k_0}),$$

being C_0 the constant given by Proposition 3.2.

To prove (4.10), we apply the dichotomy of Proposition 3.2 rescaled in the ball $B_{\eta^{k_0}}$. Specifically, we obtain from Proposition 3.2, exploiting (3.4), (3.5), and (3.6), that (4.10) holds true, unless $a(\eta^{k_0+1})\leqslant\frac{a(\eta^{k_0})}{2}$, but this contradicts (4.9). This proves (4.10).

Now, we want to apply Corollary 3.4 rescaled in the ball $B_{\eta^{k_0+1}}$, namely, with B_1 replaced by $B_{\eta^{k_0+1}}$ and a by $a(\eta^{k_0+1})$. For this purpose, we need to check that the assumptions of Corollary 3.4 are verified in this rescaled situation. First, we note that, according to (4.8) and (4.9),

$$a(\eta^{k_0+1}) \geqslant \frac{M}{2}.$$

Also, since $k_0 \in \mathcal{K}$,

$$a(\eta^{k_0+1}) \leqslant \eta^{-Q/2} a(\eta^{k_0}) \leqslant \eta^{-Q/2}\big(C(\eta)M + 2^{-k_0}a(1)\big) \leqslant \eta^{-Q/2}\big(C(\eta)M + a(1)\big).$$

These two conditions yields (3.47) in this rescaled setting with

$$a_0 := \frac{M}{2} \qquad \text{and} \qquad a_1 := \eta^{-Q/2}\big(C(\eta)M + a(1)\big). \tag{4.11}$$

Moreover, from (4.9) and (4.10), it holds

$$\left(\fint_{B_{\eta^{k_0+1}}} |\nabla_{\mathbb{G}} u(x) - \mathbf{q}(x)|^2 \, dx\right)^{1/2} \leqslant 2\varepsilon a(\eta^{k_0+1}),$$

so (3.20) is satisfied (here with 2ε in place of ε). Finally, we achieve from (4.9) and (4.10)

$$\frac{\eta^{Q/2} a(\eta^{k_0+1})}{4} \leqslant \frac{a(\eta^{k_0})}{4} < |\mathbf{q}| \leqslant C_0 a(\eta^{k_0}) \leqslant 2C_0 a(\eta^{k_0+1}),$$

which gives that (3.6) is also fulfilled (even if with different structural constants). To summarize, all the hypotheses of Corollary 3.4 in this rescaled context are satisfied.

We can thus exploit an adapted version of Corollary 3.4. This tells us that there exist ε_0 (possibly different from the one coming from Proposition 3.2) and κ_0, depending on Q, β, a_0 and a_1, such that if

$$\widetilde{\kappa} \leqslant \kappa_0 \varepsilon_0^2, \tag{4.12}$$

then we have

$$\|\nabla_{\mathbb{G}} u\|_{L^\infty(B_{\eta^{k_0+1}/2})} \leqslant \bar{C} a(\eta^{k_0}), \tag{4.13}$$

for some universal constant $\bar{C} > 0$. We stress that (4.12) holds by choosing s in (4.1) small enough, i.e., taking $s := \left(\frac{\kappa_0 \varepsilon_0^2}{2\kappa}\right)^{\frac{1}{\beta}}$. Notice, in particular, that

$$s \text{ depends on } Q, \kappa, \beta, \text{ and } \|\nabla_{\mathbb{G}} u\|_{L^2(B_1)}, \tag{4.14}$$

by virtue of (4.11).

Using (4.13), we then get, for all $r \in \left(0, \frac{\eta^{k_0+1}}{2}\right)$, recalling the definition of k_0,

$$\begin{aligned} a(r) = \left(\frac{1}{|B_r|}\int_{B_r} |\nabla_{\mathbb{G}} u(x)|^2 \, dx\right)^{1/2} &\leqslant \bar{C} a(\eta^{k_0}) \leqslant \bar{C}\big(C(\eta)M + 2^{-k_0}a(1)\big) \\ &\leqslant \bar{C}\big(C(\eta)M + a(1)\big) \leqslant C(M,\eta)(1 + a(1)), \end{aligned} \tag{4.15}$$

as long as $C(M,\eta) \geqslant \bar{C}(C(\eta)M + 1)$.

Otherwise, if $r \in \left[\frac{\eta^{k_0+1}}{2}, \eta\right]$, we take $k_r \in \mathbb{N}$ such that $\eta^{k_r+1} < r \leqslant \eta^{k_r}$. Note that

$$\frac{1}{\eta^{k_r}} \leqslant \frac{1}{r} \leqslant \frac{2}{\eta^{k_0+1}},$$

whence

$$k_r \leqslant k_0 + C_\star, \tag{4.16}$$

where $C_\star := 1 + \frac{\log 2}{\log(1/\eta)}$. We then distinguish two cases depending on whether $k_r \in \{0, \dots, k_0\}$. If this holds, then $k_r \in \mathcal{K}$ and therefore

$$a(\eta^{k_r}) \leqslant C(\eta)M + 2^{-k_r} a(1),$$

from which, recalling (4.4), we obtain

$$\begin{aligned} a(r) &\leqslant \left(\frac{1}{|B_{\eta^{k_r+1}}|} \int_{B_{\eta^{k_r}}} |\nabla_{\mathbb{G}} u(x)|^2 \, dx\right)^{1/2} = \eta^{-Q/2} a(\eta^{k_r}) \\ &\leqslant \eta^{-Q/2}\big(C(\eta)M + 2^{-k_r} a(1)\big) \leqslant C(M, \eta)(1 + a(1)). \end{aligned} \tag{4.17}$$

If instead $k_r > k_0$, by (4.16), we have

$$\begin{aligned} a(r) &\leqslant \left(\frac{1}{|B_{\eta^{k_0+C_\star+1}}|} \int_{B_{\eta^{k_0}}} |\nabla_{\mathbb{G}} u(x)|^2 \, dx\right)^{1/2} = \eta^{-Q(C_\star+1)/2} a(\eta^{k_0}) \\ &\leqslant \eta^{-Q(C_\star+1)/2}\big(C(\eta)M + 2^{-k_0} a(1)\big) \leqslant \eta^{-Q(C_\star+1)/2}\big(C(\eta)M + a(1)\big) \\ &\leqslant C(M, \eta)(1 + a(1)), \end{aligned}$$

as long as $C(M, \eta)$ is chosen sufficiently large. Combining this with (4.17) and (4.15), we reach (4.3).

Step 2: conclusions. Up to scaling and translations, we can achieve that (4.3) is true in all balls with center x_0 in $B_{1/2}$ and sufficiently small radius. In other words, for all $r \in (0, \eta]$, it holds

$$a(r, x_0) \leqslant C(M, \eta)(1 + a(1)),$$

where

$$a(r, x_0) := \left(\fint_{B_r(x_0)} |\nabla_{\mathbb{G}} u(x)|^2 \, dx\right)^{1/2}.$$

Consequently, in view of the Lebesgue Differentiation Theorem (see, e.g., [49]), recalling that $u \in HW^{1,2}(B_1)$, we obtain, for almost every $x_0 \in B_{1/2}$,

$$|\nabla_{\mathbb{G}} u(x_0)| = \lim_{r \to 0} a(r, x_0) \leqslant C(M, \eta)(1 + a(1)) = C\Big(1 + \|\nabla_{\mathbb{G}} u\|_{L^2(B_1)}\Big),$$

for some $C > 0$, and thus

$$\|\nabla_{\mathbb{G}} u\|_{L^\infty(B_{1/2})} \leqslant C\Big(1 + \|\nabla_{\mathbb{G}} u\|_{L^2(B_1)}\Big).$$

We now show that the second claim in Theorem 1.2 is also true. To this end, we can suppose that

$$\{u = 0\} \cap B_{s/100} \neq \varnothing, \tag{4.18}$$

and we consider a Lebesgue point $\bar{x} \in B_{s/100}$ for $\nabla_{\mathbb{G}} u$. Up to a left-translation, we can take that $\bar{x} = 0$ and modify (4.18) into

$$\{u = 0\} \cap B_{s/50} \neq \varnothing. \tag{4.19}$$

We claim that, in this case,

$$\mathcal{K} = \mathbb{N}. \tag{4.20}$$

Indeed, suppose this does not hold and let k_0 as before (recall (4.7)). Hence, according to (4.10), we can apply Corollary 3.4 (rescaled as above) and get, from (3.51), that $u > 0$ in $B_{s/2}$, in contradiction with (4.19). This establishes (4.20).

Therefore, in light of (4.4) and (4.20), for all $k \in \mathbb{N}$,

$$a(\eta^k) \leqslant C(\eta)M + 2^{-k} a(1),$$

which yields

$$|\nabla_{\mathbb{G}} u(\bar{x})| = |\nabla_{\mathbb{G}} u(0)| = \lim_{k \to +\infty} a(\eta^k) \leqslant \lim_{k \to +\infty} \big(C(\eta)M + 2^{-k} a(1) \big) = C(\eta)M.$$

The proof of the second claim in Theorem 1.2 is then complete by also recalling (4.14). □

Funding F.F. is partially supported by 2023-INDAM-GNAMPA-project *Equazioni completamente non lineari locali e non locali.* F.F. is supported by the PRIN research project 2022 7HX33Z - CUP J53D23003610006, "Pattern formation in nonlinear phenomena" N.F. is partially supported by 2023-INDAM-GNAMPA-project *Problemi variazionali/nonvariazionali; interazione tra metodi integrali.* E.M. is partially supported by 2023-INDAM-GNAMPA-project *Equazioni nonlocali di tipo misto e geometrico.*

Data availibility Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Acker, A.: Heat flow inequalities with applications to heat flow optimization problems. SIAM J. Math. Anal. **8**(4), 604–618 (1977)
2. Almgren, F.J., Jr.: Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints. Mem. Amer. Math. Soc. **4**(165), viii+199 (1976)
3. Alt, H.W., Caffarelli, L.A.: Existence and regularity for a minimum problem with free boundary. J. Reine Angew. Math. **325**, 105–144 (1981)
4. Alt, H.W., Caffarelli, L.A., Friedman, A.: Variational problems with two phases and their free boundaries. Trans. Amer. Math. Soc. **282**, 431–461 (1984)
5. Anzellotti, G.: On the $C^{1,\alpha}$-regularity of ω-minima of quadratic functionals. Boll. Un. Mat. Ital. C **2**(1), 195–212 (1983)
6. Aparicio, N.D., Pidcock, M.K.: On a class of free boundary problems for the Laplace equation in two dimensions. Inverse Problems **14**(1), 9–18 (1998)
7. Berestycki, H., Caffarelli, L.A., Nirenberg, L.: Uniform estimates for regularization of free boundary problems, Analysis and partial differential equations, *Lecture Notes in Pure and Appl. Math.*, vol. 122, Dekker, New York, pp. 567-619, (1990)

8. Beurling, A.: On free-boundary problems for the Laplace equation, Seminars on analytic functions, vol. I, pp. 248–263. Inst. Advanced Study, Princeton, N.J. (1957)
9. Bonfiglioli, A., Lanconelli, E., Uguzzoni, F.: Stratified Lie Groups and Potential Theory for their Sub-Laplacians. Springer, Cham (2007)
10. Caffarelli, L.A., Salsa, S.: A geometric approach to free boundary problems, Graduate Studies in Mathematics, vol. 68. American Mathematical Society, Providence, RI (2005)
11. Capogna, L., Dainelli, D., Pauls, S., Tyson, J.: An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem, Progress in Mathematics, 259. Birkhäuser Verlag, Basel (2007)
12. Capogna, L., Garofalo, N.: Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type. J. Eur. Math. Soc **5**(1), 1–40 (2003)
13. Citti, G., Mukherjee, S.: Regularity of quasi-linear equations with Hörmander vector fields of step two, *Adv. Math.*, **408**, part A, Paper No. 108593, pp. 66, (2022)
14. Corwin, L., Greenleaf, F.P.: Representations of nilpotent Lie groups and their applications, Part I: basic theory and examples, Cambridge Studies in Advanced Mathematics 18. Cambridge University Press, Cambridge (1990)
15. David, G., Engelstein, M., Toro, T.: Free boundary regularity for almost minimizers. Adv. Math. **350**, 1109–1192 (2019)
16. David, G., Toro, T.: Regularity of almost minimizers with free boundary. Calc. Var. Partial Diff. Equ. **54**(1), 455–524 (2015)
17. De Silva, D., Savin, O.: Thin one-phase almost minimizers. Nonlinear Anal. **193**, 111507, 23 (2020)
18. De Silva, D., Savin, O.: Almost minimizers of the one-phase free boundary problem. Comm. Partial Diff. Equ. **45**(8), 913–930 (2020)
19. Dipierro, S., Ferrari, F., Forcillo, N., Valdinoci, E.: Lipschitz regularity of almost minimizers in one-phase problems driven by the p-Laplace operator. Indiana Univ. Math. J. **73**(3), 813–854 (2024)
20. Dzhugan, A., Ferrari, F.: Domain variation solutions for degenerate two phase free boundary problems. Math. Eng **3**(6), . 043, 29 (2021)
21. Ferrari, F., Forcillo, N.: A counterexample to the monotone increasing behavior of an Alt-Caffarelli-Friedman formula in the Heisenberg group, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. **34**(2), 295–306 (2023)
22. Ferrari, F., Giovagnoli, D.: *Some counterexamples to Alt-Caffarelli-Friedman monotonicity formulas in Carnot groups*, *Ann. Mat. Pura Appl.*, (2024)
23. Ferrari, F., Valdinoci, E.: Density estimates for a fluid jet model in the Heisenberg group. J. Math. Anal. Appl. **382**, 448–468 (2011)
24. Folland, G.: Subelliptic estimates and function spaces on nilpotent Lie groups. Ark. Math. **13**, 161–207 (1975)
25. Folland, G. B., Stein, E. M.: *Hardy Spaces on Homogeneous Groups*, Mathematical Notes N. 28, Princeton University Press, Princeton, (1982)
26. Franchi, B., Lu, G., Wheeden, L.: Representation formulas and weighted Poincaré inequalities for Hörmander vector fields. Ann. Inst. Fourier **45**(2), 577–604 (1995)
27. Franchi, B., Lu, G., Wheeden, L.: A Relationship between Poincaré-type inequalities and representation formulas in spaces of homogeneous type. IMRN Int. Math. Res. Notices **1**, 1–14 (1996)
28. Franchi, B., Serapioni, R., Serra Cassano, F.: Rectifiability and perimeter in the Heisenberg group. Math. Ann. **321**, 479–531 (2001)
29. Franchi, B., Serapioni, R., Serra Cassano, F.: On the structure of finite perimeter sets on step 2 Carnot groups. J. Geom. Anal. **13**(3), 421–466 (2003)
30. Garofalo, N., Nhieu, D.M.: Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces. Comm. Pure Appl. Math. **49**, 1081–1144 (1996)
31. Giusti, E.: Direct methods in the calculus of variations. World Scientific Publishing Co., Inc, River Edge, NJ (2003)
32. Hörmander, L.: Hypoelliptic second-order differential equations. Acta Math. **119**, 147–171 (1969)
33. Jerison, D.: The Poincaré inequality for vector fields satisfying Hörmander's condition. Duke Math. J. **53**(2), 503–523 (1986)
34. Lacey, A.A., Shillor, M.: Electrochemical and electro-discharge machining with a threshold current. IMA J. Appl. Math. **39**(2), 121–142 (1987)
35. Le Donne, E.: A metric characterization of Carnot groups. Proc. Amer. Math. Soc. **143**, 845–849 (2015)
36. Le Donne, E.: A primer on Carnot groups: homogenous groups, CC spaces, and regularity of their isometries. Anal. Geom. Metr. Space **5**(1), 116–137 (2017)
37. Lu, G.: Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications. Revista Mat. Iberoamericana **8**, 367–439 (1992)

38. Lu, G.: The sharp Poincare inequality for free vector fields: an endpoint result. Rev. Mat. Iberoam. **10**(2), 453–466 (1994)
39. Lu, G.: Embedding theorems on Campanato-Morrey spaces for vector fields of Hormander type and applications to subelliptic PDE. C.R. Acad. Sci. Paris Sér. I Math **320**(4), 429–434 (1995)
40. Macías, R., Segovia, C.: Lipschitz functions on spaces of homogeneous type. Adv. Math. **33**, 257–270 (1979)
41. Mitchell, J.: On Carnot-Carathéeodory metrics. J. Differ. Geom. **21**, 35–45 (1985)
42. Monti, R., Serra Cassano, F.: Surface measures in Carnot-Carathéodory spaces. Calc. Var. Partial Diff. Eq. **13**, 339–376 (2001)
43. Monti, R.: Isoperimetric problem and minimal surfaces in the Heisenberg group, in *Geometric Measure Theory and Real Analysis*, pp.57-130, CRM Series, **17**, (2015)
44. Mukherjee, S., Sire, Y.: Regularity of inhomogeneous quasi-linear equations on the Heisenberg group. Anal. Theory Appl. **37**, 520–540 (2021)
45. Mukherjee, S., Zhong, X.: $C^{1,\alpha}$-regularity for variational problems in the Heisenberg group. Anal. PDE **14**(2), 567–594 (2021)
46. Petrosyan, A., Valdinoci, E.: Geometric properties of Bernoulli-type minimizers. Interfaces Free Bound. **7**(1), 55–77 (2005)
47. Petrosyan, A., Valdinoci, E.: Density estimates for a degenerate/singular phase-transition model. SIAM J. Math. Anal. **36**(4), 1057–1079 (2005)
48. Smit Vega Garcia, M.: An Overview of Almost Minimizers of Bernoulli-Type Functionals. In: Mengesha, T., Salgado, A.J. (eds) A^3N^2M : *Approximation, Applications, and Analysis of Nonlocal, Nonlinear Models*. The IMA Volumes in Mathematics and its Applications, vol 165. Springer, Cham., (2023)
49. Toledano, R.: A note on the Lebesgue differentiation theorem in spaces of homogeneous type. Real Anal. Exchange **29**(1), 335–339 (2003)
50. Valdinoci, E.: Flatness of Bernoulli jets. Math. Z. **254**(2), 257–298 (2006)
51. Varadarajan, V.S.: Lie groups, lie algebras, and their representations. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo (1974)
52. Velichkov, B.: *Regularity of the One-phase Free Boundaries*, Lecture Notes of the Unione Matematica Italiana UMILN, volume 28, (2023)