GRADIENT POTENTIAL ESTIMATES ON THE HEISENBERG GROUP

SHIRSHO MUKHERJEE AND YANNICK SIRE

Abstract. We establish pointwise estimates for the horizontal gradient of solutions to quasi-linear \( p \)-Laplacian type non-homogeneous equations with measure data in the Heisenberg Group.

Contents

1. Introduction 1
2. Preliminaries and Previous results 2
3. Estimates of the horizontal gradient 5
Acknowledgments 14
References 14

1. Introduction

In the development of non-linear potential theory, the pointwise estimates of solutions involving Wolff potentials of \( p \)-Laplacian type elliptic equations on the Euclidean spaces is well known due to Kilpeläinen-Malý [10] and complemented by gradient estimates due to [6, 12], which rely on \( C^{1,\alpha} \)-estimates of the \( p \)-Laplacian established earlier in, for instance [5, 19]. The sub-elliptic analogue of [10] was shown by Trudinger-Wang [20]. However, adequate regularity estimates of the horizontal gradient for degenerate sub-elliptic equations of \( p \)-Laplacian type on the Heisenberg group, was unavailable until [22, 16, 18, 17], in the last years. It is therefore natural to consider the associated potential estimates of the horizontal gradient and this is the purpose of the present contribution.

In this paper, we consider equations of the type

\[
-(\text{div}_H a(x, X u)) = \mu \quad \text{in } \Omega \subset \mathbb{H}^n,
\]

where \( \mu \) is a Radon measure with \( \mu(\Omega) < \infty \) and \( \mu(\mathbb{H}^n \setminus \Omega) = 0 \); hence the equation can be considered as defined in all of \( \mathbb{H}^n \). Here we denote \( X f = (X_1 f, \ldots, X_{2n} f) \) the horizontal gradient of \( f : \Omega \to \mathbb{R} \) (see Section 2).

We shall take up the following structural assumptions throughout the paper: the continuous function \( a : \Omega \times \mathbb{R}^{2n} \to \mathbb{R}^{2n} \) is assumed to be \( C^1 \) in the gradient variable and satisfies the following standard structure condition for every \( x,y \in \Omega \) and \( z, \xi \in \mathbb{R}^{2n} \),

\[
(|z|^2 + s^2)^{-\frac{p}{2}} |\xi|^2 \leq \langle D_z a(x, z) \xi, \xi \rangle \leq L(|z|^2 + s^2)^{\frac{p-2}{2}} |\xi|^2;
\]

\[
|a(x, z) - a(y, z)| \leq L'|z|(|z|^2 + s^2)^{\frac{p-2}{2}} |x - y|^{\alpha},
\]

where \( L, L' \geq 1, s \geq 0, \alpha \in (0, 1] \) and \( p \geq 2 \). The sub-elliptic \( p \)-Laplacian equation with measure data, given by

\[
-(\text{div}_H (|X u|^{p-2} X u)) = \mu,
\]

arXiv:1904.03778v1 [math.AP] 7 Apr 2019
is a prototype of the equation (1.1) with (1.2) for the case $s = 0$ and $s$ is introduced in [1.2] for regularization purposes. In order to develop the potential theory, one has to introduce the sub-elliptic analogue of the classical Wolff potentials, i.e.

\begin{equation}
W_{\beta,p}^\mu(x_0,R) := \int_0^R \left| \frac{\mu(B_\beta(x_0))}{\gamma q^{-\beta p}} \right|^{1/p} \frac{d\rho}{\rho} \quad \forall \beta \in (0,Q/p],
\end{equation}

where $Q$ is the homogeneous dimension of $\mathbb{H}^n$. Now we state our main result.

**Theorem 1.1.** Let $u \in C^1(\Omega)$ be a solution of equation (1.1), with $\mu \in L^1(\Omega)$, $p \geq 2$ and $a : \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ satisfying the structure condition (1.2). Then there exist constants $c = c(n,p,L) > 0$ and $\hat{R} = \hat{R}(n,p,L',\alpha,\text{dist}(x_0,\partial\Omega)) > 0$, such that the estimate

\begin{equation}
|xu(x_0)| \leq cW_{\beta,p}^\mu(x_0,2\hat{R}) + c\int_{B_{2\hat{R}}(x_0)} (|xu| + s)\,dx
\end{equation}

holds for any $x_0 \in \mathbb{H}^n$, whenever $B_{2\hat{R}}(x_0) \subset \Omega$ and $0 < \hat{R} \leq \hat{R}$. Furthermore, if $a(x,z)$ is independent of $x$, then (1.5) hold for any $0 < R < \frac{1}{2}\text{dist}(x_0,\partial\Omega)$.

By a standard approximation argument, one can relax the regularity assumption on the solution $u$ and on the data $\mu$. In fact, using the well-known concept of Solutions Obtained by Limiting Approximations (SOLA) one can deal with very weak solutions of (1.1) and a general measure $\mu$ with finite mass. We refer the reader to [6] for more details.

The proof of Theorem 1.1 relies on a well-known strategy introduced in past years after the work of Kilpeläinen-Malý, based on suitable comparison estimates, or in other words suitable harmonic replacements. In the present case, several adaptations to the sub-elliptic setting are in order. It is well-known fact that the regularity theory for sub-elliptic PDEs is difficult due to the lack of ellipticity of the operators under consideration. A key idea in the latest developments of higher order regularity theory for such equations is that the homogeneous sub-elliptic equation behaves like the inhomogeneous elliptic equation of the Euclidean setting. This aspect makes quantitative regularity estimates harder at the gradient level and one needs to estimate carefully extra-terms coming from commutators. An instance of this fact appears in Proposition 3.1 for the integral decay estimate, where the extra term involving $\chi$ in (3.1) appears unavoidably. Similar integral estimates have been obtained previously in the Euclidean setting in [6, 14] etc. where the homogeneous equation would yield an estimate similar to (3.31), with $\chi = 0$. However, in our case $\chi$ is non-zero and its source goes back to the Caccioppoli type estimates of [22, 16, 17], where the extra terms containing the commutator $Tu$ are locally majorized by supremum norm of gradient $Xu$ due to an integrability estimate of $Tu$ from [22]. Consequently, two scales $R$ and $\hat{R}$ appear in the crucial Lemma 3.6 contrary to the corresponding Euclidean case in [6]. Nevertheless, it does not make any difference in the pointwise estimate, which is accomplished by a double iteration in the proof of Theorem 1.1 thereby resolving the difficulty.

2. Preliminaries and Previous results

2.1. The Heisenberg Group. Here we provide the definition and properties of Heisenberg group that would be useful in this paper. For more details, we refer to [23, 4], etc. The Heisenberg Group, denoted by $\mathbb{H}^n$ for $n \geq 1$, is identified to the Euclidean space $\mathbb{R}^{2n+1}$ with the group operation

\begin{equation}
x \circ y := (x_1 + y_1, \ldots, x_{2n} + y_{2n}, t + s + \frac{1}{2}\sum_{i=1}^n (x_i y_{n+i} - x_{n+i} y_i))
\end{equation}

for every $x = (x_1, \ldots, x_{2n}, t)$, $y = (y_1, \ldots, y_{2n}, s) \in \mathbb{H}^n$. 

Thus, $\mathbb{H}^n$ with $\circ$ of (2.1) forms a non-Abelian Lie group, whose left invariant vector fields corresponding to the canonical basis of the Lie algebra, are

$$X_i = \partial_{x_i} - \frac{x_{n+i}}{2} \partial_{t_i}, \quad X_{n+i} = \partial_{x_{n+i}} + \frac{x_i}{2} \partial_t,$$

for every $1 \leq i \leq n$ and the only non zero commutator $T = \partial_t$. We have

$$(2.2) \quad [X_i, X_{n+i}] = T \quad \text{and} \quad [X_i, X_j] = 0 \forall j \neq n + i,$$

and we call $X_1, \ldots, X_{2n}$ as horizontal vector fields and $T$ as the vertical vector field.

Given any scalar function $f : \mathbb{H}^n \to \mathbb{R}$, we denote $\mathcal{X}f = (X_1 f, \ldots, X_{2n} f)$ the horizontal gradient and $\mathcal{X}\mathcal{X}f = (X_i (X_j f))_{i,j}$ as the horizontal Hessian. Also, the sub-Laplacian operator is denoted by $\Delta_H f = \sum_{j=1}^{2n} X_j X_j f$. For a vector valued function $F = (f_1, \ldots, f_{2n}) : \mathbb{H}^n \to \mathbb{R}^{2n}$, the horizontal divergence is defined as

$$\text{div}_H(F) = \sum_{i=1}^{2n} X_i f_i.$$

The Euclidean gradient of a scalar function $g : \mathbb{R}^k \to \mathbb{R}$, shall be denoted by $\nabla g = (D_1 g, \ldots, D_k g)$ and the Hessian matrix by $D^2 g$.

The Carnot-Carathéodory metric (CC-metric) is defined as the length of the shortest horizontal curves connecting two points, see [3], and is denoted by $d$. This is equivalent to the homogeneous metric, denoted as $d_{\mathbb{H}^n}(x, y) = \|y^{-1} \circ x\|_{2n}$, where the homogeneous norm for $x = (x_1, \ldots, x_{2n}, t) \in \mathbb{H}^n$ is

$$(2.3) \quad \|x\|_{2n} := \left( \sum_{i=1}^{2n} x_i^2 + |t| \right)^{\frac{1}{2}}.$$

Throughout this article we use the CC-metric balls $B_r(x) = \{ y \in \mathbb{H}^n : d(x, y) < r \}$ for $r > 0$ and $x \in \mathbb{H}^n$. However, by virtue of the equivalence of the metrics, all assertions for CC-balls can be restated to any homogeneous metric balls.

The Haar measure of $\mathbb{H}^n$ is just the Lebesgue measure of $\mathbb{R}^{2n+1}$. For a measurable set $E \subset \mathbb{H}^n$, we denote the Lebesgue measure as $|E|$. For an integrable function $f$, we denote

$$(f)_E = \int_E f dx = \frac{1}{|E|} \int_E f dx.$$

The Hausdorff dimension with respect to the metric $d$ is also the homogeneous dimension of the group $\mathbb{H}^n$, which shall be denoted as $Q = 2n + 2$, throughout this paper. Thus, for any CC-metric ball $B_r$, we have that $|B_r| = c(n)r^Q$.

For $1 \leq p < \infty$, the Horizontal Sobolev space $HW^{1,p}(\Omega)$ consists of functions $u \in L^p(\Omega)$ such that the distributional horizontal gradient $\mathcal{X}u$ is in $L^p(\Omega, \mathbb{R}^{2n})$. $HW^{1,p}(\Omega)$ is a Banach space with respect to the norm

$$(2.4) \quad \|u\|_{HW^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\mathcal{X}u\|_{L^p(\Omega, \mathbb{R}^{2n})}.$$

We define $HW^{1,p}_{\text{loc}}(\Omega)$ as its local variant and $HW^0_{\text{loc}}(\Omega)$ as the closure of $C_0^\infty(\Omega)$ in $HW^{1,p}(\Omega)$ with respect to the norm in (2.4). The Sobolev Embedding theorem has the following version in the setting of Heisenberg group, see [9, 13, 4] etc.

**Theorem 2.1** (Sobolev Inequality). Let $B_r \subset \mathbb{H}^n$ and $1 < q < Q$. For all $u \in HW^0_{\text{loc}}(B_r)$, there exists constant $c = c(n, q) > 0$ such that, we have

$$(2.5) \quad \left( \int_{B_r} |u|^{\frac{Qp}{Q-p}} dx \right)^{\frac{Q-p}{Q}} \leq c \left( \int_{B_r} |\mathcal{X}u|^q dx \right)^{\frac{1}{q}}.$$

Hölder spaces with respect to homogeneous metrics have been defined in Folland-Stein [17] and therefore, are sometimes known as Folland-Stein classes and denoted
by $\Gamma^\alpha$ or $\Gamma^{0,\alpha}$ in some literature. However, as in \cite{16} \cite{10}, here we continue to maintain the classical notation and define

$$C^{0,\alpha}(\Omega) = \{ u \in L^\infty(\Omega) : |u(x) - u(y)| \leq c d(x, y)^\alpha \ \forall \ x, y \in \Omega \}$$

for $0 < \alpha \leq 1$, which are Banach spaces with the norm

$$\|u\|_{C^{0,\alpha}(\Omega)} = \|u\|_{L^\infty(\Omega)} + \sup_{x, y \in \Omega} \frac{|u(x) - u(y)|}{d(x, y)^\alpha}.$$  

These have standard extensions to classes $C^{k,\alpha}(\Omega)$ for $k \in \mathbb{N}$, comprising functions having horizontal derivatives up to order $k$ in $C^{0,\alpha}(\Omega)$; their local counterparts are denoted as $C^{k,\alpha}_{\text{loc}}(\Omega)$. The Morrey embedding theorem is the following.

**Theorem 2.2** (Morrey Inequality). Let $B_r \subset \mathbb{H}^n$ and $q > Q$. For all $u \in H^{1,q}_{0}(B_r) \cap C(B_r)$, there exists constant $c = c(n, q) > 0$ such that, we have

$$\|u(x) - u(y)\| \leq c d(x, y)^{1-Q/q} \left( \int_{B_r} |\nabla u|^q \, dx \right)^{1/q}, \quad \forall \ x, y \in B_r.$$  

2.2. Sub-elliptic equations. Here, we enlist some of the properties and results previously known for sub-elliptic equations of the form \eqref{1.1}.

First, we recall that the structure condition \eqref{1.2} implies the monotonicity and ellipticity inequalities, as follows:

$$\langle a(x, z_1) - a(x, z_2), z_1 - z_2 \rangle \geq c(|z_1|^2 + |z_2|^2 + s^2)^{\frac{p}{p-2}} |z_1 - z_2|^2$$

for some $c = c(n, p, L) > 0$. This ensures existence and local uniqueness of weak solution $u \in H^{1,p}(\Omega)$ of equation \eqref{1.1} from the classical theory of monotone operators, see \cite{11}. We denote $u$ as the precise representative, hereafter.

The regularity and apriori estimates of the homogeneous equation corresponding to \eqref{1.1} with freezing of the coefficients, is necessary. Therefore, for any $x_0 \in \mathbb{H}^n$, we consider the equation

$$\text{div}_n a(x_0, \nabla u) = 0 \quad \text{in} \ \Omega.$$  

Now we recall the following zero-order potential estimate due to Trudinger-Wang \cite[Theorem 5.1]{20}, which is the sub-elliptic analogue of the classical Wolff potential estimate of Kilpeläinen-Malý \cite{10}.

**Theorem 2.3.** If $u$ is a weak solution of the equation \eqref{2.11} with $a(x_0, z)$ satisfying the condition \eqref{1.2}, then there exists $c = c(n, p, L) > 0$ and $R = R(n, p, L) > 0$ such that the estimate

$$|u(x_0)| \leq c W^p_{1,p}(x_0, 2R) + c \sup_{\partial B_R} u$$

holds for any $x_0 \in \Omega$ and $0 < R \leq R$, whenever $B_{2R}(x_0) \subset \Omega$.

The following theorem ensures that the weak solutions are $C^1$ and provides necessary estimates. We refer to \cite{15} \cite{16} \cite{17} for the proofs.

**Theorem 2.4.** If $u$ is a weak solution of the equation \eqref{2.11} with $a(x_0, z)$ satisfying the condition \eqref{1.2}, then $\nabla u$ is locally Hölder continuous. Moreover, there exists constants $c = c(n, p, L) > 0$ and $\beta = \beta(n, p, L) \in (0, 1/p)$ such that

$$\sup_{B_R} |\nabla u|^p \leq c (1 - \tau)^{-Q} \int_{B_R} (|\nabla u|^2 + s^2)^{\frac{p}{2}} \, dx \quad \text{for any} \ \tau \in (0, 1);$$

$$\int_{B_R} |\nabla u - (\nabla u)|_{B_{\rho}}|^p \, dx \leq c (\rho/R)^{\beta} \int_{B_R} (|\nabla u|^2 + s^2)^{\frac{p}{2}} \, dx,$$

for every concentric $B_{\rho} \subset B_R \subset \Omega$ and $0 < \rho < R$. 


In addition, the sub-elliptic reverse Hölder inequality, see [21][15], and Gehring’s lemma, implies that there exists \( \chi_0 = \chi_0(n,p,L) > 1 \) such that we have

\[
(2.15) \quad \left( \int_{B_{r/2}} (|x|^2 + s)^{\chi_0} \, dx \right)^{\frac{1}{\chi_0}} \leq c \int_{B_r} (|x|^2 + s)^p \, dx
\]

which, together with (2.13), yields

\[
(2.16) \quad \sup_{B_{r/2}} |x|^2 \leq c \left( \int_{B_{3r/4}} (|x|^2 + s)^p \, dx \right)^{\frac{1}{p}} \leq c \int_{B_r} (|x|^2 + s)^p \, dx.
\]

We end this section by recalling the notion of De Giorgi’s class of functions in this setting. This would be required for Proposition 3.1 in Section 3. Given a metric ball \( B_{\rho_0} = B_{\rho_0}(x_0) \subset \mathbb{H}^n \), the De Giorgi’s class \( DG^+(B_{\rho_0}) \) consists of functions \( v \in HW^{1,2}(B_{\rho_0}) \cap L^\infty(B_{\rho_0}) \), which satisfy the inequality

\[
(2.17) \quad \int_{B_{\rho}} |x(v - k) + |x||^2 \, dx \leq \frac{\gamma}{(\rho' - \rho)^2} \int_{B_{\rho'}} |x - k + |x||^2 \, dx + \frac{\chi^2}{n,p,L} |x(v - k) + |x||^2 + \frac{\chi^2}{n,p,L} \frac{1}{2} \chi^2 + \varepsilon
\]

for some \( \gamma, \chi, \varepsilon > 0 \), where \( A_{k,\rho}^+ = \{ x \in B_{\rho} : (v - k)^+ = \max(v - k, 0) > 0 \} \) for any arbitrary \( k \in \mathbb{R} \), the balls \( B_{\rho'}, B_{\rho} \) and \( B_{\rho_0} \) are concentric with \( 0 < \rho' < \rho < \rho_0 \). The class \( DG^+(B_{\rho_0}) \) is similarly defined and \( DG(B_{\rho_0}) = DG^+(B_{\rho_0}) \cap DG^-(B_{\rho_0}) \). All properties of classical De Giorgi class functions, also hold for these classes.

3. Estimates of the horizontal gradient

In this section, we show several comparison estimates along the lines of [10][6] ultimately leading to a pointwise estimate of the horizontal gradient. Here onwards we fix \( x_0 \in \mathbb{H}^n \) and denote \( B_{\rho} = B_{\rho}(x_0) \) for every \( \rho > 0 \).

In the following, first we prove an integral decay estimate of solutions of the equation (2.11), that is sharper than (2.14). Similar estimates have been shown in the Euclidean setting in [13][6] etc. We remark that the pointwise oscillation estimates for the gradient obtained in [22][16][17] are slightly different from that in [5][19][13], which makes the proof of the following proposition significantly shorter.

**Proposition 3.1.** Let \( B_{r_0} \subset \Omega \) and \( u \in C^1(\Omega) \) be a solution of equation (2.11). Then there exists \( \beta = \beta(n,p,L) \in (0,1/p) \) and \( c = c(n,p,L) > 0 \), such that for all \( 0 < \rho < r < r_0 \), we have

\[
(3.1) \quad \int_{B_{\rho}} |x|^2 \leq c \left( \frac{\rho}{r} \right)^{\beta} \left[ \int_{B_{\rho}} |x|^2 \, dx + \chi \rho^2 \right]
\]

with \( \chi = M(r_0)/r_0^\beta \), where \( M(r_0) = \max_{1 \leq i \leq 2n} \sup_{B_{r_0}} |X_i u| \).

**Proof.** Given \( B_{\rho} \subset \Omega \), let us denote \( M(\rho) = \max_{1 \leq i \leq 2n} \sup_{B_{\rho}} |X_i u| \) and

\[
(3.2) \quad \omega(\rho) = \max_{1 \leq i \leq 2n} \sup_{B_{\rho}} |X_i u| \quad \text{and} \quad I(\rho) = \int_{B_{\rho}} |x|^2 \, dx
\]

for every \( 0 < \rho < r_0 \). Hence, note that \( \omega(\rho) \leq 2M(\rho) \). Now, we recall the oscillation lemma previously proved, see [17] Lemma 4.14, that there exists \( s = s(n,p,L) \geq 0 \) such that for every \( 0 < \rho < r \leq r_0/16 \), we have

\[
(3.3) \quad \omega(r) \leq (1 - 2^{-s}) \omega(8r) + 2^s M(r_0) \left( \frac{r}{r_0} \right)^\beta,
\]

for some \( \beta = \beta(n,p,L) \in (0,1/p) \). A standard iteration on (3.3), see for instance [8] Lemma 7.3, implies that for every \( 0 < \rho < r \leq r_0 \), we have

\[
(3.4) \quad \omega(\rho) \leq c \left( \frac{\rho}{r} \right)^\beta \omega(r) + \chi \rho^\beta = c \left( \frac{\rho}{r} \right)^\beta \left[ \omega(r) + \chi \rho^\beta \right]
\]
where \( \chi = M(r_0)/r_0^2 \) and \( c = c(n, p, L) > 0 \). If \( \vartheta \leq \delta r \) for some \( \delta \in (0, 1) \), it is easy to see from \([3.4]\), that for some \( c = c(n, p, L) > 0 \), we have

\[
I(\vartheta) \leq c \omega(\vartheta) \leq c \delta^{-\beta} \left( \frac{\vartheta}{r} \right)^\beta [\omega(\delta r) + \chi r^\beta].
\]

Now we claim that, there exists \( \delta = \delta(n, p, L) \in (0, 1) \) such that, the inequality

\[
\omega(\delta r) \leq c[I(r) + \chi r^\beta]
\]

holds for some \( c = c(n, p, L) > 0 \). Then \([3.5]\) and \([3.6]\) together, yields \([3.1]\); hence proving the claim \([3.6]\) is enough to complete the proof.

To this end, let us denote \( r' = \delta r \), where \( \delta \in (0, 1) \) is to be chosen later. Notice that, to prove the claim \([3.6]\), we can make the apriori assumption:

\[
\omega(r) \geq M(r_0)(r/r_0)\alpha \quad \text{with} \quad \alpha = 1/p \quad \text{for} \quad p \geq 2, \quad \text{and} \quad \alpha = 1/2 \quad \text{for} \quad 1 < p < 2,
\]

since, otherwise \([3.6]\) holds trivially with \( \beta = \alpha \). Now, we consider the following complementary cases. This is very standard for elliptic estimates, see \([5, 19, 14, 10]\).

**Case 1:** For at least one index \( l \in \{1, \ldots, 2n\} \), we have either

\[
|B_{4r'} \cap \left\{ X_l u < \frac{M(4r')}{4} \right\}| \leq \theta |B_{4r'}| \quad \text{or} \quad |B_{4r'} \cap \left\{ X_l u > -\frac{M(4r')}{4} \right\}| \leq \theta |B_{4r'}|.
\]

It has been shown in \([10, 17]\) that under assumption \([3.7]\), if Case 1 holds with choice of a small enough \( \theta = \theta(n, p, L) > 0 \), then \( X_l u \in DG(B_{2r'}) \) for every \( i \in \{1, \ldots, 2n\} \). Then, the standard local boundedness estimates of De Giorgi class functions \([3, \text{Theorem 7.2 and 7.3}] \); follow; the fact that \( X_l u \) belongs to \( DG^+(B_{2r'}) \) and \( DG^-(B_{2r'}) \), yields the following respective estimates for any \( \vartheta < M(r') \):

\[
\sup_{B_{4r'}}(X_l u - \vartheta) \leq c \left[ \int_{B_{2r'}} (X_l u - \vartheta)^+ \, dx + \chi r'^\beta \right],
\]

\[
\sup_{B_{4r'}}(\vartheta - X_l u) \leq c \left[ \int_{B_{2r'}} (\vartheta - X_l u)^+ \, dx + \chi r'^\beta \right],
\]

for every \( i \in \{1, \ldots, 2n\} \). Adding \([3.8]\) and \([3.9]\) with \( \vartheta = (X_l u)_{B_{4r'}} \), we get

\[
osc_{B_{4r'}} X_l u \leq c \left[ \int_{B_{2r'}} |X_l u - (X_l u)_{B_{4r'}}| \, dx + \chi r'^\beta \right] \leq c[I(r) + \chi r^\beta]
\]

for some \( c = c(n, p, L) > 0 \) and \( \delta < 1/2 \), which further implies \([3.6]\) for this case.

**Case 2:** With \( \theta = \theta(n, p, L) > 0 \) as in Case 1, for every \( i \in \{1, \ldots, 2n\} \), we have

\[
|B_{4r'} \cap \left\{ X_l u < \frac{M(4r')}{4} \right\}| > \theta |B_{4r'}| \quad \text{and} \quad |B_{4r'} \cap \left\{ X_l u > -\frac{M(4r')}{4} \right\}| > \theta |B_{4r'}|.
\]

First, we notice that the above assertions respectively imply \( \inf_{B_{4r'}} X_l u \leq \frac{M(4r')}{4} \) and \( \sup_{B_{4r'}} X_l u \geq -\frac{M(4r')}{4} \) for every \( i \in \{1, \ldots, 2n\} \). These further imply that

\[
\omega(4r') \geq M(4r') - M(4r')/4 = 3M(4r')/4.
\]

Now, let us denote \( L = \max_{1 \leq i \leq 2n} \|X_k u\|_{B_{r'}} = \|X_k u\|_{B_{r'}} \) for some \( k \in \{1, \ldots, 2n\} \). Then note that, if \( L > 2\omega(4r') \) then using \([3.10]\), we have

\[
\|X_k u\|_{B_{r'}} \geq 2\omega(4r') - M(4r') \geq M(4r')/2 \quad \text{in} \quad B_{4r'},
\]

which, together with the choice of \( \delta < 1/4 \), further implies

\[
I(r) \geq c(n) \int_{B_{4r'}} |X_k u - (X_k u)_{B_{4r'}}| \, dx \geq \frac{c(n)}{2} M(4r') \geq \frac{c(n)}{4} \omega(4r').
\]
If $L \leq 2\omega(4r') = 2\omega(4\delta r)$ then, we choose $\delta < 1/8$ so that using $\omega(r/2) \leq 2M(r/2)$ and (2.16) i.e. $M(r/2) \leq c\int B_r |\nabla u| dx$ respectively on (3.4), we obtain

$$\omega(4\delta r) \leq c(8\delta)^3[\omega(r/2) + \chi r^3] \leq c\delta^3 \left[ \int_{B_r} |\nabla u| dx + \chi r^3 \right]$$

$$\leq c_1\delta^3[I(r) + L + \chi r^3] \leq c_1\delta^3[I(r) + 2\omega(4\delta r) + \chi r^3]$$

for some $c_1 = c_1(n,p,L) > 0$, where the second last inequality of the above is a consequence of triangle inequality and the definition of $I$ and $L$. Now we make a further reduction of $\delta$, such that $2c_1\delta^3 < 1$, so that (3.12) imply

$$\omega(4\delta r) \leq \frac{c_1\delta^3}{1 - 2c_1\delta^3}[I(r) + \chi r^3].$$

Thus (3.11) and (3.13) together shows that (3.6) holds for Case 2, as well. Therefore, we have shown that claim (3.6) holds for both cases and the proof is finished. $\square$

### 3.1. Comparison estimates.

Here, we prove certain comparison estimates that are essential for the proof of Theorem 1.1 by localizing the equations (1.1) and (2.11). Here onwards, we fix $R > 0$ such that $B_{2R} \subset \Omega$.

Letting $u \in C^1(\Omega)$ as a solution of (1.1), we consider the Dirichlet problem

$$\begin{cases}
div_a a(x,\nabla u) = 0 & \text{in } B_{2R} \\
w - u \in HW^1_0(B_{2R}).
\end{cases}$$

The following is the first comparison lemma where the density of the Wolff potential (1.4) appears in the estimates. The proof is similar to that of [6], see also [1].

**Lemma 3.2.** Given a solution of equation (1.1) $u \in C^1(\Omega)$, if $w \in HW^1_0(B_{2R})$ is a weak solution of the equation (3.14) and $p \geq 2$, then there exists $c = c(n,p,L) > 0$ such that

$$\int_{B_{2R}} |\nabla w - \nabla u| dx \leq c \left( \frac{\mu(B_{2R})}{R^{Q-1}} \right)^{1/p}.$$

**Proof.** By testing equation (3.14) with $\varphi \in HW^1_0(B_{2R})$ and using equation (1.1), we have the weak formulation

$$\int_{B_{2R}} (a(x,\nabla u) - a(x,\nabla w), \nabla \varphi) dx = \int_{B_{2R}} \varphi \, d\mu$$

which we estimate with appropriate choices of $\varphi$, in order to show (3.15).

First, we assume $2 \leq p \leq Q$. For any $j \in \mathbb{N}$, we denote the following truncations

$$\psi_j = \max \left\{ - \frac{j}{R^3}, \min \left\{ - \frac{m - w}{m}, \frac{j}{R^3} \right\} \right\}, \quad \varphi_j = \max \left\{ - \frac{1}{R^3}, \min \left\{ - \frac{m - w}{m} - \psi_j, \frac{1}{R^3} \right\} \right\},$$

where the scaling constants $m, \gamma \geq 0$ are to be chosen later. Notice that, for each $j \in \mathbb{N}$, we have $|\varphi_j| \leq 1/R^3$ and $X \varphi_j = \frac{1}{m} (Xu - Xw) 1_{E_j}$ where

$$E_j = \{ mj/R^3 < |u - w| \leq m(j + 1)/R^3 \}.$$

Thus, taking $\varphi = \varphi_j$ in (3.16), it is easy to obtain

$$\int_{B_{2R} \cap E_j} |X w - X u|^p dx \leq \frac{cm}{R^q} \mu(B_{2R}).$$
Now, we estimate the whole integral using (3.19) and (3.18), as follows.

\[(3.18)\]

\[|w| = \int B_{2R} |x w - x u| dx \leq |E_j|^{\frac{1}{p}} \left( \int_{B_{2R} \cap E_j} |x w - x u|^p dx \right)^{\frac{1}{p}} \]

\[\leq c |E_j|^{\frac{1}{p}} (m/R^*)^\frac{1}{p} |\mu|(B_{2R})^\frac{1}{p} \]

\[\leq c (m/R^*)^\frac{1}{p} |\mu|(B_{2R})^\frac{1}{p} \left[ \frac{1}{(m_j/R^*)^\alpha} \int_{B_{2R} \cap E_j} |u - w|^\kappa dx \right]^{\frac{p-1}{p}} \]

with \(\kappa = Q/(Q - 1)\), where the last inequality of the above follows from the fact that \(|u - w|^\kappa > (m_j/R^*)^\kappa\) in \(E_j\). Also from (3.17), note that for any \(N \in \mathbb{N}\),

\[(3.19)\]

\[\int_{B_{2R} \cap \{|u - w| \leq mN/R^*\}} |x w - x u|^p dx = \sum_{j=0}^{N-1} \int_{B_{2R} \cap E_j} |x w - x u|^p dx \leq c m^{\alpha/(p - 1)} N |\mu|(B_{2R}).\]

Now, we estimate the whole integral using (3.19) and (3.18), as follows.

\[\int_{B_{2R}} |x w - x u| dx = \int_{B_{2R} \cap \{|u - w| \leq mN/R^*\}} |x w - x u| dx \]

\[+ \int_{B_{2R} \cap \{|u - w| > mN/R^*\}} |x w - x u| dx \]

\[\leq |B_{2R}|^{\frac{p-1}{p}} \left( \int_{B_{2R} \cap \{|u - w| \leq mN/R^*\}} |x w - x u|^p dx \right)^{\frac{1}{p}} \]

\[+ \sum_{j=N}^{\infty} \int_{B_{2R} \cap E_j} |x w - x u| dx \]

\[\leq c (m/R^*)^\frac{1}{p} |\mu|(B_{2R})^\frac{1}{p} \left[ |B_{2R}|^{\frac{p-1}{p}} N^\kappa \right] \]

\[+ \sum_{j=N}^{\infty} \left[ \frac{1}{(m_j/R^*)^\alpha} \int_{B_{2R} \cap E_j} |u - w|^\kappa dx \right]^{\frac{p-1}{p}} \]

Using Sobolev inequality (2.5) on the second term of the above, we obtain

\[\int_{B_{2R}} |x w - x u| dx \leq c (m/R^*)^\frac{1}{p} |\mu|(B_{2R})^\frac{1}{p} |B_{2R}|^{\frac{p-1}{p}} N^\kappa \]

\[+ c (m/R^*)^\frac{1}{p} - \frac{\kappa(p-1)}{p} |\mu|(B_{2R})^\frac{1}{p} \epsilon(N)^\frac{1}{p} \left( \int_{B_{2R}} |x u - x w| dx \right)^{\frac{p-1}{p}} \]

where \(\epsilon(N) = \sum_{j=N}^{\infty} 1/j^{\kappa(p-1)}\), \(\kappa = Q/(Q - 1)\) and \(c = c(n, p, L) > 0\).

Now, first we consider the case \(p < Q\), so that we have \(\kappa(p - 1)/p < 1\). Then, by applying Young’s inequality, we obtain

\[\int_{B_{2R}} |x w - x u| dx \leq c \left( \frac{m}{R^*} \right)^\frac{1}{p} |\mu|(B_{2R})^\frac{1}{p} |B_{2R}|^{\frac{p-1}{p}} N^\kappa \]

\[+ \epsilon(N)^\frac{1}{p\kappa(p-1)} \left( \int_{B_{2R}} |x u - x w| dx \right)^{\frac{p-1}{p}} \]

for some \(c = c(n, p, L) > 0\), by using (2.9) and \(p \geq 2\). Now, using Hölder’s inequality and (3.17), we obtain
for some \( c = c(n, p, L) > 0 \). Now, we make the following choice of the scaling constants,

\[
m = |\mu|(B_{2R})^{\frac{1}{p-1}} \quad \text{and} \quad \gamma = (Q-p)/(p-1)
\]
such that the first two terms of the above are the same. Also note that, since \( p \geq 2 > 1 + 1/k \), we have \( c(p-1) > 1 \) and hence, \( \sum_{j=1}^{\infty} 1/j^{c(p-1)} = \zeta(k(p-1)) < \infty \). Thus, for some large enough \( N \in \mathbb{N} \)

\[
e(N) = \sum_{j=N}^{\infty} 1/j^{c(p-1)} < 1/2^{c(p-1)}
\]

and the last term of the estimate can be absolved in the right hand side. With these choices of \( m, \gamma, N \), we finally obtain

(3.20)

\[
\int_{B_{2R}} |xw - Xu|^p \, dx \leq c|\mu|(B_{2R})^{\frac{1}{p-1}} R^{\frac{Q-2Q+1}{p-1}}
\]

for some \( c = c(n, p, L) > 0 \), which immediately implies (3.15).

For the case of \( p = Q \), the estimate (3.20) also follows similarly with a possibly larger \( N \) and the same choices of scaling constants, i.e. \( m = |\mu|(B_{2R})^{1/(Q-1)} \) and \( \gamma = 0 \); except here we absolve the last term to the right hand side directly, without using Young’s inequality.

Now we assume the \( p \geq Q \). Here we simply choose \( \varphi = u - w \) in (3.16) and use (2.9) together with Morrey’s inequality (2.8) to obtain

\[
\int_{B_{2R}} |xw - Xu|^p \, dx \leq c \int_{B_{2R}} |u - w| \, d\mu \leq c|\mu|(B_{2R}) \sup_{B_{2R}} |u - w|
\]

\[
\leq c|\mu|(B_{2R}) R^{1 - \frac{Q}{p}} \left( \int_{B_{2R}} |xw - Xu|^p \, dx \right)^{\frac{1}{p}},
\]

which, upon using Young’s inequality, yields

(3.21)

\[
\int_{B_{2R}} |xw - Xu|^p \, dx \leq c|\mu|(B_{2R})^{\frac{p}{p-1}} R^{\frac{p-Q}{p-1}}.
\]

Then, using Hölder’s inequality and (3.21), we obtain

\[
\int_{B_{2R}} |xw - Xu| \, dx \leq |B_{2R}|^{\frac{1}{P}} \left( \int_{B_{2R}} |xw - Xu|^p \, dx \right)^{\frac{1}{p}} \leq c|\mu|(B_{2R})^{\frac{p}{p-1}} R^{\frac{Q-2Q+1}{p-1}}
\]

which, just as before, implies (3.15). Thus, the proof is finished. \( \square \)

Remark 3.3. It is evident that by using the sub-elliptic Sobolev or Morrey inequality on (3.15), we can obtain the estimate

\[
\int_{B_{2R}} |u - w| \, dx \leq c \left( \frac{|\mu|(B_{2R})}{R^{Q-p}} \right)^{\frac{1}{p-1}}
\]

where \( u \) and \( w \) are the functions stated in Lemma 3.2.

For the next comparison estimate, we require the Dirichlet problem with freezing of the coefficients. Letting \( w \in HW^{1,p}(B_{2R}) \) as weak solution of (3.14), we consider

(3.22)

\[
\begin{cases}
\text{div} \, a(x, \nabla v) = 0 & \text{in } B_R;
\vspace{2mm}
v - w \in HW^{1,p}(B_R).
\end{cases}
\]

Lemma 3.4. Given weak solution \( w \in HW^{1,p}(B_{2R}) \) of (3.14), if \( v \in HW^{1,p}(B_R) \) is the weak solution of equation (3.22), then there exists \( c = c(n, p, L) > 0 \) such that

(3.23)

\[
\int_{B_R} |xv - Xw|^p \, dx \leq cL^2 R^{2\alpha} \int_{B_R} (|xw| + s)^p \, dx.
\]
Proof. First, note that by testing equation (3.23) with \( w - v \) and using the ellipticity (2.10), it is not difficult to show the following inequality,

\[
(3.24) \quad \int_{B_R} |\nabla v|^p \, dx \leq c \int_{B_R} (|\nabla w| + s)^p \, dx,
\]

for some \( c = c(n, p, L) \); the proof is standard, see [17, Lemma 5.1] for instance. Also, testing both equations (3.14) and (3.22) with \( w - v \), we have that

\[
\int_{B_{2R}} \langle a(x, \nabla w), \nabla w - \nabla v \rangle \, dx = 0 = \int_{B_R} \langle a(x_0, \nabla v), \nabla w - \nabla v \rangle \, dx.
\]

Using the above together with (2.9) and (1.2), we obtain

\[
c \int_{B_R} \left( |\nabla w|^2 + |\nabla v|^2 + s^2 \right)^{\frac{p}{2}} |\nabla w - \nabla v|^2 \, dx
\]

\[
\leq \int_{B_R} \langle a(x_0, \nabla w) - a(x_0, \nabla v), \nabla w - \nabla v \rangle \, dx
\]

\[
= \int_{B_R} \langle a(x_0, \nabla w) - a(x, \nabla w), \nabla w - \nabla v \rangle \, dx
\]

\[
\leq cL^2 R^2 \int_{B_R} (|\nabla w|^2 + |\nabla v|^2 + s^2)^{\frac{p}{2}} |\nabla w - \nabla v| \, dx
\]

Using Young’s inequality on the last integral of the above, it is easy to get

\[
\int_{B_R} (|\nabla w|^2 + |\nabla v|^2 + s^2)^{\frac{p}{2}} |\nabla w - \nabla v|^2 \, dx \leq c(L^2 R^2)^{\frac{p}{2}} \int_{B_R} (|\nabla w|^2 + |\nabla v|^2 + s^2)^{\frac{p}{2}} \, dx.
\]

This, together with (3.24), is enough to prove (3.23). \( \square \)

Combining Lemma 3.2 and Lemma 3.4 we obtain the following comparison estimate of solution \( u \) of (1.1) and weak solution \( v \) of (3.22).

**Corollary 3.5.** Given a solution \( u \in C^1(\Omega) \) of equation (1.1), if \( w \in HW^{1,p}(B_{2R}) \) is a weak solution of the equation (3.14) and \( v \in HW^{1,p}(B_R) \) is the weak solution of equation (3.22), then there exists \( c = c(n, p, L) > 0 \) such that

\[
\int_{B_{2R}} |\nabla w - \nabla u| \, dx \leq c(1 + (L'R^\gamma)^{\frac{1}{\gamma}}) \left( \int_{B_{2R}} |\nabla w - \nabla v|^{\frac{p}{2}} \, dx \right)^{\frac{1}{p}} + c(L'R^\gamma)^{\frac{1}{\gamma}} \int_{B_{2R}} (|\nabla u| + s) \, dx.
\]

**Proof.** First, notice that Hölder’s inequality and (3.23) imply

\[
\int_{B_{2R}} |\nabla v - \nabla w| \, dx \leq c(L'R^\gamma)^{\frac{1}{\gamma}} \left( \int_{B_{2R}} (|\nabla w| + s)^p \, dx \right)^{\frac{1}{p}}.
\]

Hence, using (3.15) and (3.25), we obtain

\[
\int_{B_{2R}} |\nabla v - \nabla u| \, dx \leq \int_{B_{2R}} |\nabla w - \nabla u| \, dx + \int_{B_{2R}} |\nabla v - \nabla w| \, dx
\]

\[
\leq c \left( \frac{|\nabla (B_{2R})|}{R^{q-1}} \right)^{\frac{1}{q}} + c(L'R^\gamma)^{\frac{1}{\gamma}} \left( \int_{B_{2R}} (|\nabla w| + s)^p \, dx \right)^{\frac{1}{p}}.
\]

We estimate the last integral using reverse Hölder’s inequality and Gehring’s lemma [21] similarly as (2.15), (2.16), to obtain

\[
\left( \int_{B_{2R}} (|\nabla w| + s)^p \, dx \right)^{\frac{1}{p}} \leq \int_{B_{2R}} (|\nabla w| + s) \, dx
\]

\[
\leq \int_{B_{2R}} (|\nabla u| + s) \, dx + \int_{B_{2R}} |\nabla u - \nabla w| \, dx
\]

\[
\leq \int_{B_{2R}} (|\nabla u| + s) \, dx + c \left( \frac{|\nabla (B_{2R})|}{R^{q-1}} \right)^{\frac{1}{q}}.
\]
where the last inequality follows from (3.15). Now it is easy to see that combining (3.26) and (3.27), the proof is finished. □

3.2. Proof of Theorem 1.1 The comparison estimates of the last subsection culminate to an integral decay estimate of the horizontal gradient of the solution of equation (1.1). The pointwise estimate of $\nabla u$ shall follow thereafter, by iteration and limiting argument, thereby proving Theorem 1.1.

The following is the integral estimate of the horizontal gradient $\nabla u$ which is induced by the integral estimate (3.1) of $\nabla v$ and the previous comparison estimate.

**Lemma 3.6.** Let $u \in C^1(\Omega)$ be a solution of the equation (1.1) and let $B_{2R} \subset \Omega$ for some $R > 0$. Then there exists $\beta = \beta(n, p, L) \in (0, 1)$ and $c = c(n, p, L) > 0$ such that, for every $0 < \varrho < R \leq \tilde{R}/2$, the following estimate holds:

$$
\int_{B_{\varrho}} |\nabla u - (\nabla u)_{B_{\varrho}}| \, dx 
\leq c \left( \frac{\varrho}{R} \right)^\beta \int_{B_{2\varrho}} |\nabla u - (\nabla u)_{B_{2\varrho}}| \, dx + c \left( \frac{\varrho}{R} \right)^\beta \int_{B_{2\varrho}} (|\nabla u| + s) \, dx 
+ c \left( \frac{R}{\varrho} \right)^Q \left[ \left( 1 + (L' R^\alpha)^\frac{2}{p} \right) \left( \frac{|\mu|_{(B_{2R})}}{R^{Q-1}} \right)^{\frac{1}{2}} + (L' R^\alpha)^\frac{2}{p} \int_{B_{2\varrho}} (|\nabla u| + s) \, dx \right] 
+ c \left( \frac{R}{\varrho} \right)^\beta \left[ \left( 1 + (L' R^\alpha)^\frac{2}{p} \right) \left( \frac{|\mu|_{(B_{2R})}}{R^{Q-1}} \right)^{\frac{1}{2}} + (L' R^\alpha)^\frac{2}{p} \int_{B_{2\varrho}} (|\nabla u| + s) \, dx \right].
$$

**Proof.** Given $u \in C^1(\Omega)$ and $B_{\varrho} \subset \Omega$, we define the comparison functions $w$ and $v$ as weak solutions of equations (3.14) and (3.22), as before. Then we have

$$
\int_{B_{\varrho}} |\nabla u - (\nabla u)_{B_{\varrho}}| \, dx 
\leq 2 \int_{B_{\varrho}} |\nabla u - (\nabla v)_{B_{\varrho}}| \, dx 
+ 2 \int_{B_{\varrho}} |\nabla u - \nabla v| \, dx.
$$

(3.28)

Now, we shall estimate both terms of the right hand side of (3.28) separately.

Using $r = R$ and $r_0 = \tilde{R}/2$ in (3.11), we estimate the first term of (3.28) as follows,

$$
\int_{B_{\varrho}} |\nabla v - (\nabla v)_{B_{\varrho}}| \, dx 
\leq c \left( \frac{\varrho}{R} \right)^\beta \left[ \int_{B_{2\varrho}} |\nabla v - (\nabla v)_{B_{2\varrho}}| + (R/\tilde{R})^\beta \sup_{B_{R/2}} |\nabla v| \right] 
\leq c \left( \frac{\varrho}{R} \right)^\beta \left[ \int_{B_{2\varrho}} |\nabla u - (\nabla u)_{B_{2\varrho}}| \, dx + 2 \int_{B_{2\varrho}} |\nabla u - \nabla v| \, dx + (R/\tilde{R})^\beta \sup_{B_{R/2}} |\nabla v| \right].
$$

The second term of (3.28) is estimated simply as

$$
\int_{B_{\varrho}} |\nabla u - \nabla v| \, dx 
\leq c \left( \frac{R}{\varrho} \right)^Q \int_{B_{2\varrho}} |\nabla u - \nabla v| \, dx.
$$

Using the above two estimates in (3.28), we obtain

$$
\int_{B_{\varrho}} |\nabla u - (\nabla u)_{B_{\varrho}}| \, dx 
\leq c \left( \frac{\varrho}{R} \right)^\beta \int_{B_{2\varrho}} |\nabla u - (\nabla u)_{B_{2\varrho}}| \, dx 
+ c \left( \frac{R}{\varrho} \right)^Q \int_{B_{2\varrho}} |\nabla u - \nabla v| \, dx + (R/\tilde{R})^\beta \sup_{B_{R/2}} |\nabla v|.
$$

(3.29)
The last term is estimated using (2.16) as
\[ \sup_{B_{R/2}} |\mathbf{X}v| \leq c \int_{B_R} (|\mathbf{X}v| + s) \, dx \leq c \int_{B_R} (|\mathbf{X}u| + s) \, dx + c \int_{B_R} |\mathbf{X}u - \mathbf{X}v| \, dx, \]
which combined with (3.29), yields
\[ \int_{B_0} |\mathbf{X}u - (\mathbf{X}u)_{B_0}| \, dx \leq c \left( \frac{\theta}{R} \right)^\beta \int_{B_R} |\mathbf{X}u - (\mathbf{X}u)_{B_R}| \, dx + c \left( \frac{R}{\theta} \right) \int_{B_R} (|\mathbf{X}u| + s) \, dx \]
\[ + c \left( \frac{R}{\theta} \right)^Q \int_{B_R} |\mathbf{X}u - \mathbf{X}v| \, dx + c \left( \frac{R}{\theta} \right) \int_{B_R} |\mathbf{X}u - \mathbf{X}v| \, dx. \]

The second last term of the above is estimated from Corollary 3.5. Now note that, since $2R \leq \tilde{R}$ and $B_{2\tilde{R}} \subset \Omega$, the comparison functions $w$ and $v$ can be defined in $\{u\} + HW^{1,p}_0(B_{\tilde{R}})$ by extension and we can derive all the previous comparison estimates in the scale of $\tilde{R}$. Then we estimate the last term of the above using the $\tilde{R}$-scaled version of Corollary 3.5. Then, together with the elementary inequality
\[ \int_{B_R} |\mathbf{X}u - (\mathbf{X}u)_{B_R}| \, dx \leq 2^{Q+1} \int_{B_{2R}} |\mathbf{X}u - (\mathbf{X}u)_{B_{2R}}| \, dx, \]
the proof is finished. \(\square\)

Before the proof of Theorem 1.1, we provide the following estimate of the density of Wolff potential. We refer to [6] for the proof.

**Lemma 3.7.** Given any $h > 1, x_0 \in \mathbb{H}^n$ and $r > 0$, if $r_i = r/(2h)^i$ for every $i \in \{0, 1, 2, \ldots\}$, then for any $N \in \mathbb{N}$, we have
\[ \sum_{i=0}^{N-1} \left( \frac{|\mu|(B_{r_i}(x_0))}{r_i^{Q-1}} \right)^{\frac{1}{\alpha}} \leq \left( \frac{2^{\frac{2}{\alpha}-1}}{\log(2)} + \frac{(2h)^{\frac{Q-1}{\alpha}}}{\log(2h)} \right) W_{\frac{p}{\alpha}}^{\mu}(x_0, 2r). \]

Finally, now we are ready prove the main theorem, Theorem 1.1.

**Proof of Theorem 1.1.** In this proof, we shall fix some arbitrary $x_0 \in \mathbb{H}^n$ and denote the metric balls $B_R = B_{R}(x_0)$ and $B_{\tilde{R}} = B_{\tilde{R}}(x_0)$ as before. We shall also assume
\[ (3.30) \quad 0 < R \leq \tilde{R} \leq \hat{R} = \hat{R}(n, p, L, L', \alpha, \delta(\Omega)) \]
where $\Omega$ can be chosen as small as required.

To begin with, we consider $\hat{R} \leq \min\{\delta(\Omega), L'^{-1/\alpha}\}$ so that, we have $B_R \subset B_{\hat{R}} \subset \Omega$ and $L' R^c \leq L'^{-c} \leq 1$. We choose $h = h(n, p, L) > 1$ large enough such that we have $(c/h)^\beta \leq 1/2$ for $c = c(n, p, L) > 0$ as in Lemma 3.6. Then we apply the Lemma 3.6 with $R, \hat{R}$ in place of $2R, 2\hat{R}$ and $h = R/4h$, to obtain
\[ \int_{B_{R/4h}} |\mathbf{X}u - (\mathbf{X}u)_{B_{R/4h}}| \, dx \]
\[ \leq \frac{1}{2} \int_{B_R} |\mathbf{X}u - (\mathbf{X}u)_{B_R}| \, dx + \frac{1}{2} \left( \frac{R}{\hat{R}} \right)^\beta \int_{B_R} (|\mathbf{X}u| + s) \, dx \]
\[ + c_0 \left( \frac{|\mu|(B_R)}{R^{Q-1}} \right)^{\frac{1}{\alpha}} \int_{B_R} (|\mathbf{X}u| + s) \, dx \]
\[ + c_0 \left( \frac{|\mu|(B_{\hat{R}})}{\hat{R}^{Q-1}} \right)^{\frac{1}{\alpha}} \int_{B_{\hat{R}}} (|\mathbf{X}u| + s) \, dx \]
for some $c_0 = c_0(n, p, L) \geq 1$, which we fix temporarily. Let us denote
\[ (3.32) \quad \theta(r) = \left( \frac{|\mu|(B_r)}{r^{Q-1}} \right)^{\frac{1}{\alpha}} \quad \forall \ 0 < r < \hat{R}. \]
We run an iteration on (3.31) with the sequence \( R_i = R/(4i h)^i \) and \( \tilde{R}_i = R/(4i h)^i \) for some \( 1 < h < h \) and \( i \in \{0, 1, 2, \ldots \} \). Let us denote \( \tau = \tau(n, p, L) \in (0, 1) \) as \( \tau = (h/h)^i \) and

\[
(3.33) \quad A_i = \int_{B_{R_i}} |x - (x u)_{B_{R_i}}| \, dx, \quad K_i = \int_{B_{R_i}} |x u| \, dx, \quad \tilde{K}_i = \int_{B_{R_i}} |x u| \, dx.
\]

Notice that, substituting \( \tau = (\tilde{\tau}_i) \) for some \( 1 < R_i \)

\[
A_i \leq \frac{1}{2} A_{i-1} + \frac{\tau^{i-1}}{2} (\tilde{K}_{i-1} + s) + c_0 \left[ \theta(R_i-1) + \theta(\tilde{R}_{i-1}) \right] + c_0 (L'R_i^{0})^{\frac{n}{2}} (K_{i-1} + s) + c_0 (L'R_i^{0})^{\frac{n}{2}} (\tilde{K}_{i-1} + s).
\]

for every \( i \in \{1, 2, \ldots \} \). For any \( N \in \mathbb{N} \), we sum over \( i \in \{1, 2, \ldots, N\} \) and add \( A_0 \) to both sides, to obtain

\[
\sum_{i=0}^{N} A_i \leq A_0 + \frac{1}{2} \sum_{i=0}^{N-1} A_i + \frac{1}{2} \sum_{i=0}^{N-1} \tau^i (\tilde{K}_i + s) + c_0 \sum_{i=0}^{N-1} (\theta(R_i) + \theta(\tilde{R}_i)) + c_0 \left[ \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (K_i + s) + \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (\tilde{K}_i + s) \right].
\]

Then, iterating on the terms \( \sum_{i=0}^{N} A_i \) and adding appropriate positive terms on the right hand side of the above, we obtain

\[
\sum_{i=0}^{N} A_i \leq \sum_{i=0}^{N} \left[ A_0 + \frac{1}{2} \sum_{i=0}^{N-1} \tau^i (\tilde{K}_i + s) + c_0 \sum_{i=0}^{N-1} (\theta(R_i) + \theta(\tilde{R}_i)) \right] + c_0 \sum_{i=0}^{N-1} \left[ \frac{1}{2} \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (K_i + s) + \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (\tilde{K}_i + s) \right]
\]

\[
(3.34) \quad \leq 2A_0 + \sum_{i=0}^{N-1} \tau^i (\tilde{K}_i + s) + 2c_0 \sum_{i=0}^{N-1} (\theta(R_i) + \theta(\tilde{R}_i)) + 2c_0 \left[ \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (K_i + s) + \sum_{i=0}^{N-1} (L'R_i^{0})^{\frac{n}{2}} (\tilde{K}_i + s) \right].
\]

Now, recalling (3.33), it is easy to see that

\[
K_{N+1} = K_0 + \sum_{i=0}^{N} (K_{i+1} - K_i) \leq K_0 + \sum_{i=0}^{N} \int_{B_{R_{i+1}}} |x - (x u)_{B_{R_i}}| \, dx
\]

\[
(3.35) \quad \leq K_0 + \sum_{i=0}^{N} (4h)^Q \int_{B_{R_i}} |x - (x u)_{B_{R_i}}| \, dx = K_0 + (4h)^Q \sum_{i=0}^{N} A_i,
\]
where $h = h(n, p, L) \geq 1$, as chosen. Using (3.34) in (3.35), we get
\[
K_{N+1} \leq K_0 + (4h)^Q \left( 2A_0 + \sum_{i=0}^{N-1} \tau(\bar{K}_i + s) + 2c_0 \sum_{i=0}^{N-1} \left( \theta(R_i) + \theta(\bar{R}_i) \right) \\
+ 2c_0 \left[ \sum_{i=0}^{N-1} (L'R_i)^2 (\bar{K}_i + s) + \sum_{i=0}^{N-1} (L'R_i)^2 (\bar{K}_i + s) \right] \right)
\]
\[
\leq K_0 + cA_0 + c \sum_{i=0}^{N-1} \tau(\bar{K}_i + s) + cW_{1/p}^u(x_0, 2R) \\
+ c \sum_{i=0}^{N-1} (L'R_i)^2 (\bar{K}_i + s) + c \sum_{i=0}^{N-1} (L'R_i)^2 (\bar{K}_i + s)
\]
\[
\leq K_0 + cA_0 + c \sum_{i=0}^{N-1} \tau(\bar{K}_i + s) + cW_{1/p}^u(x_0, 2R)
\]
\[
\leq K_0 + cA_0 + c \sum_{i=0}^{N-1} \tau(\bar{K}_i + s) + cW_{1/p}^u(x_0, 2R)
\]

where we have used Lemma 3.7 to get (3.36) \[\theta(R_i), \sum_{i=0}^{N-1} \theta(\bar{R}_i) \leq cW_{1/p}^u(x_0, 2R)\] for some $c = c(n, p, L) > 0$. Now notice that (3.37)
\[
A_0 + K_0 \leq \int_{B_R} (|\mathcal{X}u| + s) \, dx \leq M
\]

where we denote
\[
\mathcal{M} = \int_{B_R} (|\mathcal{X}u| + s) \, dx + W_{1/p}^u(x_0, 2R)
\]

Note that $K_1 \leq K_0 + |K_1 - K_0| \leq c(A_0 + K_0) \leq c\mathcal{M}$. Also notice that $K_0 = K_0$ and let us choose $\tilde{h} = h/2$; hence similarly, $K_1 \leq c\mathcal{M}$. Now we run an induction assuming that $K_m, K_m \leq c\mathcal{M}$ for some $c = c(n, p, L) > 0$ for every $m = 0, 1, \ldots, N$. Then the estimate of $K_{N+1}$ of the above yields
\[
K_{N+1} \leq c\mathcal{M} \left[ 1 + \sum_{i=0}^{N-1} \tau(\bar{K}_i) + (L'R_i)^2 \right] \leq c\mathcal{M}(1 + c'(L'R_i)^2) = c(1 + c')\mathcal{M},
\]

for some $c' = c'(n, p, L) > 0$, since $\tau = (\tilde{h}/R) \in (0, 1)$ and $L'R_i \leq 1$ whenever $R \leq \bar{R}$. Thus $K_{N+1}$ also satisfies the hypothesis and we conclude that it is true for $K_m$ for every $m \in \mathbb{N}$. Also, with a further reduction of $\bar{R}$, one can replace $c'$ with $1/2$ on the above. This ensures a uniformly bounded growth of the constants and hence for some $c = c(n, p, L) > 0$, we have
\[
|\mathcal{X}u(x_0)| = \lim_{m \to \infty} K_m \leq c\mathcal{M},
\]

and the proof is finished. In addition, we note that if $a(x, z)$ is independent of $x$ then we can assume $L' = 0$ and the proof holds for any $\bar{R} > 0$ whenever $B_{2\bar{R}}(x_0) \subset \Omega$. \[\square\]

**Acknowledgments**

S.M. has been partially supported by the project “Variationaalist integraalit geometr”. Y.S. is partially supported by the Simons foundation.

**References**

[1] Lucio Boccardo and Thierry Gallouët. Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.*, 87(1):149–169, 1989.

[2] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie groups and potential theory for their sub-Laplacians*. Springer Monographs in Mathematics. Springer, Berlin, 2007.

[3] Luca Capogna, Donatella Danielli, and Nicola Garofalo. An embedding theorem and the Harnack inequality for nonlinear subelliptic equations. *Comm. Partial Differential Equations*, 18(9-10):1765–1794, 1993.
[4] Luca Capogna, Donatella Danielli, Scott D. Pauls, and Jeremy T. Tyson. *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, volume 259 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2007.

[5] E. DiBenedetto. $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal.*, 7(8):827–850, 1983.

[6] Frank Duzaar and Giuseppe Mingione. Gradient estimates via non-linear potentials. *Amer. J. Math.*, 133(4):1093–1149, 2011.

[7] G. B. Folland and Elias M. Stein. *Hardy spaces on homogeneous groups*, volume 28 of *Mathematical Notes*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

[8] Enrico Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[9] David Jerison. The Poincaré inequality for vector fields satisfying Hörmander’s condition. *Duke Math. J.*, 53(2):503–523, 1986.

[10] Tero Kilpeläinen and Jan Malý. The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.*, 172(1):137–161, 1994.

[11] David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 31 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1980 original.

[12] Tuomo Kuusi and Giuseppe Mingione. Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.*, 207(1):215–246, 2013.

[13] Gary M. Lieberman. Boundary regularity for solutions of degenerate elliptic equations. *Nonlinear Anal.*, 12(11):1203–1219, 1988.

[14] Gary M. Lieberman. The natural generalization of the natural conditions of Ladyzhenskaya and Ural’tseva for elliptic equations. *Comm. Partial Differential Equations*, 16(2-3):311–361, 1991.

[15] Giuseppe Mingione, Anna Zatorska-Goldstein, and Xiao Zhong. Gradient regularity for elliptic equations in the Heisenberg group. *Adv. Math.*, 222(1):62–129, 2009.

[16] S. Mukherjee and X. Zhong. $C^{1,\alpha}$-Regularity for variational problems in the Heisenberg group. [https://arxiv.org/abs/1711.04871](https://arxiv.org/abs/1711.04871)

[17] Shirsho Mukherjee. $C^{1,\alpha}$-Regularity of Quasilinear elliptic equations on the Heisenberg Group. [https://arxiv.org/abs/1805.03742](https://arxiv.org/abs/1805.03742) 2018.

[18] Shirsho Mukherjee. On local Lipschitz regularity of Quasilinear Elliptic equations in the Heisenberg Group. [https://arxiv.org/abs/1804.00751](https://arxiv.org/abs/1804.00751) 2018.

[19] Peter Tolksdorf. Regularity for a more general class of quasilinear elliptic equations. *J. Differential Equations*, 51(1):126–150, 1984.

[20] Neil S. Trudinger and Xu-Jia Wang. On the weak continuity of elliptic operators and applications to potential theory. *Amer. J. Math.*, 124(2):369–410, 2002.

[21] Anna Zatorska-Goldstein. Very weak solutions of nonlinear subelliptic equations. *Ann. Acad. Sci. Fenn. Math.*, 30(2):407–436, 2005.

[22] Xiao Zhong. Regularity for variational problems in the Heisenberg Group. [https://arxiv.org/abs/1711.05524](https://arxiv.org/abs/1711.05524)

(S. Mukherjee) Department of Mathematics, Johns Hopkins University, 3400 N. Charles Street, Baltimore MD 21218, USA

E-mail address: smukhe20@jhu.edu

(Y. Sire) Department of Mathematics, Johns Hopkins University, 3400 N. Charles Street, Baltimore MD 21218, USA

E-mail address: sire@math.jhu.edu