Mathematische Annalen manuscript No.
(will be inserted by the editor)

Regularity Results for Quasilinear Elliptic Equations in the Heisenberg Group

Juan J. Manfredi · Giuseppe Mingione

Received: date / Revised: date – © Springer-Verlag 2007

Abstract. We prove regularity results for solutions to a class of quasilinear elliptic equations in divergence form in the Heisenberg group $\mathbb{H}^n$. The model case is the non-degenerate p-Laplacean operator

$$\sum_{i=1}^{2n} X_i \left(\left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} X_i u \right) = 0,$$

where $\mu > 0$, and p is not too far from 2.

Mathematics Subject Classification (2000): 35D10, 35H05

Key words. Heisenberg group, p-Laplacean, weak solutions, regularity

---

Contents

Juan J. Manfredi
Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260,USA
e-mail: `manfredi@pitt.edu`
URL: http://www.pitt.edu/~manfredi

Giuseppe Mingione
Dipartimento di Matematica, Università di Parma, Parco Area delle Scienze 53/A, I-43100 Parma, Italy
e-mail: `giuseppe.mingione@unipr.it`
URL: www.unipr.it/~mingiu36

1. Introduction and Results

The aim of this paper is to prove the local smoothness of solutions for a class of non-linear elliptic equations in the Heisenberg group $\mathbb{H}^n$. The novelty here is the consideration of equations with super-linear growth with respect to the (horizontal) gradient. Indeed, a primary example covered by our analysis is the non-degenerate p-Laplacean equation

$$\sum_{i=1}^{2n} X_i \left(\left(\mu^2 + |\mathfrak{X}u|^2 \right)^{\frac{p-2}{2}} X_i u \right) = 0, \tag{1}$$

for a suitable, explicitly determined range of values of the growth exponent $p \geq 2$, and $\mu > 0$. Here, denoting points $x \in \mathbb{H}^n$ with the usual coordinates $x = (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_{2n}, t)$, we mean

$$X_i \equiv X_i(x) = \partial_{x_i} - \frac{x_{n+i}}{2}\partial_t, \qquad X_{n+i} \equiv X_{n+i}(x) = \partial_{x_{n+i}} + \frac{x_i}{2}\partial_t,$$

$$T \equiv T(x) = \partial_t, \qquad \mathfrak{X}u = (X_1 u, X_2 u, \ldots, X_{2n} u).$$

In section 2 we present more details about the Heisenberg group.

More generally, we shall consider elliptic equations in divergence form of the type

$$\sum_{i=1}^{2n} X_i a_i(\mathfrak{X}u) = 0, \tag{2}$$

where the vector field $a = (a_i)\colon \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is of class C^1 and satisfies the following growth and ellipticity conditions:

$$|Da(z)|(\mu^2 + |z|^2)^{\frac{1}{2}} + |a(z)| \leq L(\mu^2 + |z|^2)^{\frac{p-1}{2}}, \tag{3}$$

and

$$\nu(\mu^2 + |z|^2)^{\frac{p-2}{2}} |\lambda|^2 \leq \sum_{i,j=1}^{2n} D_{z_j} a_i(z) \lambda_i \lambda_j, \tag{4}$$

for every $z, \lambda \in \mathbb{R}^{2n}$, where $0 < 2\nu \leq L$, μ and the growth exponent p, are restricted by the non-degeneracy condition

$$0 < \mu \leq 1, \tag{5}$$

and the following bound on the growth rate:

$$2 \leq p < c(n), \tag{6}$$

where

$$c(n) = \begin{cases} 4 & \text{if } n = 1,2 \\ \frac{10}{3} & \text{if } \ n = 3 \\ 1 + n - \sqrt{n^2 - 2n - 3} & \text{if } \ n \geq 4. \end{cases} \tag{7}$$

When $n = 3$ the end point $p = 10/3$ is actually also included; see Remark 7 below. Observe that $c(n) > 2$ and $\lim_{n\to\infty} c(n) = 2$. The crucial point in (5) is that $\mu \neq 0$, while the limitation $\mu \leq 1$ is assumed here only as an avoidable normalization condition. The choice

$$a_i(z) = (\mu^2 + |z|^2)^{\frac{p-2}{2}} z_i \quad \text{for} \quad i \in \{1, \ldots, 2n\}$$

makes the equation (1) fall into the class considered in (2). Our results also apply to minima of variational integrals of the type

$$u \mapsto \int_\Omega f(\mathfrak{X}u)\, dx, \tag{8}$$

provided $f \colon \mathbb{R}^{2n} \to \mathbb{R}$ is a differentiable function such that the vector field $(a_i(z)) = (D_{z_i} f(z))$ satisfies the assumptions specified in (3)–(6). In this case, the regularity of minima is obtained by passing to the Euler-Lagrange equation of the functional in (8) above

$$\sum_{i=1}^{2n} X_i \left(D_{z_i} f(\mathfrak{X}u)\right) = 0, \tag{9}$$

which is of type (2). A typical example in this case is, of course,

$$u \mapsto \int_\Omega (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}}\, dx, \tag{10}$$

with μ and p satisfying (5) and (6) respectively.

Under the previous assumptions, we shall first prove Lipschitz continuity of weak solutions and related explicit, a priori estimates: this is actually the focal point of the theory and the real hard part of our work. After such a step we apply a technique of Capogna [2],[4] to conclude that weak solutions are smooth. The $C^{1,\alpha}$-regularity of solutions for the case $p = 2$ has been obtained by Capogna in [2],[3],[4]. We are not aware of any result concerning the case $p \neq 2$, except from results obtained via Cordes perturbation techniques, when p lies in a small, not explicitly determined, neighborhood of 2; see [9] and [10]. On the other hand we note that in these two papers (5) is not needed and the degenerate p-Laplacean, given by (1) when $\mu = 0$, is also included. As far as the lower order regularity theory is concerned, that is $C^{0,\alpha}$-regularity, we mention the papers

[5],[23],[24],[32] for sub-elliptic equations, and the paper [31], where variational integrals of the type in (8) are considered. See also Theorem 4 below. Regularity results for minima of general vectorial and strictly convex functionals of the type (8), including the one in (9), have been proved by Capogna & Garofalo [6], relying on explicit a priori estimates for solutions to constant coefficients systems. A new and interesting approach to such estimates, has been recently given by Föglein in [12], where the case of elliptic systems with general p-growth is considered. Both in [6] and [13] partial regularity results are proved i.e. regularity of the gradient of solutions outside a closed negligible set, in fact called the singular set.

Our main result is the following:

Theorem 1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(6). *Then the Euclidean gradient Du is locally Hölder continuous in Ω.*

Once the Hölder continuity of the gradient is achieved, higher smoothness of the vector field (a_i) implies higher regularity of solutions. As an example of such application of Theorem 1 we have

Theorem 2. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (1) *under the assumptions* (5)-(6). *Then $u \in C^\infty(\Omega)$.*

A couple of words on the bound (6). We do not know if the number $c(n)$ determined in (7) is already optimal or not, but we feel that some improvements can still be done. We think that a first important step toward establishing the maximal regularity of solutions for more general values of p would be to show that $c(n)$ is actually independent of n. In this respect, the value $c(n) = 4$ appears to be a natural threshold number to investigate, see also [8]. Eventually proving results for all possible values of p, that is proving or disproving regularity, still remains a very open issue. We are planning to take up such issues in future work.

The proof of Theorem 1 is divided in various steps. Its focal point is, as mentioned above, the proof of the boundedness of the full Euclidean gradient Du of the solution u. In the standard, Euclidean case, this is achieved using the fact that certain non-linear functions of the gradient turn out to be a sub-solutions of a suitable, linear elliptic equation with bounded and measurable coefficients, obtained by various manipulations based on differentiating the starting equation. This is not the case in the present setting. Indeed the horizontal differentiation of (2) leads to the appearance of extra terms, containing the vertical part Tu, and needing a new, suitable treatment. It is worthwhile remarking that proving the a priori boundedness of Du is not needed when $p = 2$ [2],[3], since the differentiated equation exhibits coefficients that are automatically bounded;

see Remark 11 below. In order to prove that Du is bounded we use a delicate boot-strap argument getting in turn regularity for the horizontal derivatives from regularity of the vertical one and vice-versa. Note that in all the previous approaches these two stages were separate [21],[2],[3],[4]. We begin with the results of Marchi [26] and Domokos [8], asserting that $Tu \in L^p$, provided $2 \le p < 4$. We use this regularity of the vertical derivative to show suitable higher integrability for $\mathfrak{X}u$ in certain L^s spaces, for $s > p$. Using this last fact, we shall go back to Tu to prove its boundedness via a suitable, anisotropic, application of the Moser's iteration technique. It is at this point that we need assumptions (5) and (6), essentially to by-pass the fact that the equation does not control the Tu part. As a side benefit we shall also obtain an explicit L^∞-a priori estimate for Tu, which reveals its precise dependence with respect to the parameter μ, and also reflects the interplay between horizontal and vertical regularity through the appearance of the L^s norm of $\mathfrak{X}u$, see (41) below. With the Tu boundedness in our hands, we can now estimate the horizontal gradient $\mathfrak{X}u$ again. This needs, in turn, the horizontal differentiation of the equation and the treatment of the right-side extra terms generated by the non-commutativity of the vector fields X_i. At this stage we shall need a careful integration-by-parts technique to put all the new terms in the correct, ready-to-estimate form, taking now strong advantage that Tu is bounded. We will also need a peculiar choice of suitable test functions in order to overcome certain technical problems. At the end, once again we run Moser's iteration to obtain the a priori L^∞- estimate for $\mathfrak{X}u$ in (52), which extends the Euclidean ones, and depends on the L^∞-norm of Tu. Eventually, this will also lead us to establish a more precise form of the L^∞-a-priori estimate for Tu, revealing an interesting duality with the estimates for $\mathfrak{X}u$; see Remark 10 below. To highlight the interaction between the vertical and horizontal regularization procedures described up to now, we summarize the various steps of our proof of the boundedness of Du in the following scheme:

$$\begin{gathered} Tu \in L^p_{\mathrm{loc}} \implies |\mathfrak{X}u|^{\frac{p}{2}} \in HW^{1,2}_{\mathrm{loc}} \implies \\ T\left(|\mathfrak{X}u|^{\frac{p}{2}}\right) \in L^2_{\mathrm{loc}} \implies \mathfrak{X}u \in L^{\frac{Np}{N-2}}_{\mathrm{loc}} \implies \\ \mathfrak{X}u \in L^{\frac{Qp}{Q-p}-\varepsilon}_{\mathrm{loc}} \implies Tu \in L^\infty_{\mathrm{loc}} \implies \mathfrak{X}u \in L^\infty_{\mathrm{loc}}. \end{gathered} \tag{11}$$

In the last line $\varepsilon > 0$ can be picked arbitrarily small. The starting information, together with the first arrow, can be retrieved from the papers [8],[26] of Domokos and Marchi, respectively, while all the other implications are actually the content of this paper. After obtaining the Lipschitz regularity of solutions a further differentiation of the equation and a modification of Capogna's technique [2],[3],[4] leads to $C^{1,\alpha}$ regularity, and eventually to C^∞-smoothness in the case of a smooth vector field (a_i). Some kind of explicit a priori estimates will be finally demonstrated in Section 8.

2. Preliminaries

We adopt the convention of denoting by c a general constant, possibly varying from line to line in the same chain of inequalities, while the relevant dependence will be specified, and possibly denoted in a more peculiar way; e.g.: c_0, c_1, and so on. In the following Ω will denote an open, bounded subset of $\mathbb{R}^{2n+1}$, where $n \in \mathbb{N}$. We also adopt the following convention: when dealing with a measurable function f, by saying that f is bounded, we usually mean that f is essentially bounded; in other words we identify sup with esssup. Moreover, we dealing with functions spaces of vector valued maps $f : \Omega \to \mathbb{R}^n$, we shall make the usual identification $L^s(\Omega, \mathbb{R}^n) = L^s(\Omega)$, $W^{1,s}(\Omega, \mathbb{R}^n) = W^{1,s}(\Omega)$ and so on.

2.1. The Heisenberg group; CC-distance, CC-balls.

We identify the Heisenberg group $\mathbb{H}^n$ with $\mathbb{R}^{2n+1}$, see also (14) below. Points in $\mathbb{H}^n$ are denoted by

$$x = (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_{2n}, t). \tag{12}$$

The group multiplication is given by

$$\begin{aligned}(x_1, ..., x_{2n}, t) &\cdot (y_1, ..., y_{2n}, u)\\ &= (x_1 + y_1, ..., x_{2n} + y_{2n}, t + u + \frac{1}{2}\sum_{i=1}^{n}(x_i y_{n+i} - x_{n+i} y_i)).\end{aligned}$$

For $1 \leq i \leq n$ we have the canonical left invariant vector fields

$$X_i = \partial_{x_i} - \frac{x_{n+i}}{2}\partial_t, \qquad \text{and} \qquad X_{n+i} = \partial_{x_{n+i}} + \frac{x_i}{2}\partial_t,$$

the only non-trivial commutator being

$$T = \partial_t = [X_i, X_{n+i}].$$

We call $X_1, X_2, \ldots, X_{2n}$ horizontal vector fields and T the vertical vector field. The horizontal gradient of a function $u \colon \mathbb{H}^n \mapsto \mathbb{R}$ is the vector

$$\mathfrak{X}u = (X_1 u, X_2 u, \ldots, X_{2n} u).$$

The second horizontal derivatives are given by the $2n \times 2n$ matrix $\mathfrak{X}\mathfrak{X}u = \mathfrak{X}^2 u$ with entries

$$(\mathfrak{X}(\mathfrak{X}u))_{i,j} = \left(\mathfrak{X}^2(u)\right)_{i,j} = X_i(X_j(u)).$$

Note that such a matrix is not symmetric due to the non-commutativity of the horizontal vector fields X_i. The standard Euclidean gradient of a function u will be denoted by

$$Du = (D_1 u, D_2 u, \ldots, D_{2n} u, D_{2n+1} u).$$

The Euclidean dimension and the Homogeneous dimension of $\mathbb{H}^n$ will be denoted by

$$N = 2n + 1, \qquad \text{and} \qquad Q = 2n + 2, \tag{13}$$

respectively. For notational convenience we shall also denote

$$Y_s = X_{s+n}, \qquad \text{and} \qquad y_s = x_{s+n}, \qquad \text{for } s \in \{1, \ldots, n\}.$$

The Heisenberg Lie algebra $\mathfrak{h}^n$ is as step 2 nilpotent Lie algebra. This means that $\mathfrak{h}_n$ admits a decomposition as a vector space sum

$$\mathfrak{h}^n = \mathfrak{h}_0 \oplus \mathfrak{h}_1$$

such that

$$[\mathfrak{h}_0, \mathfrak{h}_0] = \mathfrak{h}_1.$$

The horizontal part $\mathfrak{h}_0$ is generated by $\{X_1, \ldots, X_n, Y_1, \ldots, Y_n\}$ and the vertical part $\mathfrak{h}_1$ by T. Note that $\mathfrak{h}^n$ is generated as a Lie algebra by $\mathfrak{h}_0$.

The exponential mapping $\exp\colon \mathfrak{h}^n \mapsto \mathbb{H}^n$ is a global diffeomorphism. A point $x \in \mathbb{H}^n$ has exponential coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n, t)$ if

$$x = \exp\left(\left(\sum_{j=1}^{n} x_i X_i + y_i Y_i\right) + tT\right). \tag{14}$$

The identification between $\mathbb{H}^n$, $\mathfrak{h}^n$, and $\mathbb{R}^{2n+1}$ is precisely the use of exponential coordinates in $\mathbb{H}^n$, and it is already used in (12).

The horizontal tangent space at a point $x \in \mathbb{H}^n$ is the $2n$-dimensional subspace

$$T_{\mathrm{h}}(x) = \text{linear span}\{X_1(x), \ldots, X_n(x), Y_1(x), \ldots, Y_n(x)\}.$$

A piecewise smooth curve $t \mapsto \gamma(t)$ is horizontal if $\gamma'(t) \in T_h(\gamma(t))$ whenever $\gamma'(t)$ exists. Given two points $x, y \in \mathbb{H}^n$ denote by

$$\Gamma(x, y) = \{\text{horizontal curves joining } x \text{ and } y\}.$$

Chow's accessibility theorem [7] implies that $\Gamma(x, y) \neq \emptyset$.

For convenience, we fix an ambient Riemannian metric in $\mathbb{H}^n$ so that $\mathfrak{h}_0 = \{X_1, \ldots, X_n, Y_1, \ldots, Y_n\}$ is a left invariant orthonormal frame and the Riemannian volume element and group Haar measure agree, and are equal to the Lebesgue measure in $\mathbb{R}^{2n+1}$.

The Carnot-Carathèodory metric is then defined by

$$d_{cc}(x, y) = \inf\{\text{length}(\gamma) : \gamma \in \varGamma(x, y)\}.$$

It depends only on the restriction of the ambient Riemannian metric to the horizontal distribution generated by the horizontal tangent space.

The Carnot gauge is $|x|_{cc} = d_{cc}(x, 0)$. While explicit formulas are available [1], sometimes it is more convenient to work with an equivalent gauge, smooth away from the origin, called the Heisenberg gauge:

$$|x|_{\mathbb{H}^n} = \left(\left(\sum_{j=1}^{n} x_i^2 + y_i^2 \right)^2 + t^2 \right)^{\frac{1}{4}}.$$

These gauges are indeed comparable [1]

$$|x|_{cc} \approx |x|_{\mathbb{H}^n} \approx \left(\sum_{j=1}^{n} |x_i| + |y_i| \right) + |t|^{\frac{1}{2}}.$$

The non-isotropic dilations are the group homorphisms given by

$$\delta_r \left(x_1, \ldots, x_n, y_1, \ldots y_n, t\right) = \left(rx_1, \ldots, rx_n, ry_1, \ldots ry_n, r^2 t\right),$$

where $r > 0$. In this paper all the balls will be considered with respect to the Carnot-Carathèodory distance (CC-distance):

$$B(x_0, r) = \{y \in \mathbb{H}^n \colon d_{cc}(x_0, y) < r\} = \{y \in \mathbb{H}^n \colon |y^{-1} \cdot x_0|_{cc} < r\},$$

but we could equally have used the smooth gauge to define equivalent balls

$$B'(x_0, r) = \{y \in \mathbb{H}^n \colon |y^{-1} \cdot x_0|_{\mathbb{H}^n} < r\}.$$

The point is that in both cases we get the ball centered at the origin of radius $r > 0$ by applying the nonisotropic dilation δ_r to the unit ball centered at the origin

$$B(0, r) = \delta_r B(0, 1).$$

In particular, the volume can be estimated by

$$|B(x_0, r)| \approx r^Q. \tag{15}$$

When clear, or not essential in the context, we will omit the center of the ball $B_r = B(x_0, r)$ and, if not otherwise stated, when considering several balls simultaneously, they will be concentric. The doubling property of the balls B_r now easily follows from (15). More precisely for every compact subset $K \subset \mathbb{R}^{2n+1}$ there exist a constant $C < \infty$ and a radius $R_0 > 0$, both depending on K, such that

$$|B(x,\lambda r)| \leq C|B(x,r)|\lambda^Q \qquad \text{whenever} \quad \lambda r \leq R_0 \text{ and } \ x \in K. \tag{16}$$

In the rest of the paper we shall use this result with the choice $K = \overline{\Omega}$, with $C = C(\Omega)$ and $R_0 = R_0(\Omega)$ denoting the respective quantities.

In the following, the average of a function $u \in L^1(B(x_0,r))$, over a ball $B(x_0,r)$, is denoted by

$$(u)_r = (u)_{x_0,r} = \frac{1}{|B(x_0,r)|}\int_{B(x_0,r)} u\,dx = \fint_{B(x_0,r)} u\,dx,$$

and once again, when clear from the context, we shall not specify the center.

2.2. Sub-elliptic function spaces.

Let us now recall a few definitions of certain function spaces, that can be retrieved, for instance, in [2],[3],[4]; in the following A will denote a smooth open subset of $\Omega \subset \mathbb{H}^n$. The Folland-Stein class $\Gamma^\alpha(A)$ consists of all Hölder continuous functions, with exponent $\alpha \in (0,1]$, with respect to the CC-distance. Therefore we say that $f \in \Gamma^\alpha(A)$ if and only if

$$[f]_\alpha \equiv [f]_{\alpha,A} := \sup_{x,y\in A, x\neq y} \frac{|f(x)-f(y)|}{d_{cc}(x,y)^\alpha} < \infty, \qquad \alpha \in (0,1]. \tag{17}$$

A function $f \in L^p(A)$ lies in the (Heisenberg) Morrey space $M^{p,\lambda}(A)$ if and only if

$$\sup_{x\in A, 0<r<b} \fint_{B(x,r)\cap A} |f|^p\,dx \leq Cr^{p(\lambda-1)} \text{ where } b = \min\{R_0, \operatorname{diam}(A)\}, \tag{18}$$

for a fixed constant $C < \infty$. In the previous definition the radius $R_0 > 0$ is the one from (16), with the choice $K \equiv \bar{\Omega}$. The local variants $\Gamma^\alpha_{\text{loc}}(A)$ and $M^{p,\lambda}_{\text{loc}}(A)$ are defined saying that $u \in \Gamma^\alpha_{\text{loc}}(A)$ (resp. $u \in M^{p,\lambda}_{\text{loc}}(A)$) if and only if $u \in \Gamma^\alpha(A')$ (resp. $u \in M^{p,\lambda}(A')$), for every smooth open subset $A' \subset\subset A$. Morrey spaces are useful when proving the Hölder continuity of functions via integral estimates, see for instance the classical Morrey's embedding Theorem in Chapter 2 of [16]. This has been extended to the sub-elliptic setting by Capogna [4]. We shall use the following implication from the results in [4]:

$$|\mathfrak{X}u| \in M^{2,\lambda}(A) \Longrightarrow u \in \Gamma^{\alpha}_{\mathrm{loc}}(A)\ . \tag{19}$$

We refer again to [2],[4] for more information on Folland-Stein classes and Morrey type spaces.

The horizontal Sobolev space $HW^{1,p}(\Omega)$ consists of those functions u in $L^p(\Omega)$ whose horizontal distributional derivatives are in turn in L^p, that is $\mathfrak{X}u \in L^p(\Omega)$. $HW^{1,p}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{HW^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\mathfrak{X}u\|_{L^p(\Omega)},$$

for $p \geq 1$. The closure of $C_0^\infty(\Omega)$ in $HW^{1,p}(\Omega)$ is denoted by $HW^{1,p}_0(\Omega)$, while the local variant $HW^{1,p}_{\mathrm{loc}}(\Omega)$ is defined by saying that $u \in HW^{1,p}_{\mathrm{loc}}(\Omega)$ if and only if $u \in HW^{1,p}(\Omega')$, for every open subset $\Omega' \subset\subset \Omega$.

The Homogeneous dimension Q defined in (13) comes into the play also when proving the sub-elliptic version of the classical Sobolev embedding theorem. The following is a particular case of the more general results available in [5],[23].

Theorem 3. *Let $u \in HW^{1,q}(\Omega)$ with $1 < q < Q$, and let $B_r \subset\subset \Omega$ be a ball of radius r. Then $u \in L^{\frac{qQ}{Q-q}}(B_r)$, and moreover there exists a constant c, depending only on n and p, such that*

$$\left(\fint_{B_r} |u|^{\frac{qQ}{Q-q}}\, dx \right)^{\frac{Q-q}{qQ}} \leq c \left(\fint_{B_r} |\mathfrak{X}u|^q\, dx \right)^{\frac{1}{q}} .$$

The previous theorem will be repeatedly used in the rest of the paper with the choice $q = 2$.

2.3. Difference quotients.

Now we shall recall a few basic properties of the difference quotient operators in the Heisenberg group.

Definition 1. *Let Z be a vector field in $\mathbb{H}^n$. The difference quotient of u at the point x is*

$$D_h^Z u(x) = \frac{u(xe^{hZ}) - u(x)}{h},$$

where $h \neq 0$.

The relationship between difference quotients along vector fields and derivatives along vector fields is given by the following lemma of Hörmander **[18]**, valid for general vector fields, not necessarily left-invariant.

Lemma 1. *Let K be a compact set included in Ω open set in $\mathbb{H}^n$. Let Z be vector field and $u \in L^p_{\text{loc}}(\Omega)$, $p > 1$. If there exist σ and C two positive constants such that*

$$\sup_{0<|h|<\sigma} \int_K \left| D_h^Z u(x) \right|^p \, dx \leq C^p$$

then $Zu \in L^p(K)$ and $\|Zu\|_{L^p(K)} \leq C$. Conversely, if $Zu \in L^p(K)$ then for some $\sigma > 0$

$$\sup_{0<|h|<\sigma} \int_K \left| D_h^Z u(x) \right|^p \, dx \leq \Big(2\|Zu\|_{L^p(K)} \Big)^p.$$

Remark 1. Beside the previous result we shall repeatedly use in the following the fact that once $X_i u \in L^q(\Omega)$, for $q > 1$, then $D_h^{X_i} u \to X_i u$ strongly in $L^q_{\text{loc}}(\Omega)$, as $h \to 0$. The proof of this basic fact goes exactly as in the standard Euclidean case, via a density and approximation argument whcih is still available in the Heisenberg group setting; see, for instance [16], Chapter 8.

2.4. Sub-elliptic equations.

In the following we shall also need to consider more general vector fields of the type $a = (a_i)\colon \Omega \times \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$, which are measurable in the first variable, and continuous in the second one. They shall satisfy a growth condition of the type

$$|a(x,z)| \leq L(1 + |z|^2)^{\frac{p-1}{2}}, \qquad L, p > 1, \tag{20}$$

and a monotonicity condition of the type

$$\sum_{i=1}^{2n} a_i(x,z) z_i \geq \nu |z|^p - L, \qquad \nu > 0, \tag{21}$$

for every $x \in \Omega$ and $z \in \mathbb{R}^{2n}$. We observe that assumptions (3)-(4) imply assumptions (20) and (21), eventually choosing a different value for the constants ν and L.

Definition 2. *A weak solution to the equation*

$$\sum_{i=1}^{2n} X_i a_i(x, \mathfrak{X} u) = 0, \tag{22}$$

in the bounded open subset $\Omega \subset \mathbb{H}^n$, under the assumption (20)*, and with* $p > 1$*, is a function* $u \in HW^{1,p}(\Omega)$ *such that*

$$\int_\Omega \sum_{i=1}^{2n} a_i(x, \mathfrak{X}u) X_i \varphi \, dx = 0, \qquad \textit{for all } \ \varphi \in HW_0^{1,p}(\Omega).$$

We shall need a few known regularity result results for solutions of equations as in (2). The first can be found in [5],[23].

Theorem 4. *Let* $p > 1$*, and let* $u \in HW^{1,p}(\Omega)$ *be a weak solution of the equation* (22)*, in a domain* $\Omega \subset \mathbb{H}^n$*, where the vector field* $a = (a_i)\colon \Omega \times \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ *satisfies the assumptions* (20)–(21)*. Then for every* $A \subset\subset \Omega$*, there exists constants* $c \equiv c(n, p, L/\nu, \operatorname{dist}(A, \partial\Omega)) < \infty$*,* $\alpha \equiv \alpha(n, p, L/\nu, \operatorname{dist}(A, \partial\Omega)) > 0$*, such that*

$$||u||_{L^\infty(A)} + [u]_{\alpha,A} \le c.$$

The previous theorem will be crucially used in the last section of the paper, applied for the linear case i.e. $a_i(x, z) := \sum_{i=1}^{2n} b_{i,j}(x) z_j$, see (139) below. We explicitly observe that Theorem 4 applies to solutions of the equation (2), under the assumptions (3) and (4), as already explained above. The second regularity result we shall need in the paper is the following theorem, due to Domokos [**8**].

Theorem 5. *Let* $2 \le p < 4$ *and let* $u \in HW^{1,p}(\Omega)$ *be a weak solution of equation* (2) *in a domain* $\Omega \subset \mathcal{H}^n$*. Then* $Tu \in L^p_{\text{loc}}(\Omega)$*,* $\mathfrak{X}^2 u \in L^2_{\text{loc}}(\Omega)$*. Moreover, if* $B_r = B(x_0, r) \subset\subset \Omega$*, we have the estimates*

$$\int_{B_{\alpha r}} |Tu|^p \, dx \le \frac{c}{r^p} \int_{B_r} \left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p}{2}} \, dx + \frac{c}{r^{2p}} \int_{B_r} |u|^p \, dx, \tag{23}$$

and

$$\begin{aligned} &\int_{B_{\alpha r}} \left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} |\mathfrak{X}^2 u|^2 \, dx \\ &\qquad \le \frac{c}{r^2} \int_{B_r} \left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p}{2}} \, dx + \frac{c}{r^{p+2}} \int_{B_r} |u|^p \, dx. \end{aligned} \tag{24}$$

Here $\alpha \in (0, 1)$ *is a fixed constant depending only on* p*, and* $c > 0$ *is a fixed constant depending only on* p *and* n*.*

Domokos' original statement is actually referred to solutions of (1). The proof is written in such a way that it readily applies to solutions of (2). The precise and explicit dependence of the constants upon r stated in Theorem 5 can be obtained by Domokos' proof via a scaling argument by noticing the the constants in Domokos' theorem are independent of the non-degeneracy parameter μ.

Remark 2. From Theorem 5 and estimate (24), or just from the estimates in [8] and Lemma 3 below, we easily infer that

$$(\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{4}} \mathfrak{X}u \in HW^{1,2}_{\mathrm{loc}}(\Omega). \tag{25}$$

In particular we have

$$\mathfrak{X}u \in HW^{1,2}_{\mathrm{loc}}(\Omega). \tag{26}$$

2.5. Algebraic results.

We finally conclude with three algebraic lemmata. The proof of the first twos can be found, for instance, in [17].

Lemma 2. *Let $1 < p < \infty$. There exists a constant $c = c(n,p) > 0$, independent of $\mu \in [0,1]$, such that, for any $z_1, z_2 \in \mathbb{R}^n$*

$$\begin{aligned} c^{-1}\Big(\mu^2 + |z_1|^2 + |z_2|^2\Big)^{\frac{p-2}{2}} &\le \int_0^1 (\mu^2 + |z_2 + \tau z_1|^2)^{\frac{p-2}{2}}\, d\tau \\ &\le c\Big(\mu^2 + |z_1|^2 + |z_2|^2\Big)^{\frac{p-2}{2}}. \end{aligned}$$

Lemma 3. *Let $1 < p < \infty$. There exists a constant $c \equiv c(n,p) > 0$, independent of $\mu \in [0,1]$, such that, for any $z_1, z_2 \in \mathbb{R}^n$*

$$\begin{aligned} &c^{-1}\Big(\mu^2 + |z_1|^2 + |z_2|^2\Big)^{\frac{p-2}{2}} |z_2 - z_1|^2 \\ &\qquad \le \Big|(\mu^2 + |z_2|^2)^{\frac{p-2}{4}} z_2 - (\mu^2 + |z_1|^2)^{\frac{p-2}{4}} z_1\Big|^2 \\ &\qquad\qquad \le c\Big(1 + |z_1|^2 + |z_2|^2\Big)^{\frac{p-2}{2}} |z_2 - z_1|^2. \end{aligned}$$

The proof of the following iteration lemma can be found in [16], Chapter 6.

Lemma 4. *Let $r > 0$ and $I\colon [r/2, r] \to \mathbb{R}$ be a non-negative, bounded function such that, for any $r/2 \le t < s < r$*

$$I(t) \le \theta I(s) + \frac{A}{(s-t)^q} + B,$$

where $A, B \ge 0$, and $\theta \in (0,1)$, $q \ge 1$ are fixed constants. Then there exists a constant $c \equiv c(\theta)$ such that

$$I(r/2) \le \frac{cA}{r^q} + cB.$$

3. Basic higher integrability

From now on, and for the rest of the paper, we shall denote by u a weak solution to the equation (2), that is a function $u \in HW^{1,p}(\Omega)$ such that

$$\int_\Omega \sum_{i=1}^{2n} a_i(\mathfrak{X}u) X_i\varphi \, dx = 0, \qquad \text{for all } \ \varphi \in HW_0^{1,p}(\Omega). \tag{27}$$

The main aim of this section is to prove the higher integrability of the horizontal gradient of solutions. More precisely, we have:

Theorem 6. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)–(5)*, with $p \in [2,4)$. Then*

$$\mathfrak{X}u \in L^{\frac{pN}{N-2}}_{\mathrm{loc}}(\Omega). \tag{28}$$

This will follow from the existence of vertical derivatives of certain non-linear quantities involving $\mathfrak{X}u$, according to the scheme outlined in (11). We state the result in a way that makes it an extension of the standard higher differentiability results in the Euclidean case. See for instance [16], Chapter 8.

Theorem 7. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)–(5)*, and assume that $2 \le p < 4$. Then*

$$T((\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{4}}\mathfrak{X}u) \in L^2_{\mathrm{loc}}(\Omega) \qquad \textit{and} \qquad Tu \in HW^{1,2}_{\mathrm{loc}}(\Omega). \tag{29}$$

Moreover, let $x_0 \in \Omega$ and $r > 0$ such that $B_r = B(x_0, r) \subset\subset \Omega$. Then, we have

$$\int_{B_{\frac{r}{2}}} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} |T\mathfrak{X}u|^2 \, dx \le \frac{c}{r^2} \int_{B_r} (\mu^2 + |\mathfrak{X}u|^p + |Tu|^p) \, dx, \tag{30}$$

for every $B_\rho = B(x_0, \rho) \subset B_r$. Here the constant c only depends on $n, L/\nu$ and p, being independent of the particular solution u, the constant μ, and the vector field (a_i).

Proof. The proof will be a bit more involved than necessary, since we will eventually build on it for the proof of Theorem 8 below. From now on, without loss of generality, we shall assume that $\nu = 1$ via a standard scaling argument that consists of in considering the vector field $\nu^{-1}(a_i)$ instead of (a_i). Therefore in the constant dependence the letter ν will disappear, but at the end, when re-scaling back, it will reappear in the form L/ν, as in the statement of Theorems 8 and 7. To begin the proof, in the weak form of the equation (27), we take $D^T_{-h}\varphi$ instead of φ. Here

$|h| > 0$ is small enough, depending on the quantity dist(supp $\varphi, \partial\Omega$) to guarantee that $D^T_{-h}\varphi$ has still compact support in Ω. Using discrete integration by parts for difference quotients as in [2] and [8] we obtain

$$\int_\Omega \sum_{i=1}^{2n} D_h^T(a_i(\mathfrak{X}u))X_i\varphi\, dx = 0. \tag{31}$$

Next, we write

$$\begin{aligned} D_h^T(a_i(\mathfrak{X}u))(x) &= \int_0^1 \sum_{j=1}^{2n} D_{z_j}a_i\left(\mathfrak{X}u(x) + shD_h^T\mathfrak{X}u(x)\right) D_h^T X_j u(x)\, ds \\ &= \sum_{j=1}^{2n} A_{i,j}(x)X_j D_h^T u(x), \end{aligned} \tag{32}$$

where we have set

$$A_{i,j}(x) := \int_0^1 D_{z_j}a_i\left(\mathfrak{X}u(x) + shD_h^T\mathfrak{X}u(x)\right)\, ds.$$

We have used the fact that X_i and T commute for every $i \in \{1, \ldots, 2n\}$, to deduce that X_i and D_h^T commute too. We can use the algebraic lemma 2, together with the ellipticity and growth assumptions (4) and (3), respectively. This yields the following lower bound:

$$\begin{aligned} c_1(n,p)(\mu^2 + |\mathfrak{X}u(x)|^2 &+ |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}}|\mathfrak{X}D_h^T u(x)|^2 \\ &\leq \int_0^1 (\mu^2 + |\mathfrak{X}u(x) + shD_h^T\mathfrak{X}u(x)|^2)^{\frac{p-2}{2}}\, ds\ |\mathfrak{X}D_h^T u(x)|^2 \\ &\leq \sum_{i,j=1}^{2n} A_{i,j}(x)X_j D_h^T u(x) X_i D_h^T u(x), \end{aligned} \tag{33}$$

and the following upper bound:

$$|A_{i,j}(x)| \leq c_2(n,p,L)(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}}. \tag{34}$$

Using the identity (32), equation (31) can be rewritten as

$$\int_\Omega \sum_{i,j=1}^{2n} A_{i,j}X_j D_h^T u X_i\varphi\, dx = 0.$$

In the previous equality we use the test function

$$\varphi = \eta^2 g((D_h^T u)^2)D_h^T u, \qquad \text{where} \qquad g(t) \equiv g_\alpha(t) := t^\alpha \qquad \alpha \geq 0, \tag{35}$$

and $\eta \in C_0^\infty(\Omega)$ is a smooth cut-off function with compact support in Ω, to be chosen later, and such that $0 \le \eta \le 1$. This is an admissible test function since u is locally bounded in Ω by Theorem 4. It would also be possible to use another choice of test function here, in order to avoid the use of such result, and give a self contained proof of the local boundedness of Tu (see the choice of the function g in (65) below, the procedure followed in the next section, and Remark 6 below). For simplicity, we prefer to use the results of [5],[23]. Expanding terms in (35) we have

$$\begin{aligned}\int_\Omega \sum_{i,j=1}^{2n} A_{i,j} X_j D_h^T u \big[& 2\eta X_i \eta g((D_h^T u)^2) D_h^T u \\ & + 2\eta^2 g'((D_h^T u)^2)(D_h^T u)^2 X_i D_h^T u + \eta^2 g((D_h^T u)^2) X_i D_h^T u\big]\, dx = 0.\end{aligned}$$

We used again the fact that X_i and D_h^T commute for every $i \in \{1, \dots, 2n\}$. Using (33) we find

$$\begin{aligned}& c\int_\Omega \eta^2 (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha} |\mathfrak{X} D_h^T u|^2\, dx \\ & \le \int_\Omega \eta^2 \left[2g'((D_h^T u)^2)(D_h^T u)^2 + g((D_h^T u)^2)\right] \sum_{i,j=1}^{2n} A_{i,j} X_j D_h^T u X_i D_h^T u\, dx\end{aligned}$$

while making use of (34) and Young's inequality we have, for any $\varepsilon \in (0,1)$

$$\begin{aligned}& \left| \int_\Omega 2\eta \left(\sum_{i,j=1}^{2n} A_{i,j} X_j D_h^T u X_i \eta \right) g((D_h^T u)^2) D_h^T u\, dx \right| \\ & \le c \int_\Omega \eta |\mathfrak{X}\eta| (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha} |D_h^T u| |\mathfrak{X} D_h^T u|\, dx \\ & \le \varepsilon \int_\Omega \eta^2 (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha} |\mathfrak{X} D_h^T u|^2\, dx \\ & \quad + \frac{c\, C_\eta}{\varepsilon} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha+2}\, dx,\end{aligned}$$

where we have set

$$C_\eta = \|\mathfrak{X}\eta\|^2_{L^\infty(\Omega)}, \tag{36}$$

and used the standard identification

$$(D_h^T u)^{2\gamma} = ((D_h^T u)^2)^\gamma, \qquad \text{for all } \ \gamma \ge 0. \tag{37}$$

Taking ε small enough, depending on n, p and L only and reabsorbing terms we have

$$\int_\Omega \eta^2 (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha} |\mathfrak{X} D_h^T u|^2\, dx$$

$$\leq c\,C_\eta \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha+2}\, dx, \quad (38)$$

where $c = c(n, p, L)$. In the previous inequality we take now $\alpha = 0$, and we choose $\eta \in C_0^\infty(B_r)$ to be such that $\eta \equiv 1$ on $B_{r/2}$, $0 \leq \eta \leq 1$, and $\|\mathfrak{X}\eta\|_{L^\infty(\Omega)} \leq c(n)/r$; the existence of such a function can be inferred as in [5]. Moreover we use again that $\mathfrak{X} D_h^T u = D_h^T \mathfrak{X} u$, getting

$$\int_{B_{\frac{r}{2}}} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} |D_h^T \mathfrak{X}u|^2\, dx$$

$$\leq c\,C_\eta \int_{B_r} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2 + (D_h^T u)^2)^{\frac{p}{2}}\, dx, \quad (39)$$

where we also applied Young's inequality on the right hand side. Using the fact that $Tu \in L^p(B_r)$, that $p \geq 2$, and that $\mu > 0$, we have that the quantity $\int |D_h^T \mathfrak{X}u|^2$ stays bounded uniformly with respect to h and then Hörmander's Lemma 1 on difference quotients tells us that $T\mathfrak{X}u \in L^2(B_r)$, and the second inclusion in (29) follows. Once the existence of $T\mathfrak{X}u$ a.e. is achieved, we can let $h \to 0$ in the previous inequality, and a standard application of Fatou's lemma to treat the left hand side finally yields (30). As for the first assertion in (29), it is sufficient to remark that using Lemma 3 in combination with (39), we have that

$$\int_{B_{\frac{r}{2}}} \left| D_h^T \left((\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{4}} \mathfrak{X}u \right) \right|^2 dx$$

$$\leq c(n,p) \int_{B_{\frac{r}{2}}} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{p-2}{2}} |D_h^T \mathfrak{X}u|^2\, dx,$$

from which also the first assertion in (29) immediately follows.

A straightforward application of the standard, Euclidean Sobolev embedding theorem now gives the desired higher integrability of $\mathfrak{X}u$ asserted by Theorem 6.

Proof (of Theorem 6). It is a standard fact. Combining (29) with the assertion in (25) yields

$$D((\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{4}} \mathfrak{X}u) \in L^2_{\mathrm{loc}}(\Omega),$$

therefore by the standard Sobolev embedding theorem we have that

$$(\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{4}} \mathfrak{X}u \in L^{\frac{2N}{N-2}}_{\mathrm{loc}}(\Omega),$$

from which (28) immediately follows, again via Hormander's Lemma 1.

4. Vertical estimates

In this section we continue to follow the path outlined in (11). We consider the fifth arrow. More precisely, we shall prove that the higher integrability of the horizontal gradient $\mathfrak{X}u$ with a certain exponent $s \geq p$, will force the local boundedness of Tu, provided p is not far from 2, in a way that is determined by the size of s itself, via (40) below. Eventually, using the bound in (6), we shall apply such a fact with the choice $s \approx Qp/(Q-p)$. In Section 7 we shall also let $s \to \infty$ to get a certain a priori estimate (135) for Tu. The main result of this section is therefore the following:

Theorem 8. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)–(5)*, with $2 \leq p < 4$. Assume also that $\mathfrak{X}u \in L^s_{\mathrm{loc}}(\Omega)$, where $s \geq p$ is such that*

$$\chi = \frac{Q}{Q-2}\frac{s-p+2}{s} > 1. \tag{40}$$

Then we have $Tu \in L^\infty_{\mathrm{loc}}(\Omega)$. Moreover, let $x_0 \in \Omega$ and $r > 0$ be such that $B_r = B(x_0, r) \subset\subset \Omega$, then we have

$$\|Tu\|_{L^\infty(B_\rho)} \leq \left(\frac{c}{r-\rho}\right)^{\frac{\chi}{\chi-1}} \left(\frac{\|\mu + |\mathfrak{X}u|\|_{L^s(B_r)}}{\mu}\right)^{\frac{(p-2)\chi}{2(\chi-1)}} \|Tu\|_{L^{\frac{2s}{s-p+2}}(B_r)}, \tag{41}$$

for every $B_\rho = B(x_0, \rho) \subset B_r$. Here the constant c only depends on $n, L/\nu$ and p, being independent of the particular solution u, the constant μ, and the vector field (a_i), and s.

Remark 3. Estimate (41) exhibits a few interesting features. First of all the rate of "blow-up" with respect to the parameters μ, χ and the $L^s(B_r)$ norm of $\mathfrak{X}u$ is exactly quantified. In particular, the estimate worsens when either $\mu \to 0$, that is the equation becomes degenerate, or when $\chi \to 1$, that is the inequality in (6) becomes an equality. Moreover, in the non-degenerate case $p = 2$ we notice the identities

$$\frac{s}{s-p+2} = 1 \qquad\qquad \frac{\chi}{\chi-1} = \frac{Q}{2} \tag{42}$$

and (41) reduces to

$$\|Tu\|_{L^\infty(B_\rho)} \leq \frac{c}{(r-\rho)^{\frac{Q}{2}}}\|Tu\|_{L^2(B_r)},$$

which is the analog of the usual estimates obtained in the Euclidean case when considering uniformly elliptic equations. For instance, taking

$\rho = r/2$ and taking into account that $|B_r| \approx r^Q$, we get the weak-type Harnack inequality

$$\sup_{B_{\frac{r}{2}}} |Tu| \leq c \left(\fint_{B_r} |Tu|^2 \, dx \right)^{\frac{1}{2}}. \tag{43}$$

For another form of the a priori estimate (41) see (135) and (136) below, where we shall use the fact that the constant c in (41) does not depend on s. Anyway, we like to remark that even in the case $p = 2$, estimate (43) is anyway new in this setting.

Proof (of Theorem 8). We restart from the proof of Theorem 7, and more precisely go back to (38), valid for any $\alpha \geq 0$; using the fact that $p \geq 2$, and discarding terms on the left hand side we have

$$\begin{aligned} &\int_\Omega \eta^2 (D_h^T u)^{2\alpha} |\mathfrak{X} D_h^T u|^2 \, dx \\ &\leq \frac{c\, C_\eta}{\mu^{p-2}} \int_{\operatorname{supp} \eta} (\mu^2 + |\mathfrak{X} u(x)|^2 + |\mathfrak{X} u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha+2} \, dx. \end{aligned} \tag{44}$$

We now observe

$$\left| \mathfrak{X} \left(\eta [(D_h^T u)^2]^{\frac{\alpha+1}{2}} \right) \right|^2 \leq c |\mathfrak{X} \eta|^2 (D_h^T u)^{2\alpha+2} + c(\alpha+1)^2 \eta^2 (D_h^T u)^{2\alpha} |\mathfrak{X} D_h^T u|^2. \tag{45}$$

Merging this last inequality with (44), and using the sub-elliptic Sobolev embedding Theorem 3 with $q = 2$, we obtain

$$\begin{aligned} &\left(\int_\Omega \eta^{\frac{2Q}{Q-2}} (D_h^T u)^{\frac{(2\alpha+2)Q}{Q-2}} \, dx \right)^{\frac{Q-2}{Q}} \\ &\leq \frac{c(\alpha+1)^2 C_\eta}{\mu^{p-2}} \int_{\operatorname{supp} \eta} (\mu^2 + |\mathfrak{X} u(x)|^2 + |\mathfrak{X} u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha+2} \, dx. \end{aligned} \tag{46}$$

Note that, referring to (45), we have estimated

$$\begin{aligned} &\int_\Omega |\mathfrak{X} \eta|^2 (D_h^T u)^{2\alpha+2} \, dx \\ &\leq \frac{c C_\eta}{\mu^{p-2}} \int_{\operatorname{supp} \eta} (\mu^2 + |\mathfrak{X} u(x)|^2 + |\mathfrak{X} u(xe^{hT})|^2)^{\frac{p-2}{2}} (D_h^T u)^{2\alpha+2} \, dx, \end{aligned}$$

since $p \geq 2$, and then we have used this last estimate to obtain (46). We recall that the number C_η has been defined in (36), and we are using the convention in (37). In order to estimate the right hand side of (46) we first consider the case $p > 2$, when we apply Hölder's inequality with conjugate exponents $s/(p-2)$ and $s/(s-p+2)$, and obtain

$$\begin{aligned}\left(\int_{\Omega} \eta^{\frac{2Q}{Q-2}} (D_h^T u)^{\frac{(2\alpha+2)Q}{Q-2}}\, dx\right)^{\frac{Q-2}{Q}} \\ \leq \frac{c(\alpha+1)^2 C_\eta \tilde{C}_h}{\mu^{p-2}} \left(\int_{\operatorname{supp}\eta} (D_h^T u)^{\frac{(2\alpha+2)s}{s-p+2}}\, dx\right)^{\frac{s-p+2}{s}}. \qquad (47)\end{aligned}$$

We have set

$$\tilde{C}_h := \left(\int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2)^{\frac{s}{2}}\, dx\right)^{\frac{p-2}{s}};$$

note that

$$\lim_{h\to 0} \tilde{C}_h = \left(\int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u(x)|^2)^{\frac{s}{2}}\, dx\right)^{\frac{p-2}{s}}. \qquad (48)$$

When $p = 2$ inequality (47) immediately follows by (46), taking into account the first identity in (42). The point is that now (47) yields an improvement of the integrability of $D_h^T u$ since (40) implies

$$\frac{(\alpha+1)Q}{Q-2} > \frac{(\alpha+1)s}{s-p+2} \qquad \text{for all} \quad \alpha \geq 0.$$

This is the starting point for running a suitable version of Moser's iteration technique. From now on all the balls considered will be concentric, centered at the given, but arbitrary point $x_0 \in \Omega$. With $\rho < R$ as in the statement, we define the family of interpolating balls $B_\rho \subset B_{\rho_{k+1}} \subset B_{\rho_k} \subset B_R$ accordingly to the following choice of the radii:

$$\rho_k = \rho + \frac{R-\rho}{2^k}, \qquad k \geq 0, \qquad (49)$$

and note that $\rho_0 = R$, while $\rho_k \to \rho$. We next take a family of smooth cut-off functions $\{\eta_k\}_k \subset C_0^\infty(B_{\rho_k})$ in such a way that $0 \leq \eta_k \leq 1$, $\eta_k \equiv 1$ on $B_{\rho_{k+1}}$, and $\|\mathfrak{X}\eta_k\|_{L^\infty(\Omega)} \leq \gamma(n)^k (R-\rho)^{-1}$, for every $k \in \mathbb{N}$, and where $\gamma(n) \geq 2$ is an absolute constant. The existence of such a family of functions can be inferred from [5]. Now we inductively define the sequence $\{\alpha_k\}$ according to

$$\begin{cases} \alpha_{k+1} = \chi\alpha_k + (\chi - 1) \\ \alpha_0 = 0, \end{cases}$$

for the choice of χ done in (40). It follows that

$$\alpha_k = (\chi - 1)\sum_{j=0}^{k-1} \chi^j = \chi^k - 1, \qquad \text{for all} \quad k \geq 1.$$

Finally we set

$$A_k = \left(\int_{B_{\rho_k}} (D_h^T u)^{\frac{2s(\alpha_k+1)}{s-p+2}} \, dx \right)^{\frac{1}{\alpha_k+1}} = \left(\int_{B_{\rho_k}} (D_h^T u)^{\frac{2s\chi^k}{s-p+2}} \, dx \right)^{\frac{1}{\chi^k}} .$$

We now iterate inequality (47) with the choices $\alpha = \alpha_k$ and $\eta = \eta_k$. An elementary calculation gives that

$$A_{k+1} \le \left[\frac{c_1^k \chi^{2k} \tilde{C}_h}{\mu^{p-2}(r-\rho)^2} \right]^{\frac{Q}{Q-2}\frac{1}{\chi^{k+1}}} A_k, \qquad \text{for all } \ k \ge 0.$$

Here $c_1 \ge 2$ depends only on $n, p, L/\nu$, and, in particular, is independent of h and s. Using elementary induction, the previous inequality also gives

$$\begin{aligned} A_k &\le (c_1\chi^2)^{\frac{Q}{Q-2}\sum_{j=1}^{\infty}\frac{j}{\chi^j}} \left[\frac{\tilde{C}_h}{\mu^{p-2}(r-\rho)^2} \right]^{\frac{Q}{Q-2}\sum_{j=1}^{\infty}\frac{1}{\chi^j}} A_0 \\ &\le \left[c_1 \left(\frac{Q}{Q-2} \right)^2 \right]^{\frac{Q}{Q-2}\sum_{j=1}^{\infty}\frac{j}{\chi^j}} \left[\frac{\tilde{C}_h}{\mu^{p-2}(r-\rho)^2} \right]^{\frac{Q}{Q-2}\sum_{j=1}^{\infty}\frac{1}{\chi^j}} A_0, \end{aligned}$$

for all $k \in \mathbb{N}$ since, by the definition of $\chi = \chi(s)$ in (40), we have that $\chi \le Q/(Q-2)$ when $s \ge p \ge 2$. Now we let $k \to \infty$, then $\alpha_k \to \infty$ since $\chi > 1$, and obtain

$$\sup_{B_\rho} (D_h^T u)^{\frac{2s}{s-p+2}} \le \left(\frac{c}{r-\rho} \right)^{\frac{2Q}{Q-2}\frac{1}{\chi-1}} \left(\frac{\tilde{C}_h}{\mu^{p-2}} \right)^{\frac{Q}{Q-2}\frac{1}{\chi-1}} \int_{B_r} (D_h^T u)^{\frac{2s}{s-p+2}} \, dx,$$

and c here depends only on n, p, L, being independent of both h and s. In the previous estimate the constant c only depends on the fixed quantities n, p and L (by the initial scaling we are assuming $\nu = 1$), and is in particular independent of h. In order to pass $h \to 0$ in the above inequality, we recall now the following facts: $s \ge p \ge 2$ implies that

$$\frac{2s}{s-p+2} \le p$$

and therefore, since $Tu \in L^p_{\mathrm{loc}}(\Omega)$ by Theorem 5, then $D_h^T u \to Tu$, strongly in $L_{\mathrm{loc}}^{\frac{2s}{s-p+2}}(\Omega)$. Keeping into account (48), letting $h \to 0$, a standard lower semicontinuity convergence argument yields

$$\sup_{B_\rho} (Tu)^{\frac{2s}{s-p+2}} \le \left(\frac{c}{r-\rho} \right)^{\frac{2Q}{Q-2}\frac{1}{\chi-1}} \left(\frac{\|\mu + |\mathfrak{X}u|\|_{L^s(B_r)}}{\mu} \right)^{\frac{Q}{Q-2}\frac{p-2}{\chi-1}}$$

$$\cdot \int_{B_r} (Tu)^{\frac{2s}{s-p+2}}\, dx.$$

Estimate (41) immediately follows from this last inequality after an elementary manipulation, with the specified dependence upon the various constants. Finally the fact that $Tu \in L^\infty_{\mathrm{loc}}(\Omega)$ follows via a standard covering argument, since the ball $B(x_0, r) \subset\subset \Omega$ was arbitrary. This concludes the proof of Theorem 8.

Remark 4. In order to clarify what comes next, we remark that previous theorem immediately gives an application. Indeed, using it with $s = p$ and assuming that

$$p < \frac{2Q}{Q-2} = \frac{2n+2}{n}, \tag{50}$$

in order to meet (40), we get the local boundedness of Tu. The bound in (50) is worse than the one in (6), and therefore, as already mentioned at the beginning of the section, we first obtain higher integrability of $\mathfrak{X}u$, in Theorems 6, and 10, and then apply Theorem 8 in a more efficient way, that is with $s > p$. In other words, the higher the integrability exponent of $\mathfrak{X}u$ is, the farer we are allowed to pick p away from 2 via (40).

5. Horizontal estimates

5.1. Main statement of the section

The aim of this section is to derive the local boundedness of horizontal derivatives of solutions from that of Tu; we are actually concerned with the last arrow in the scheme (11). The main result here is therefore

Theorem 9. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(5)*, with the exponent $p \in [2,4)$ satisfying also*

$$p \le \frac{2(N-2)}{N-4}, \qquad \textit{when} \quad n \ge 3. \tag{51}$$

Finally, assume that $Tu \in L^\infty_{\mathrm{loc}}(\Omega)$. Then we have $\mathfrak{X}u \in L^\infty_{\mathrm{loc}}(\Omega)$. Moreover, let $x_0 \in \Omega$ and $r > 0$ be such that $B_r = B(x_0, r) \subset\subset \Omega$, then we also have

$$\|\mathfrak{X}u\|_{L^\infty(B_\rho)} \le \left(\frac{c}{r-\rho}\right)^{\frac{Q}{p}} \|1 + |Tu|\|^{\frac{Q}{2}}_{L^\infty(B_r)} \left(\int_{B_r} (1 + |\mathfrak{X}u|^p)\, dx\right)^{\frac{1}{p}} \tag{52}$$

for every $B_\rho = B(x_0, \rho) \subset B_r$. Here the constant c only depends on $n, L/\nu$ and p, being independent of the particular solution u, the constant μ, and the vector field (a_i).

Remark 5. Taking $\rho = r/2$ in the previous estimate and observing that $|B_r| \approx r^Q$, we obtain

$$\sup_{B_{\frac{r}{2}}} |\mathfrak{X}u| \leq c\|1 + |Tu|\|_{L^\infty(B_r)}^{\frac{Q}{2}} \left(\fint_{B_r} (1 + |\mathfrak{X}u|^p)\, dx \right)^{\frac{1}{p}}, \tag{53}$$

is in some sense dual to the one in (43), and would be the classical $L^\infty - L^p$ reverse Harnack type inequality for solutions to standard elliptic equations, but for the presence of the L^∞-norm of Tu. Also observe that assuming $\|Tu\|_{L^\infty}$ finite, the previous estimates is stable when $\mu \to 0$, while, in the general case, that is without assuming the a-priori boundedness of Tu, the dangerous dependence on μ is hidden in the presence of $\|Tu\|_{L^\infty}$ via estimates of the type in (41). This is the reason why, in the final statement of Theorem 1, we need to assume that $\mu > 0$.

5.2. Some auxiliary functions.

We shall need the following family of auxiliary functions:

$$g_{\alpha,k}(t) = \frac{k(1+t)^\alpha}{k + (1+t)^\alpha} \qquad t, \alpha \geq 0 \qquad k \in \mathbb{N}. \tag{54}$$

We notice that

$$0 \leq g_{\alpha,k}(t) \leq \min\{k, (1+t)^\alpha\}, \tag{55}$$

and moreover

$$0 \leq g_{\alpha,k}(t) \leq g_{\alpha,k+1}(t) \qquad \text{for all } \; k \in \mathbb{N}, \tag{56}$$

and

$$\lim_{k\to\infty} g_{\alpha,k}(t) = (1+t)^\alpha. \tag{57}$$

We have

$$g'_{\alpha,k}(t) = \left[\frac{k}{k + (1+t)^\alpha}\right]^2 \alpha(1+t)^{\alpha-1}$$

and

$$\begin{aligned} g''_{\alpha,k}(t) = & \left[\frac{k}{k + (1+t)^\alpha}\right]^2 \alpha(\alpha-1)(1+t)^{\alpha-2} \\ & -2\left[\frac{k}{k + (1+t)^\alpha}\right]^3 \frac{\alpha^2}{k}(1+t)^{2\alpha-2} \\ = & \, h^1_{\alpha,k}(t) + h^2_{\alpha,k}(t). \end{aligned}$$

We are interested in the following properties of $g_{\alpha,k}(t)$ and its derivatives:

$$g'_{\alpha,k}(t)(1+t) \le \alpha g_{\alpha,k}(t), \tag{58}$$

and

$$|g''_{\alpha,k}(t)|(1+t) \le 3(\alpha+1)g'_{\alpha,k}(t). \tag{59}$$

Estimate (58) is trivial since it immediately reduces to $k \le k+(1+t)^\alpha$. In order to prove (59) we observe

$$|h^1_{\alpha,k}(t)|(1+t) \le (\alpha+1)g'_{\alpha,k}(t),$$

and

$$\begin{aligned}
|h^2_{\alpha,k}(t)|(1+t) &\le 2\alpha^2 \left[\frac{k}{k+(1+t)^\alpha}\right]^2 \frac{(1+t)^{2\alpha-1}}{k+(1+t)^\alpha} \\
&\le 2\alpha \left[\frac{(1+t)^\alpha}{k+(1+t)^\alpha}\right] g'_{\alpha,k}(t) \\
&\le 2(\alpha+1)g'_{\alpha,k}(t),
\end{aligned}$$

so that (59) follows from the last two estimates. We shall also deal with the following family of functions:

$$W_{\alpha,k}(t) := 2g'_{\alpha,k}(t)t + g_{\alpha,k}(t), \qquad t, \alpha \ge 0 \qquad k \in \mathbb{N}. \tag{60}$$

Using (58) and then (55), together with the fact that $g'_{\alpha,k}(t) \ge 0$, we find

$$g_{\alpha,k}(t) \le W_{\alpha,k}(t) \le (2\alpha+1)g_{\alpha,k}(t) \le (2\alpha+1)k. \tag{61}$$

Moreover, taking into account (59), and then again (58), we also find

$$|W'_{\alpha,k}(t)|t \le |W'_{\alpha,k}(t)|(1+t) \le 3(\alpha+1)W_{\alpha,k}(t). \tag{62}$$

Let us again notice that $g'_{\alpha,k}(t) \le g'_{\alpha,k+1}(t)$ for every k, α and t and therefore we also gain, taking into account (56)

$$W_{\alpha,k}(t) \le W_{\alpha,k+1}(t) \qquad\qquad \text{for all } k \in \mathbb{N}. \tag{63}$$

Moreover, by (57)

$$\begin{aligned}
(1+t)^\alpha \le \lim_{k\to\infty} W_{\alpha,k}(t) &= (1+t)^{\alpha-1}[2\alpha t + (1+t)] \\
&\le 3(\alpha+1)(1+t)^\alpha.
\end{aligned} \tag{64}$$

5.3. Identities for certain test functions.

Here we point out a few elementary identities for some test functions we need to use later; these are essentially based on properties of difference quotients in the Heisenberg group and we report them here in a certain detail for future convenience. We adopt a natural notation. Any function ϕ is usually meant to be evaluated at a point $x \in \Omega$ as in $\phi(x)$. If this is not the case we will use a more explicit notation, for example: $\phi(xe^{-hX_s})$, $\phi(xe^{hX_s})$, and so on. With u denoting the fixed solution form Theorem 9, we let

$$\phi_1 \equiv \phi_1^s := D_{-h}^{X_s}(\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}u) \in HW_0^{1,p}(\Omega) \qquad \text{for } s \in \{1,\dots,n\}, \tag{65}$$

where $g\colon \mathbb{R} \mapsto \mathbb{R}$ is a C^∞-function such that

$$|g'(t)t| + |g(t)| \le M \qquad \text{for } t \ge 0 \tag{66}$$

for a certain constant $M \ge 0$, and $\eta \in C_0^\infty(\Omega)$ is a smooth function with compact support in Ω. As usual here, in order to have ϕ_1 to be an admissible test function in (27), and in particular having compact support in Ω, we need to take $|h|$ suitably small, depending on the quantity $\operatorname{dist}(\operatorname{supp}\eta, \partial\Omega) > 0$. The fact that $\phi_1 \in HW_0^{1,p}(\Omega)$ is now basically a consequence of two facts: the first is the bound in (66), the second being the fact that $Tu \in L^p_{\mathrm{loc}}(\Omega)$; differently from the Euclidean setting, for an arbitrary function f we have in fact that $\mathfrak{X}f \in L^p$ does not imply automatically that $\mathfrak{X}D_h^{X_s}f, \mathfrak{X}D_h^{Y_s}f \in L^p$, basically due to the lack of commutativity of (X_s, Y_s). See formulas (68) and (72) below. Now, if $i \neq s+n$ we have

$$\begin{aligned} X_i\phi_1 &= D_{-h}^{X_s}X_i(\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}u) \\ &= D_{-h}^{X_s}\Big[2\eta X_i\eta g((D_h^{X_s}u)^2)D_h^{X_s}u + 2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 X_i D_h^{X_s}u \\ &\qquad\qquad + \eta^2 g((D_h^{X_s}u)^2)X_i D_h^{X_s}u\Big], \end{aligned} \tag{67}$$

and this last function is clearly in $L^p(\Omega)$, by (66). Here we essentially used the fact that $D_{-h}^{X_s}X_i = X_i D_{-h}^{X_s}$ whenever $i \neq s+n$. In order to treat the case $X_i = X_{s+n} = Y_s$ we recall the following formula:

$$Y_s(D_h^{X_s}u)(x) = D_h^{X_s}(Y_s u)(x) - Tu(xe^{hX_s}) \qquad h \neq 0, \tag{68}$$

which is a consequence of the fact that $X_sY_s - Y_sX_s = T$, and have instead

$$\begin{aligned} Y_s\phi_1 &= D_{-h}^{X_s}\Big[Y_s(\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}u)\Big] - T\Big[\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}u\Big](xe^{-hX_s}) \\ &= D_{-h}^{X_s}\Big[2\eta Y_s\eta g((D_h^{X_s}u)^2)D_h^{X_s}u + 2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 Y_s D_h^{X_s}u \end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\eta^2 g((D_h^{X_s}u)^2)Y_s D_h^{X_s}u\Big] \\
&\quad -\Big[(2\eta T\eta g((D_h^{X_s}u)^2)D_h^{X_s}u)(xe^{-hX_s}) \\
&\qquad +(2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 D_h^{X_s}Tu)(xe^{-hX_s}) \\
&\qquad +(\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}Tu)(xe^{-hX_s})\Big] \\
&= D_{-h}^{X_s}\Big[2\eta Y_s\eta g((D_h^{X_s}u)^2)D_h^{X_s}u + 2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 D_h^{X_s}Y_s u \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}Y_s u\Big] \\
&\quad -D_{-h}^{X_s}\Big[2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 Tu(xe^{hX_s}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\eta^2 g((D_h^{X_s}u)^2)Tu(xe^{hX_s})\Big] \\
&\quad -\Big[(2\eta T\eta g((D_h^{X_s}u)^2)D_h^{X_s}u)(xe^{-hX_s}) \\
&\qquad +(2\eta^2 g'((D_h^{X_s}u)^2)(D_h^{X_s}u)^2 D_h^{X_s}Tu)(xe^{-hX_s}) \\
&\qquad +(\eta^2 g((D_h^{X_s}u)^2)D_h^{X_s}Tu)(xe^{-hX_s})\Big],
\end{aligned} \tag{69}
$$

and we can see that all the terms on the right hand side of the previous equality are in $L^p(\Omega)$ since $Tu \in L^p_{\mathrm{loc}}(\Omega)$, η has compact support in Ω, and (66) is in force. In a completely similar way we define

$$
\phi_2 \equiv \phi_2^s := D_{-h}^{Y_s}(\eta^2 g((D_h^{Y_s}u)^2)D_h^{Y_s}u) \in HW_0^{1,p}(\Omega) \quad \text{for } s \in \{1,\dots,n\}. \tag{70}
$$

As above, if $i \neq s$ we have

$$
\begin{aligned}
X_i\phi_2 &= D_{-h}^{Y_s}X_i(\eta^2 g((D_h^{Y_s}u)^2)D_h^{Y_s}u) \\
&= D_{-h}^{Y_s}\Big[2\eta X_i\eta g((D_h^{Y_s}u)^2)D_h^{Y_s}u + 2\eta^2 g'((D_h^{Y_s}u)^2)(D_h^{Y_s}u)^2 X_i D_h^{Y_s}u \\
&\qquad\qquad\qquad\qquad\qquad +\eta^2 g((D_h^{Y_s}u)^2)X_i D_h^{Y_s}u\Big],
\end{aligned} \tag{71}
$$

while in the case $X_i \equiv X_s$ we recall that

$$
X_s(D_h^{Y_s}u)(x) = D_h^{Y_s}(X_s u)(x) + Tu(xe^{hY_s}) \qquad h \neq 0, \tag{72}
$$

and have

$$
\begin{aligned}
X_s\phi_2 &= D_{-h}^{Y_s}\Big[X_s(\eta^2 g((D_h^{Y_s}u)^2)D_h^{Y_s}u)\Big] + T\Big[\eta^2 g((D_h^{Y_s}u)^2)D_h^{Y_s}u\Big](xe^{-hY_s}) \\
&= D_{-h}^{Y_s}\Big[2\eta X_s\eta g((D_h^{Y_s}u)^2)D_h^{Y_s}u + 2\eta^2 g'((D_h^{Y_s}u)^2)(D_h^{Y_s}u)^2 X_s D_h^{Y_s}u \\
&\qquad\qquad\qquad\qquad\qquad\qquad +\eta^2 g((D_h^{Y_s}u)^2)X_s D_h^{Y_s}u\Big]
\end{aligned}
$$

$$+\Big[(2\eta T\eta g((D_h^{Y_s}u)^2)D_h^{Y_s}u)(xe^{-hY_s})$$
$$+(2\eta^2 g'((D_h^{Y_s}u)^2)(D_h^{Y_s}u)^2 D_h^{Y_s}Tu)(xe^{-hY_s})$$
$$+(\eta^2 g((D_h^{Y_s}u)^2)D_h^{Y_s}Tu)(xe^{-hY_s})\Big].$$

Using (72) to develop the right hand side in a way similar to that for (69), we can also check that $X_s\phi_2 \in L^p(\Omega)$. We conclude that both ϕ_1 and ϕ_2 are admissible test functions in (27).

5.4. The proof of Theorem 9

Step 1: Basic elliptic estimates. Select g to be always positive and non-decreasing, and satisfying (66). Take $\varphi = \phi_1$ in the weak form (27), where ϕ_1 is defined in (65). Taking into consideration (67)-(69) we find, using discrete integration by parts for difference quotients,

$$\int_\Omega \sum_{i=1}^{2n} D_h^{X_s}(a_i(\mathfrak{X}u))\Big[2\eta X_i\eta g((D_h^{X_s}u)^2)D_h^{X_s}u \tag{73}$$
$$+\eta^2 W((D_h^{X_s}u)^2)X_iD_h^{X_s}u\Big]\,dx$$
$$= -\int_\Omega a_{s+n}(\mathfrak{X}u)\Big[(2\eta T\eta g((D_h^{X_s}u)^2)D_h^{X_s}u)(xe^{-hX_s})$$
$$+(\eta^2 W((D_h^{X_s}u)^2)D_h^{X_s}Tu)(xe^{-hX_s})\Big]\,dx. \tag{74}$$

In the previous formula we have denoted

$$W(t) = 2g'(t)t + g(t).$$

In a similar way, testing (27) with $\varphi = \phi_2$, with ϕ_2 being defined in (70), we get

$$\int_\Omega \sum_{i=1}^{2n} D_h^{Y_s}(a_i(\mathfrak{X}u))\Big[2\eta X_i\eta g((D_h^{Y_s}u)^2)D_h^{Y_s}u + \eta^2 W((D_h^{Y_s}u)^2)X_iD_h^{Y_s}u\Big]\,dx$$
$$= \int_\Omega a_s(\mathfrak{X}u)\Big[(2\eta T\eta g((D_h^{Y_s}u)^2)D_h^{Y_s}u)(xe^{-hY_s})$$
$$+(\eta^2 W((D_h^{Y_s}u)^2)D_h^{Y_s}Tu)(xe^{-hY_s})\Big]\,dx. \tag{75}$$

We next define

$$A_{i,j}^s(x) := \int_0^1 D_{z_j}a_i\Big(\mathfrak{X}u(x) + \tau h D_h^{X_s}\mathfrak{X}u(x)\Big)\,d\tau, \tag{76}$$

so that

$$
\begin{aligned}
D_h^{X_s}(a_i(\mathfrak{X}u))(x) &= \int_0^1 \sum_{j=1}^{2n} D_{z_j} a_i \left(\mathfrak{X}u(x) + \tau h D_h^{X_s} \mathfrak{X}u(x) \right) D_h^{X_s} X_j u(x)\, d\tau \\
&= \sum_{j=1}^{2n} A_{i,j}^s(x) D_h^{X_s} X_j u(x).
\end{aligned}
$$

The use of formula (68) yields

$$
D_h^{X_s}(a_i(\mathfrak{X}u))(x) = \sum_{j=1}^{2n} A_{i,j}^s(x) X_j D_h^{X_s} u(x) + A_{i,s+n}^s(x) Tu(xe^{hX_s}). \quad (77)
$$

Exactly in the same way, via formula (72), we have

$$
D_h^{Y_s}(a_i(\mathfrak{X}u))(x) = \sum_{j=1}^{2n} A_{i,j}^{s+n}(x) X_j D_h^{Y_s} u(x) - A_{i,s}^{s+n}(x) Tu(xe^{hY_s}), \quad (78)
$$

and this time, according to (76)

$$
A_{i,j}^{s+n}(x) := \int_0^1 D_{z_j} a_i \left(\mathfrak{X}u(x) + \tau h D_h^{Y_s} \mathfrak{X}u(x) \right) d\tau.
$$

We can again use Lemma 2, together with the ellipticity assumption (4), as in (33). This yields the following lower bounds for every $s \in \{1, \ldots, n\}$:

$$
\begin{aligned}
& c(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} |\mathfrak{X} D_h^{X_s} u(x)|^2 \\
& \qquad \le \sum_{i,j=1}^{2n} A_{i,j}^s(x) X_j D_h^{X_s} u(x) X_i D_h^{X_s} u(x), \qquad (79)
\end{aligned}
$$

and

$$
\begin{aligned}
& c(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} |\mathfrak{X} D_h^{Y_s} u(x)|^2 \\
& \qquad \le \sum_{i,j=1}^{2n} A_{i,j}^{s+n}(x) X_j D_h^{Y_s} u(x) X_i D_h^{Y_s} u(x), \qquad (80)
\end{aligned}
$$

with $c = c(n,p) > 0$ being independent of μ. From above we have, via (3) and Lemma 2 again,

$$
|A_{i,j}^s(x)| \le c_2(n,p,L)(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}}, \quad (81)
$$

and

$$
|A_{i,j}^{s+n}(x)| \le c_2(n,p,L)(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}}. \quad (82)
$$

It is worthwhile remarking that in the inequalities (79)-(82) the constants involved do not depend on μ. Now we use identities (77)-(78) in (74)-(75), respectively. Summing up over $s \in \{1, \ldots, n\}$ finally yields

$$
\begin{aligned}
&I_1 + I_2 := \qquad (83)\\
&\int_\Omega \eta^2 \sum_{s=1}^{n} \sum_{i,j=1}^{2n} A^s_{i,j} X_j D_h^{X_s} u X_i D_h^{X_s} u W((D_h^{X_s} u)^2)\, dx \\
&+ \int_\Omega \eta^2 \sum_{s=1}^{n} \sum_{i,j=1}^{2n} A^{s+n}_{i,j} X_j D_h^{Y_s} u X_i D_h^{Y_s} u W((D_h^{Y_s} u)^2)\, dx \\
&= - \int_\Omega 2\eta \sum_{s=1}^{n} \sum_{i,j=1}^{2n} A^s_{i,j} X_j D_h^{X_s} u X_i \eta g((D_h^{X_s} u)^2) D_h^{X_s} u\, dx \\
&\quad - \int_\Omega 2\eta \sum_{s=1}^{n} \sum_{i,j=1}^{2n} A^{s+n}_{i,j} X_j D_h^{Y_s} u X_i \eta g((D_h^{Y_s} u)^2) D_h^{Y_s} u\, dx \\
&\quad - \int_\Omega \sum_{s=1}^{n} \sum_{i=1}^{2n} A^s_{i,s+n} Tu(xe^{hX_s}) \Big[2\eta X_i \eta g((D_h^{X_s} u)^2) D_h^{X_s} u \\
&\qquad\qquad + \eta^2 W((D_h^{X_s} u)^2) X_i D_h^{X_s} u \Big]\, dx \\
&\quad + \int_\Omega \sum_{s=1}^{n} \sum_{i=1}^{2n} A^{s+n}_{i,s} Tu(xe^{hY_s}) \Big[2\eta X_i \eta g((D_h^{Y_s} u)^2) D_h^{Y_s} u \\
&\qquad\qquad + \eta^2 W((D_h^{Y_s} u)^2) X_i D_h^{Y_s} u \Big]\, dx \\
&\quad - \int_\Omega \sum_{s=1}^{n} a_{s+n}(\mathfrak{X} u) \Big[(2\eta T \eta g((D_h^{X_s} u)^2) D_h^{X_s} u)(xe^{-hX_s}) \\
&\qquad\qquad + (\eta^2 W((D_h^{X_s} u)^2) D_h^{X_s} Tu)(xe^{-hX_s}) \Big]\, dx \\
&\quad + \int_\Omega \sum_{s=1}^{n} a_s(\mathfrak{X} u) \Big[(2\eta T \eta g((D_h^{Y_s} u)^2) D_h^{Y_s} u)(xe^{-hY_s}) \\
&\qquad\qquad + (\eta^2 W((D_h^{Y_s} u)^2) D_h^{Y_s} Tu)(xe^{-hY_s}) \Big]\, dx \\
&=: \sum_{k=3}^{8} I_k. \qquad (84)
\end{aligned}
$$

We shall now estimate integrals $I_1, I_2, \ldots, I_8$ separately.

Using (79)-(80) we have that

$$c^{-1} I_1 + I_2 \geq$$

$$\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u(x)|^2\,dx$$
$$+\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} W((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u(x)|^2\,dx.$$

While using the upper bounds in (81)-(82) and Young's inequality twice we have, for any $\varepsilon \in (0,1)$

$$\begin{aligned}
&|I_3| + |I_4| \\
&\le c\int_\Omega \eta|\mathfrak{X}\eta| \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} \cdot \\
&\qquad\qquad \cdot |D_h^{X_s}u| g((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|\,dx \\
&\quad + c\int_\Omega \eta|\mathfrak{X}\eta| \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} \cdot \\
&\qquad\qquad \cdot |D_h^{Y_s}u| g((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u|\,dx \\
&\le \varepsilon\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\,dx \\
&\quad + \varepsilon\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} W((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u|^2\,dx \\
&+ c(n,p,L)\frac{C_\eta}{\varepsilon}\int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} g((D_h^{X_s}u)^2)\,dx \\
&\quad + c(n,p,L)\frac{C_\eta}{\varepsilon}\int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p}{2}} g((D_h^{Y_s}u)^2)\,dx \\
&\quad + c(n,p,L)\frac{C_\eta}{\varepsilon}\int_{\operatorname{supp}\eta} \sum_{s=1}^n \left[|D_h^{X_s}u|^p g((D_h^{X_s}u)^2) + |D_h^{Y_s}u|^p g((D_h^{Y_s}u)^2)\right]dx.
\end{aligned}$$

We observe that we have used the fact that g is non-decreasing to estimate $g \le W$ and we have set, similarly to (36)

$$C_\eta = \|\mathfrak{X}\eta\|^2_{L^\infty(\Omega)} + \|T\eta\|_{L^\infty(\Omega)} + 1. \tag{85}$$

We now estimate I_5 by treating separately the two resulting integrals. We have, using the bound in (81) and Young's inequality twice

$$\left|\int_\Omega 2\eta \sum_{s=1}^n \sum_{i=1}^{2n} A^s_{i,s+n} X_i\eta Tu(xe^{hX_s}) g((D_h^{X_s}u)^2) D_h^{X_s}u\,dx\right|$$

$$
\begin{aligned}
&\leq c(n,p,L)\int_{\operatorname{supp}\eta}|\mathfrak{X}\eta|\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}}\\
&\qquad\qquad\cdot|Tu(xe^{hX_s})||D_h^{X_s}u|g((D_h^{X_s}u)^2)\,dx\\
&\leq c(n,p,L)\|\mathfrak{X}\eta\|_{L^\infty(\Omega)}^{\frac{p}{p-1}}\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}\frac{p}{p-1}}\\
&\qquad\qquad\cdot|D_h^{X_s}u|^{\frac{p}{p-1}}g((D_h^{X_s}u)^2)\,dx\\
&\quad+c(n,p,L)\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}|Tu(xe^{hX_s})|^p g((D_h^{X_s}u)^2)\,dx\\
&\leq c\,C_\eta\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}}g((D_h^{X_s}u)^2)\,dx\\
&\quad+c\,C_\eta\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}|D_h^{X_s}u|^p g((D_h^{X_s}u)^2)\,dx\\
&\quad+c\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}|Tu(xe^{hX_s})|^p g((D_h^{X_s}u)^2)\,dx,
\end{aligned}\tag{86}
$$

where we estimated $\|\mathfrak{X}\eta\|_{L^\infty(\Omega)}^{\frac{p}{p-1}}\leq c\,C_\eta$ via Young's inequality and the very definition of C_η in (85), since $p\geq 2$. As for the second integral spreading from I_5, again using (81) and Young's inequality twice, with $\varepsilon\in(0,1)$, we have

$$
\begin{aligned}
&\left|\int_\Omega \eta^2\sum_{s=1}^{n}\sum_{i=1}^{2n}A^s_{i,s+n}X_iD_h^{X_s}u\,Tu(xe^{hX_s})W((D_h^{X_s}u)^2)\,dx\right|\\
&\leq c\int_\Omega\eta^2\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}}|\mathfrak{X}D_h^{X_s}u|\\
&\qquad\qquad\cdot|Tu(xe^{hX_s})|W((D_h^{X_s}u)^2)\,dx\\
&\leq \varepsilon\int_\Omega\eta^2\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}}W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\,dx\\
&\quad+\frac{c(n,p,L)}{\varepsilon}\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}}W((D_h^{X_s}u)^2)\,dx\\
&\quad+\frac{c(n,p,L)}{\varepsilon}\int_{\operatorname{supp}\eta}\sum_{s=1}^{n}|Tu(xe^{hX_s})|^pW((D_h^{X_s}u)^2)\,dx.
\end{aligned}\tag{87}
$$

Connecting the last two inequalities, and eventually changing ε, we find the final estimate for I_5, which is

$$
\begin{aligned}
&|I_5| \\
&\le \varepsilon \int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\, dx \\
&\quad + c\left(\frac{1}{\varepsilon} + C_\eta\right) \int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} \cdot \\
&\qquad\qquad\qquad\qquad\qquad\qquad \cdot W((D_h^{X_s}u)^2)\, dx \\
&\quad + c\, C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^n |D_h^{X_s}u|^p W((D_h^{X_s}u)^2)\, dx \\
&\quad + \frac{c}{\varepsilon} \int_{\operatorname{supp}\eta} \sum_{s=1}^n |Tu(xe^{hX_s})|^p W((D_h^{X_s}u)^2)\, dx,
\end{aligned}
$$

where the constant $c = c(n,p,L)$ does not depend on h, and $\varepsilon \in (0,1)$ is arbitrary. We have also used the trivial inequality $g \le W$. The estimate for I_6 is entirely similar, more precisely

$$
\begin{aligned}
&|I_6| \\
&\le \varepsilon \int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} W((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u|^2\, dx \\
&\quad + c\left(\frac{1}{\varepsilon} + C_\eta\right) \int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p}{2}} W((D_h^{Y_s}u)^2)\, dx \\
&\quad + c\, C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^n |D_h^{Y_s}u|^p W((D_h^{Y_s}u)^2)\, dx \\
&\quad + \frac{c}{\varepsilon} \int_{\operatorname{supp}\eta} \sum_{s=1}^n |Tu(xe^{hY_s})|^p W((D_h^{Y_s}u)^2)\, dx,
\end{aligned}
$$

with the same dependence upon the various constants for c and $\varepsilon \in (0,1)$. The estimate of the remaining integrals in (83) requires greater care in that we shall first need to let $h \to 0$ in the estimates we have derived up to now, and then estimate the resulting terms. We summarize the estimates obtained for the integrals $I_1 - I_6$ in the following new inequality, which is obtained choosing $\varepsilon \equiv \varepsilon(n,p,L)$ small enough and re-absorbing terms in (83):

$$II_1 + II_2 :=$$

$$
\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\,dx \\
&\quad + \int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} W((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u|^2\,dx \\
&\le c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} W((D_h^{X_s}u)^2)\,dx \\
&\quad + c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p}{2}} W((D_h^{Y_s}u)^2)\,dx \\
&\quad + c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^n \Big[|D_h^{X_s}u|^p W((D_h^{X_s}u)^2) + |D_h^{Y_s}u|^p W((D_h^{Y_s}u)^2)\Big]\,dx \\
&\quad + c \int_{\operatorname{supp}\eta} \sum_{s=1}^n \Big[|Tu(xe^{hX_s})|^p W((D_h^{X_s}u)^2) \\
&\qquad\qquad\qquad\qquad + |Tu(xe^{hY_s})|^p W((D_h^{Y_s}u)^2)\Big]\,dx \\
&\quad + |I_7| + |I_8| \\
&=: II_3 + II_4 + II_5 + II_6 + |I_7| + |I_8|. \qquad (88)
\end{aligned}
$$

Observe that in the previous inequality we have once again used the fact that $g \le W$.

Step 2: Letting $h \to 0$. Here we make the function g explicit with the choice

$$g(t) = g_{\alpha,k}(t) \qquad (89)$$

for fixed $\alpha \ge 0$ and $k \in \mathbb{N}$. With this choice we also have $W(t) = W_{\alpha,k}(t)$. Recall that the functions $g_{\alpha,k}(t)$ and $W_{\alpha,k}(t)$ have been defined in (54) and (60), respectively.

Remark 6. The choice of the function g in (89) can be also done in (35) instead of taking the unbouded function $g(t) = t^\alpha$, and this does not require the function u to be locally bounded in Ω. Therefore we can fully avoid the use of the regularity result in Theorem 4 also in the proof of vertical Lipschitz regularity from the previous section.

In the following, by "$h \to 0$" we shall mean $h_k \to 0$, where $\{h_k\}_{k\in\mathbb{N}}$ is a sequence chosen in such a way that, when $k \to \infty$ and $s \in \{1,\dots,n\}$, then

$$\begin{cases} \mathfrak{X}u(xe^{hX_s}), \mathfrak{X}u(xe^{hY_s}) \to \mathfrak{X}u(x) & \text{in} \quad L^p_{\text{loc}}(\Omega, \mathbb{R}^{2n}) \quad \text{and a.e.} \\ Tu(xe^{hX_s}), Tu(xe^{hY_s}) \to Tu(x) & \text{in} \quad L^p_{\text{loc}}(\Omega) \quad \text{and a.e.} \\ D_h^{X_s}\mathfrak{X}(x), D_h^{Y_s}\mathfrak{X}u(x) \to X_s\mathfrak{X}u(x), Y_s\mathfrak{X}u(x) & \text{in} \quad L^2_{\text{loc}}(\Omega, \mathbb{R}^{2n}) \quad \text{and a.e.} \end{cases} \tag{90}$$

This is actually possible: indeed, $(90)_2$ follows via Theorem 5, while $(90)_3$ comes from (26). Let us observe that the following additional convergence actually takes place:

$$\mathfrak{X}D_h^{X_s}u(x), \mathfrak{X}D_h^{Y_s}u(x) \to \mathfrak{X}X_s(x), \mathfrak{X}Y_su(x) \quad \text{in} \quad L^2_{\text{loc}}(\Omega, \mathbb{R}^{2n}) \quad \text{and a.e.} \tag{91}$$

Indeed, it suffices to prove the first one, the same argumentation works for the second; observe that when $k \neq s+n$ then $X_k(D_h^{X_s}u)(x) = D_h^{X_s}(X_ku)(x)$ and then the fact that $X_k(D_h^{X_s}u)(x) \to X_kX_su(x)$ follows from $(90)_3$. On the other hand, using (68) and $(90)_{2,3}$ we have $Y_s(D_h^{X_s}u)(x) = D_h^{X_s}(Y_su)(x) - Tu(xe^{hX_s}) \to X_sY_su(x) - Tu(xe^{hX_s}) = Y_sX_su(x)$, and (91) finally follows. And now we wish to pass to the limit in (88) for $h \to 0$. Using $(90)_{2,3}$, together with Fatou's lemma, we find

$$\begin{aligned} &\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \left[W((X_su)^2)|\mathfrak{X}X_su|^2 + W((Y_su)^2)|\mathfrak{X}Y_su|^2\right] dx \\ &\qquad \leq \liminf_{h\to 0} \ (II_1 + II_2). \end{aligned} \tag{92}$$

Using the fact that W is bounded, given by (61), and a well-know variant of Lebesgue's dominated convergence theorem, we gain

$$\begin{aligned} &\limsup_{h\to 0} \ (II_3 + II_4) \\ &\leq c\, C_\eta \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n \left[W((X_su)^2) + W((Y_su)^2)\right] dx. \end{aligned} \tag{93}$$

In the same way we have

$$\limsup_{h\to 0} II_5 \leq c\, C_\eta \int_{\operatorname{supp}\eta} |\mathfrak{X}u|^p \sum_{s=1}^n \left[W((X_su)^2) + W((Y_su)^2)\right] dx, \tag{94}$$

and, since $Tu \in L^p_{\text{loc}}(\Omega)$,

$$\limsup_{h\to 0} II_6 \leq c \int_{\operatorname{supp}\eta} |Tu|^p \sum_{s=1}^n \left[W((X_su)^2) + W((Y_su)^2)\right] dx. \tag{95}$$

Letting $h \to 0$ in the terms I_7 and I_8 requires greater care; in particular we shall use the upper bound in (51) for p. We shall give the details for I_7, the case of I_8 being completely analogous. Using growth conditions (3), (55) and Young's inequality we have that

$$\begin{aligned}&\left|a_{s+n}(\mathfrak{X}u)(2\eta\, T\eta\, g((D_h^{X_s}u)^2)D_h^{X_s}u)(xe^{-hX_s})\right|\\&\le c(n,p,L)k\|T\eta\|_{L^\infty(\Omega)}\left[(\mu^2+|\mathfrak{X}u(x)|^2)^{\frac{p}{2}}+|(D_h^{X_s}u)(xe^{-hX_s})|^p\right].\end{aligned}$$

Therefore, since $\mathfrak{X}u \in L^p_{\mathrm{loc}}(\Omega)$, and η has compact support in Ω, again applying a well-known variant of Lebesgue's dominated convergence theorem we get

$$\begin{aligned}\lim_{h\to 0}\int_\Omega a_{s+n}(\mathfrak{X}u)(2\eta T\eta g((D_h^{X_s}u)^2)D_h^{X_s}u)(xe^{-hX_s})\,dx\\=\int_\Omega a_{s+n}(\mathfrak{X}u)2\eta T\eta g((X_s u)^2)X_s u\,dx.\end{aligned}\tag{96}$$

As for the remaining term coming from I_7, we shall use the fact that $X_sTu \in L^2_{\mathrm{loc}}(\Omega)$ and therefore $D_h^{X_s}Tu \to X_sTu$ strongly in $L^2_{\mathrm{loc}}(\Omega)$. We shall argue as follows: start by changing variables

$$\begin{aligned}&\int_\Omega \sum_{s=1}^n a_{s+n}(\mathfrak{X}u)(\eta^2 W((D_h^{X_s}u)^2)D_h^{X_s}Tu)(xe^{-hX_s})\,dx\\&=\int_\Omega \eta^2\sum_{s=1}^n (a_{s+n}(\mathfrak{X}u))(xe^{hX_s})W((D_h^{X_s}u)^2)D_h^{X_s}Tu\,dx.\end{aligned}$$

Note that using again (3) and Young's inequality we find

$$\begin{aligned}&\left|(a_{s+n}(\mathfrak{X}u))(xe^{hX_s})W((D_h^{X_s}u)^2)D_h^{X_s}Tu\right|\\&\le c(\alpha,k,L)\left[(\mu^2+|\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{2p-2}{2}}+|D_h^{X_s}Tu|^2\right].\end{aligned}\tag{97}$$

Now we recall that $\mathfrak{X}u \in L^{\frac{pN}{N-2}}_{\mathrm{loc}}(\Omega)$ and observe that the bound in (51) implies also

$$2p-2 \le \frac{pN}{N-2}, \qquad \text{for } \ n \ge 1. \tag{98}$$

Indeed the last inequality is true for $n = 1, 2$ since $p < 4$, while for $n \ge 3$ it is equivalent to require that p satisfies (51). Consequently we always have

$$\mathfrak{X}u \in L^{2p-2}_{\mathrm{loc}}(\Omega). \tag{99}$$

Remark 7. When proving Thorem 1 the bound in (51) will be eventually replaced by the one in (6), which is more restrictive. In fact, we observe that

$$1 + n - \sqrt{n^2 - 2n - 3} < \frac{2(N-2)}{(N-4)} \qquad \text{when} \qquad n \geq 4. \tag{100}$$

While, when $n = 3$, we have

$$10/3 = 2(N-2)/(N-4) < 1 + n - \sqrt{n^2 - 2n - 3} = 4,$$

and this is the reason for the number 10/3 to appear in the definition of $c(n)$. For the same reason, when $n = 3$, we can allow to consider $p = 10/3$. Compare with the proof of Theorem 11 below given that the upper equality is allowed in (51).

By (99) we have that $\mathfrak{X}u(xe^{hX_s}) \to \mathfrak{X}u(x)$ strongly in $L^{2p-2}_{\mathrm{loc}}(\Omega)$, while by the second inclusion in (29) we have that $D_h^{X_s}Tu \to X_sTu$ in $L^2_{\mathrm{loc}}(\Omega)$. Therefore, with the help of (97) and the usual variant of Lebesgue's dominated convergence theorem, we gain

$$\begin{aligned}
&\lim_{h\to 0} \int_\Omega \sum_{s=1}^n a_{s+n}(\mathfrak{X}u)(\eta^2 W((D_h^{X_s}u)^2) D_h^{X_s}Tu)(xe^{-hX_s})\, dx \\
&= \lim_{h\to 0} \int_\Omega \eta^2 \sum_{s=1}^n (a_{s+n}(\mathfrak{X}u))(xe^{hX_s}) W((D_h^{X_s}u)^2) D_h^{X_s}Tu\, dx \\
&= \int_\Omega \eta^2 \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) W((X_su)^2) X_sTu\, dx.
\end{aligned} \tag{101}$$

Summarizing the results in (96) and (101) we obtain

$$\begin{aligned}
\lim_{h\to 0} I_7 = &-\int_\Omega 2\eta T\eta \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) g((X_su)^2) X_su\, dx \\
&- \int_\Omega \eta^2 \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) W((X_su)^2) X_sTu\, dx.
\end{aligned} \tag{102}$$

And in a completely analogous way we find that

$$\begin{aligned}
\lim_{h\to 0} I_8 = &\int_\Omega 2\eta T\eta \sum_{s=1}^n a_s(\mathfrak{X}u) g((Y_su)^2) Y_su\, dx \\
&+ \int_\Omega \eta^2 \sum_{s=1}^n a_s(\mathfrak{X}u) W((Y_su)^2) Y_sTu\, dx.
\end{aligned} \tag{103}$$

In turn, using (92), (93), (94), (95) and (102)-(103) in (88), we easily get

$$
\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^n (\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}} \left[W((X_su)^2)|\mathfrak{X}X_su|^2 + W((Y_su)^2)|\mathfrak{X}Y_su|^2\right]\, dx \\
&\le c\, C_\eta \int_{\operatorname{supp}\eta} (\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n \left[W((X_su)^2)+W((Y_su)^2)\right]\, dx \\
&\quad + c\int_{\operatorname{supp}\eta} |Tu|^p \sum_{s=1}^n \left[W((X_su)^2)+W((Y_su)^2)\right]\, dx. \\
&\qquad + \left|\int_\Omega 2\eta\, T\eta\, \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) g((X_su)^2) X_su\, dx\right| \qquad (=: III_1) \\
&\qquad + \left|\int_\Omega \eta^2 \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) W((X_su)^2) X_sTu\, dx\right| \qquad (=: III_2) \\
&\qquad + \left|\int_\Omega 2\eta\, T\eta\, \sum_{s=1}^n a_{s}(\mathfrak{X}u) g((Y_su)^2) Y_su\, dx\right| \qquad (=: III_3) \\
&\qquad + \left|\int_\Omega \eta^2 \sum_{s=1}^n a_{s}(\mathfrak{X}u) W((Y_su)^2) Y_sTu\, dx\right| \qquad (=: III_4)
\end{aligned} \tag{104}
$$

and it remains to estimate the new defined quantities $III_1, \dots, III_4$.

Step 3: Further commutations and estimates. For III_1 and III_3 it suffices to use (3) plus an elementary estimation involving $g \le W$. Indeed we have

$$
\begin{aligned}
&III_1 + III_3 \\
&\le c\|T\eta\|_{L^\infty(\Omega)} \int_{\operatorname{supp}\eta} (\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n \left[W((X_su)^2)+W((Y_su)^2)\right]\, dx.
\end{aligned} \tag{105}
$$

The estimation of III_2 and III_4 requires additional integration by parts. We shall give the details for III_2 only, the case of III_4 being completely analogous. Integrating by parts (since $X_i^* = -X_i$) gives

$$
\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) W((X_su)^2) X_sTu\, dx \\
&= -\int_\Omega \eta^2 \sum_{s=1}^n \sum_{j=1}^{2n} D_{z_j} a_{s+n}(\mathfrak{X}u) X_sX_ju W((X_su)^2) Tu\, dx \\
&\quad - \int_\Omega 2\eta \sum_{s=1}^n a_{s+n}(\mathfrak{X}u) X_s\eta W((X_su)^2) Tu\, dx
\end{aligned}
$$

$$
\begin{aligned}
&-\int_{\Omega} 2\eta^2 \sum_{s=1}^{n} a_{s+n}(\mathfrak{X}u) W'((X_s u)^2) X_s u X_s X_s u T u \, dx \\
&=: IV_1 + IV_2 + IV_3. \qquad (106)
\end{aligned}
$$

Before continuing with the estimates of the last three integrals we actually need to justify the previous integration by parts. This follows by first showing that

$$
\eta^2 a_{s+n}(\mathfrak{X}u) W((X_s u)^2) \in HW_0^{1,\frac{p}{p-1}}(\Omega) \cap L^2(\Omega) \qquad \text{for } s \in \{1, \dots, n\}, \tag{107}
$$

and then applying Lemma 5 below, with the obvious choice

$$
f = \eta^2 a_{s+n}(\mathfrak{X}u) W((X_s u)^2) \qquad \text{and} \qquad g = Tu.
$$

Obviously, f has compact support in Ω since so has η. Using (3) and (61) we have that

$$
\begin{aligned}
\int_{\Omega} |\eta^2 a_{s+n}(\mathfrak{X}u) W((X_s u)^2)|^2 \, dx &\leq c(\alpha, k, L) \int_{\Omega} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{2p-2}{2}} \, dx \\
&\leq c \int_{\Omega} (\mu^2 + |\mathfrak{X}u|)^{\frac{pN}{N-2}} \, dx < \infty,
\end{aligned}
$$

the last inequality being a consequence of (98). Next we prove that

$$
\eta^2 a_{s+n}(\mathfrak{X}u) W((X_s u)^2) \in HW_0^{1,\frac{p}{p-1}}(\Omega), \tag{108}
$$

taking into account that the derivatives of this expression have already been calculated in (106); this will finally end the proof of (107) and the justification of the integration by parts in (106). Using the fact that W is bounded (61), and taking into consideration (3), we have

$$
\begin{aligned}
&\int_{\Omega} |\eta^2 D_{z_j} a_{s+n}(\mathfrak{X}u) X_s X_j u W((X_s u)^2)|^{\frac{p}{p-1}} dx \\
&\leq c(\alpha, k, L) \int_{\operatorname{supp} \eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{(p-2)p}{2(p-1)}} |\mathfrak{X}\mathfrak{X}u|^{\frac{p}{p-1}} \, dx \\
&\leq c \int_{\operatorname{supp} \eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} |\mathfrak{X}\mathfrak{X}u|^2 + (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \, dx, \qquad (109)
\end{aligned}
$$

the last integral being finite by (24). Note that in the case $p > 2$ we used Young's inequality with conjugate exponents $2(p-1)/p$ and $2(p-1)/(p-2)$. Using (3) and (61), we have immediately

$$
\int_{\Omega} |2\eta \sum_{s=1}^{n} a_{s+n}(\mathfrak{X}u) X_s \eta W((X_s u)^2)|^{\frac{p}{p-1}} \, dx
$$

$$\leq c(\alpha,k,L)\|\mathfrak{X}\eta\|_{L^\infty(\Omega)}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\,dx.$$

Finally, using (62) and again estimating as for (109), we have

$$\begin{aligned}
&\int_\Omega |2\eta^2 a_{s+n}(\mathfrak{X}u)W'((X_su)^2)X_suX_sX_su|^{\frac{p}{p-1}}\,dx\\
&\leq c(\alpha,k,L)\int_{\operatorname{supp}\eta}\left[(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-1}{2}}\frac{W((X_su)^2)}{1+(X_su)^2}|X_su||X_sX_su|\right]^{\frac{p}{p-1}}\,dx\\
&\leq c\int_{\operatorname{supp}\eta}\left[(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}W((X_su)^2)|X_sX_su|\right]^{\frac{p}{p-1}}\,dx\\
&\leq c\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{(p-2)p}{2(p-1)}}|\mathfrak{X}\mathfrak{X}u|^{\frac{p}{p-1}}\,dx\\
&\leq c\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}|\mathfrak{X}\mathfrak{X}u|^2+(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\,dx<\infty.
\end{aligned}$$

Now, once the previous inequalities have been established, using their analogs at the level of finite difference quotients allows to prove (108) with $p/(p-1)$ replaced by 1. Here we use the fact that $a_{s+n}(\mathfrak{X}u)\in HW^{1,\frac{p}{p-1}}_{\mathrm{loc}}(\Omega)$ (as in (109)), $W((X_su)^2)\in HW^{1,2}_{\mathrm{loc}}(\Omega)$ (as $\mathfrak{X}\mathfrak{X}u\in L^2_{\mathrm{loc}}(\Omega)$), and $\mathfrak{X}u\in L^{p-2}_{\mathrm{loc}}(\Omega)$. Then the previous inequalities alllows to get the desired integrability. Thus, we have established (108), and therefore (107), as well as (106).

Remark 8. Observe that when checking (107) and then (106), we did not use the fact that $Tu\in L^\infty(B_r)$, but just used that $Tu\in L^p(B_r)$ together with the bound in (51), the higher differentiability and integrability results of Theorems 5, and Theorem 6, respectively. This will turn out to be crucial later, in the proof of Theorem 6.1 below. The fact that we use only $Tu\in L^p(B_r)$ here is reflected in that we make estimates in $HW^{1,\frac{p}{p-1}}_0(\Omega)$, rather than in $HW^{1,1}_0(\Omega)$, when considering (107).

Next we estimate the terms coming up from (106) starting with IV_1. Commuting again we find

$$\begin{aligned}
IV_1=&-\int_\Omega\eta^2\sum_{s=1}^n\sum_{j=1}^{2n}D_{z_j}a_{s+n}(\mathfrak{X}u)X_jX_suW((X_su)^2)Tu\,dx\\
&-\int_\Omega\eta^2\sum_{s=1}^n D_{z_{s+n}}a_{s+n}(\mathfrak{X}u)W((X_su)^2)(Tu)^2\,dx\\
&=:V_1+V_2.
\end{aligned}\tag{110}$$

Then, using the growth conditions (3), and Young's inequality twice, we have

$$
\begin{aligned}
|V_1| &\le c\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}|Tu|\sum_{s=1}^n W((X_su)^2)|\mathfrak{X}X_su|\,dx\\
&\le \varepsilon\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\sum_{s=1}^n W((X_su)^2)|\mathfrak{X}X_su|^2\,dx\\
&\quad +\frac{c}{\varepsilon}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\sum_{s=1}^n W((X_su)^2)\,dx\\
&\quad +\frac{c}{\varepsilon}\int_{\operatorname{supp}\eta}|Tu|^p\sum_{s=1}^n W((X_su)^2)\,dx,
\end{aligned}
$$

with $\varepsilon\in(0,1)$. Similarly, using again that $\|\mathfrak{X}\eta\|^{\frac{p}{p-1}}_{L^\infty(\Omega)}\le c\,C_\eta$ via Young's inequality, we also obtain

$$
\begin{aligned}
|V_2|+|IV_2| \le c(1+C_\eta)&\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\sum_{s=1}^n W((X_su)^2)\,dx\\
&+c\int_{\operatorname{supp}\eta}|Tu|^p\sum_{s=1}^n W((X_su)^2)\,dx.
\end{aligned}
$$

For IV_3 we shall need to use the inequality (62). In particular the fact that the constant involved there depends only on α in the explicit way reported there, and is also independent of k. This will be crucial for the subsequent iteration in Step 5 below. Using (3) and Young's inequality, we find

$$
\begin{aligned}
|IV_3| &\le c\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-1}{2}}\sum_{s=1}^n |W'((X_su)^2)||X_su||X_sX_su||Tu|\,dx\\
&\le \frac{\varepsilon}{3(\alpha+1)}\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\sum_{s=1}^n |W'((X_su)^2)||X_su|^2|\mathfrak{X}X_su|^2\,dx\\
&\quad +\frac{c(\alpha+1)}{\varepsilon}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}|Tu|^2\sum_{s=1}^n |W'((X_su)^2)|\,dx\\
&\le \varepsilon\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\sum_{s=1}^n W((X_su)^2)|\mathfrak{X}X_su|^2\,dx\\
&\quad +\frac{c(\alpha+1)}{\varepsilon}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}|Tu|^2\sum_{s=1}^n |W'((X_su)^2)|\,dx.
\end{aligned}
$$

In the last estimate we made crucial use of (62). Combining the estimates for IV_1, IV_2, IV_3, V_1 and V_2 we recover the desired estimate for III_2. A completely analogous estimate can be worked out for III_4 so that, at the end we find

$$\begin{aligned}
&|III_2| + |III_4| \\
&\le \varepsilon \int_\Omega \eta^2(\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \cdot \\
&\qquad\qquad \cdot \sum_{s=1}^n \left[W((X_s u)^2)|\mathfrak{X}X_s u|^2 + W((Y_s u)^2)|\mathfrak{X}Y_s u|^2\right] dx \\
&\quad + \frac{cC_\eta}{\varepsilon} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n \left[W((X_s u)^2) + W((Y_s u)^2)\right] dx \\
&\quad + \frac{c}{\varepsilon} \int_{\operatorname{supp}\eta} |Tu|^p \sum_{s=1}^n \left[W((X_s u)^2) + W((Y_s u)^2)\right] dx \\
&\quad + \frac{c(\alpha+1)}{\varepsilon} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} |Tu|^2 \cdot \\
&\qquad\qquad \cdot \sum_{s=1}^n \left[|W'((X_s u)^2)| + |W'((Y_s u)^2)|\right] dx. \qquad (111)
\end{aligned}$$

We point out that the constant c in the above estimate depends only on n, p and L. It is in particular independent of both α and k. The dependence on α in the previous inequality has been explicitly calculated.

Step 4: Letting $k \to \infty$. Taking into consideration the estimates for the quantities $III_1, \ldots, III_4$, using them in (104), and taking $\varepsilon \equiv \varepsilon(n, p, L)$ small enough in order to re-absorb terms on the left hand side, we gain

$$\begin{aligned}
&\int_\Omega \eta^2(\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \sum_{s=1}^n \left[W((X_s u)^2)|\mathfrak{X}X_s u|^2 + W((Y_s u)^2)|\mathfrak{X}Y_s u|^2\right] dx \\
&\le c\, C_\eta \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n \left[W((X_s u)^2) + W((Y_s u)^2)\right] dx \\
&\quad + c \int_{\operatorname{supp}\eta} |Tu|^p \sum_{s=1}^n \left[W((X_s u)^2) + W((Y_s u)^2)\right] dx \\
&\quad + c(\alpha+1)^2 \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} |Tu|^2 \cdot \\
&\qquad\qquad \cdot \sum_{s=1}^n \left[\frac{|W((X_s u)^2)|}{1 + (X_s u)^2} + \frac{|W((Y_s u)^2)|}{1 + (Y_s u)^2}\right] dx. \qquad (112)
\end{aligned}$$

As already remarked in the previous step, the constants involved in the former estimate are independent of $k \in \mathbb{N}$. In particular, the constant c only depends on n, p and L. Observe also that the last term has been obtained using again (62) in order to estimate the last integral coming from (111). We are now ready to pass to the limit in the above estimate, using (63)-(64). The left hand side is treated via Fatou's lemma, while, thanks to (63) we can use monotone convergence theorem on the right hand side, finally obtaining, after an elementary estimation based also on (64)

$$\int_{\Omega} \eta^2 \sum_{s=1}^{n} \left[|X_s u|^{p-2+2\alpha} |\mathfrak{X} X_s u|^2 + |Y_s u|^{p-2+2\alpha} |\mathfrak{X} Y_s u|^2 \right] dx$$
$$\leq c(\alpha+1)^3 C_T C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} 1 + \left[|X_s u|^{p+2\alpha} + |Y_s u|^{p+2\alpha} \right] dx. \tag{113}$$

In the previous inequality we have denoted

$$C_T = \left(1 + \|Tu\|_{L^\infty(B_r)}\right)^p. \tag{114}$$

We observe that in order to get (113) we have used the following pointwise inequality:

$$(\mu^2 + |\mathfrak{X} u|^2)^{\frac{p}{2}} \sum_{s=1}^{n} \left[(X_s u)^{2\alpha} + (Y_s u)^{2\alpha} \right]$$
$$\leq c(n,p) \sum_{s=1}^{n} 1 + \left[(X_s u)^{p+2\alpha} + (Y_s u)^{p+2\alpha} \right], \tag{115}$$

valid for every $\alpha \geq 0$, which in turn elementary follows from a repeated application of Young's inequality. It is crucial that in the above inequality the constant involved is independent of α, as it will be clear in the following computations.

Step 5: Final iteration. Inequality (113) is the starting point for running a Moser's iteration procedure with a sequence of exponents $\alpha = \alpha_k$, that will implicitly lead to to conclude that each of the integrals we consider with is finite. In the following, whenever we work with an integrand containing α, we shall argue assuming that the corresponding integral is finite. This fact can be checked, step by step, along the iteration procedure. Let us first observe that

$$\left| \mathfrak{X} \left(\eta^2 |X_s u|^{\frac{p+2\alpha}{2}} + \eta^2 |Y_s u|^{\frac{p+2\alpha}{2}} \right) \right|^2$$
$$\leq c\, \alpha^2 \eta^2 \left[|X_s u|^{p-2+2\alpha} |\mathfrak{X} X_s u|^2 + |Y_s u|^{p-2+2\alpha} |\mathfrak{X} Y_s u|^2 \right]$$

$$+c\,C_\eta\left[|X_s u|^{p+2\alpha}+|Y_s u|^{p+2\alpha}\right]. \qquad (116)$$

Combining the last inequality with (113), taking into account the dependence on α in the inequalities (113) and (116), using Theorem 3 with $q=2$, and (113), we find that

$$\left(\int_\Omega \eta^{\frac{2Q}{Q-2}}\sum_{s=1}^n\left[|X_s u|^{p+2\alpha}+|Y_s u|^{p+2\alpha}\right]^{\frac{Q}{Q-2}}\,dx\right)^{\frac{Q-2}{Q}}$$
$$\leq c(\alpha+1)^5 C_T C_\eta\int_{\operatorname{supp}\eta} 1+\sum_{s=1}^n\left[|X_s u|^{p+2\alpha}+|Y_s u|^{p+2\alpha}\right]\,dx. \qquad (117)$$

As in Theorem 8, we consider concentric balls, centered at x_0. As in the previous section, whenever $\rho<R$ we define the family of concentric interpolating balls $B_\rho\subset B_{\rho_{k+1}}\subset B_{\rho_k}\subset B_R$, according to the choice of the radii already made in (49). We next take a family of smooth cut-off functions $\{\eta_k\}_k\subset C_0^\infty(B_{\rho_k})$ in such a way that $0\leq\eta_k\leq 1$, $\eta_k\equiv 1$ on $B_{\rho_{k+1}}$, and

$$1\leq\|\mathfrak{X}\eta_k\|^2_{L^\infty(\Omega)}+\|T\eta_k\|_{L^\infty(\Omega)}\leq\frac{\gamma(n)^k}{(r-\rho)^2},$$

for every $k\in\mathbb{N}$, and where $\gamma(n)$ is an absolute constant. The existence of such a family of functions can be inferred, once again, from [5]. Keeping into account (85) this also yields

$$C_{\eta_k}=C_\eta\leq\frac{c^k}{(r-\rho)^2}. \qquad (118)$$

Now we inductively define the sequence $\{\alpha_k\}_k$ according to

$$\begin{cases}\alpha_{k+1} = \tilde{\chi}\alpha_k+\frac{p}{Q-2}\\ \alpha_0 = 0,\end{cases} \qquad (119)$$

where this time we have set

$$\tilde{\chi}=\frac{Q}{Q-2}>1.$$

We notice the following relation

$$\alpha_k=\frac{p}{Q-2}\sum_{j=0}^{k-1}\tilde{\chi}^j=\frac{p}{2}(\tilde{\chi}^k-1), \qquad \text{for every} \qquad k\geq 1,$$

so that

$$p+2\alpha_{k+1}=\tilde{\chi}(p+2\alpha_k) \qquad p+2\alpha_k=p\tilde{\chi}^k \qquad \text{for } \; k\geq 0. \qquad (120)$$

We let

$$A_k := \left(\int_{B_{\rho_k}} 1 + \sum_{s=1}^{n} \left[|X_s u|^{p+2\alpha_k} + |Y_s u|^{p+2\alpha_k} \right] \, dx \right)^{\frac{1}{p+2\alpha_k}} ,$$

so that, after a few elementary manipulations, taking into consideration (118), (119), and (120), inequality (117) gives

$$A_{k+1} \leq \left[c(n,p,L)\tilde{\chi}^5 \right]^{\frac{k}{p\tilde{\chi}^k}} \left[\frac{C_T}{(r-\rho)^2} \right]^{\frac{1}{p\tilde{\chi}^k}} A_k, \qquad \text{for } \ k \geq 0. \tag{121}$$

The constant C_T has been introduced in (114). Noticing that

$$\sum_{j=0}^{\infty} \frac{1}{\tilde{\chi}^j} = \frac{\tilde{\chi}}{\tilde{\chi}-1} = \frac{Q}{2},$$

the iteration of (121) and further elementary estimates finally yield

$$\begin{aligned} A_{k+1} &\leq \left[c(n,p,L)\tilde{\chi}^5 \right]^{\frac{1}{p} \sum_{j=0}^{\infty} \frac{j}{\tilde{\chi}^j}} \left[\frac{C_T}{(r-\rho)^2} \right]^{\frac{1}{p} \sum_{j=0}^{\infty} \frac{1}{\tilde{\chi}^j}} A_0 \\ &\leq \left(\frac{c(n,p,L)}{r-\rho} \right)^{\frac{Q}{p}} \left(1 + \|Tu\|_{L^\infty(B_r)} \right)^{\frac{Q}{2}} \left(\int_{B_r} (1+|\mathfrak{X}u|^2)^{\frac{p}{2}} \, dx \right)^{\frac{1}{p}} , \end{aligned}$$

for every $k \in \mathbb{N}$, so that (52) follows letting $k \to \infty$ in the previous relation. The local boundedness of the horizontal derivatives follows via the standard covering argument. The proof is concluded. $\square$

The following integration by parts elementary lemma has been used above, when checking the validity of (106). Its proof, based on an easy approximation argument, is standard and therefore omitted.

Lemma 5. *Let* $f \in HW_0^{1,\frac{p}{p-1}}(\Omega) \cap L^2(\Omega)$ *be such that* $\operatorname{supp} f \subset\subset \Omega$*, and* $g \in HW_0^{1,2}(\Omega) \cap L^p(\Omega)$*, where* $p > 1$*. Then*

$$\int_{\Omega} f \, (X_i g) \, dx = - \int_{\Omega} (X_i f) \, g \, dx \qquad \textit{for } i \in \{1, \ldots, 2n\}.$$

6. Further higher integrability

Here we prove further higher integrability of the horizontal gradient of solutions, concentrating on the case $p \neq 2$ for technical reasons. We are here concerned with the fourth arrow in (11):

Theorem 10. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(5)*, and assume that $p \in (2,4)$ also satisfies the restriction in* (51)*. Then*

$$\mathfrak{X}u \in L^s_{\mathrm{loc}}(\Omega) \qquad \textit{for every } s < \frac{Qp}{Q-p}. \tag{122}$$

Remark 9. Notice the fact that the previous result is not always an improvement of Theorem 6, indeed

$$\frac{Qp}{Q-p} \geq \frac{Np}{N-2} \qquad \text{if and only if} \qquad p \geq \frac{2Q}{N}, \tag{123}$$

and this is due to the fact that here we are not using the existence of vertical derivatives of $\mathfrak{X}u$ to get higher integrability, as we did for Theorem 6. Moreover, the previous result relies again on Theorem 6 itself that provides the starting higher integrability of $\mathfrak{X}u$ needed to make rigorous the integration by parts already encountered in (106), a fact that will also come into the play in the proof of Theorem 10. The bound obtained in (122) is therefore better than the one in (28) only if p is large enough, according to (123). This fact goes in the "right direction": indeed, it is exactly what we need in the following, since we are trying to increase the values of p for which regularity of solutions holds (compare with Remark 4.)

Proof (of Theorem 10). Up to a certain stage, the proof follows the one for Theorem 9 in the previous section, and we shall indicate the main modifications, keeping the notation and the terminology introduced there. We re-start from (83) and estimate the integrals $I_1, \ldots, I_4$ as in Theorem 9. The remaining terms will be estimated in a different way, using only the fact that $Tu \in L^p_{\mathrm{loc}}(\Omega)$, rather than in $Tu \in L^\infty_{\mathrm{loc}}(\Omega)$, which is not available as an assumption here. We start with I_5. For the integral in the first line of (86) we have, using Young's inequality twice

$$\begin{aligned}
&\left| \int_\Omega 2\eta \sum_{s=1}^{n} \sum_{i=1}^{2n} A^s_{i,s+n} X_i\eta Tu(xe^{hX_s}) g((D_h^{X_s}u)^2) D_h^{X_s}u \, dx \right| \\
&\leq c(n,p,L)\|\mathfrak{X}\eta\|^{\frac{p}{p-1}}_{L^\infty(\Omega)} \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}\frac{p}{p-1}} \\
&\qquad\qquad \cdot |D_h^{X_s}u|^{\frac{p}{p-1}} g((D_h^{X_s}u)^2)^{\frac{p}{p-1}} \, dx \\
&\qquad + c(n,p,L) \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} |Tu(xe^{hX_s})|^p \, dx
\end{aligned}$$

$$\leq c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} g((D_h^{X_s}u)^2)^{\frac{p}{p-2}}\,dx$$
$$+c \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} \Big[C_\eta |D_h^{X_s}u|^p + |Tu(xe^{hX_s})|^p\Big]\,dx.$$

The quantity C_η is still the one introduced in (85). As for the second integral spreading from I_5, that is the one in the first line of (87), we have, again by Young's inequality with $\varepsilon \in (0,1)$

$$\left| \int_\Omega \eta^2 \sum_{s=1}^{n} \sum_{i=1}^{2n} A^s_{i,s+n} X_i D_h^{X_s} u Tu(xe^{hX_s}) W((D_h^{X_s}u)^2)\,dx \right|$$
$$\leq \varepsilon \int_\Omega \eta^2 \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} \cdot$$
$$\cdot W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\,dx$$
$$+\frac{c(n,p,L)}{\varepsilon} \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} \cdot$$
$$\cdot W((D_h^{X_s}u)^2)^{\frac{p}{p-2}}\,dx$$
$$+\frac{c(n,p,L)}{\varepsilon} \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} |Tu(xe^{hX_s})|^p\,dx.$$

The term I_6 can be estimate in a completely similar fashion. Finally, combining the newly found estimates for I_5 and I_6 to (83), and re-absorbing terms as usually taking ε small enough, we obtain the following analog to (88):

$$\int_\Omega \eta^2 \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p-2}{2}} W((D_h^{X_s}u)^2)|\mathfrak{X}D_h^{X_s}u|^2\,dx$$
$$+\int_\Omega \eta^2 \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p-2}{2}} W((D_h^{Y_s}u)^2)|\mathfrak{X}D_h^{Y_s}u|^2\,dx$$
$$\leq c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hX_s})|^2)^{\frac{p}{2}} W((D_h^{X_s}u)^2)^{\frac{p}{p-2}}\,dx$$
$$+c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} (\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hY_s})|^2)^{\frac{p}{2}} W((D_h^{Y_s}u)^2)^{\frac{p}{p-2}}\,dx$$
$$+c\,C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{n} \Big[|D_h^{X_s}u|^p W((D_h^{X_s}u)^2)^{\frac{p}{p-2}}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad +|D_h^{Y_s}u|^p W((D_h^{Y_s}u)^2)^{\frac{p}{p-2}}\Big]\,dx \\
&\quad +c\int_{\operatorname{supp}\eta}\sum_{s=1}^n\Big[|Tu(xe^{hX_s})|^p+|Tu(xe^{hY_s})|^p\Big]\,dx \\
&\quad +|I_7|+|I_8| \\
&=: II_3+II_4+II_5+II_6+|I_7|+|I_8|.
\end{aligned}
$$

Needless to say, I_7 and I_8 are the ones defined in (83). In the present setting, this estimate replaces (88). We have used the following elementary inequalities:

$$
g\le W, \qquad\qquad W+W^{\frac{p}{p-1}}\le(4W)^{\frac{p}{p-2}}. \tag{124}
$$

Exactly as is Step 2 in the proof of Theorem 9, we are able to let $h\to 0$. This is possible since W is a globally bounded function by (61). Proceeding exactly as in Step 2 in the proof of Theorem 9, in particular, in a totally similar way for the terms I_7 and I_8, we arrive at the following replacement of (104), which comes after a few routine estimations

$$
\begin{aligned}
&\int_\Omega \eta^2\sum_{s=1}^n(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\left[W((X_su)^2)|\mathfrak{X}X_su|^2+W((Y_su)^2)|\mathfrak{X}Y_su|^2\right]\,dx \\
&\le c\,C_\eta\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\sum_{s=1}^n\Big[W((X_su)^2)^{\frac{p}{p-2}}+W((Y_su)^2)^{\frac{p}{p-2}}\Big]\,dx \\
&\qquad +c\int_{\operatorname{supp}\eta}|Tu|^p\,dx+\sum_{1=1}^4 III_i.
\end{aligned} \tag{125}
$$

Before going on, we recall that in the proof of Theorem 9, the convergence of the terms I_7 and I_8, did not require the boundedness of Tu, but just the fact $\mathfrak{X}Tu\in L^2_{\mathrm{loc}}(\Omega)$, and the bound (51), together with the higher integrability result from Theorem 6 (compare with (97)-(98)), as also observe in Remark 8. All such ingredients are available here too.

The terms III_i in (125), are the one specified in (104). These will also be estimated in a different way. We have, again using the second inequality contained in (124)

$$
\begin{aligned}
&III_1+III_3 \\
&\le c\,C_\eta\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\,\cdot \\
&\qquad\qquad \cdot\sum_{s=1}^n\Big[W((X_su)^2)^{\frac{p}{p-2}}+W((Y_su)^2)^{\frac{p}{p-2}}\Big]\,dx.
\end{aligned} \tag{126}
$$

We now estimate III_2. First we remark that the estimation of III_2 goes through the integration-by-parts procedure followed in (106). As already stressed above, to be performed, this only requires the ingredients described in Remark 8, which are available under the assumptions of the present theorem. In particular, we do not need to use the local boundedness of Tu here, it just suffices to have $Tu \in L^p_{\mathrm{loc}}(\Omega)$, which is a consequence of Theorem 5. Observe also that this is exactly the point where the bound on p assumed in the statement comes into the play; in particular this is assumed to meet condition (98). We can therefore start estimating III_2 using (106), treating the corresponding terms $IV_1, IV_2, IV_3, V_1, V_2$, taking into account the expansion in (110). We have, again by Young's inequality

$$\begin{aligned}|V_1| &\le \varepsilon \int_\Omega \eta^2 (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \sum_{s=1}^n W((X_s u)^2) |\mathfrak{X} X_s u|^2 \, dx \\ &+ \frac{c}{\varepsilon} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n W((X_s u)^2)^{\frac{p}{p-2}} \, dx \\ &+ \frac{c}{\varepsilon} \int_{\operatorname{supp}\eta} |Tu|^p \, dx,\end{aligned}$$

and, again using the second inequality in (124)

$$\begin{aligned}|V_2| + |IV_2| &\le c \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n W((X_s u)^2)^{\frac{p}{p-2}} \, dx \\ &\quad + c\|\mathfrak{X}\eta\|^{\frac{p}{p-1}} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n W((X_s u)^2)^{\frac{p}{p-1}} \, dx \\ &\quad + c \int_{\operatorname{supp}\eta} |Tu|^p \, dx \\ &\le c(1 + C_\eta) \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p}{2}} \sum_{s=1}^n W((X_s u)^2)^{\frac{p}{p-2}} \, dx \\ &\quad + c \int_{\operatorname{supp}\eta} |Tu|^p \, dx.\end{aligned}$$

Finally, again using Young's inequality twice, we obtain

$$\begin{aligned}|IV_3| &\le \frac{\varepsilon}{3(\alpha+1)} \int_\Omega \eta^2 (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \sum_{s=1}^n |W'((X_s u)^2)| |X_s u|^2 |\mathfrak{X} X_s u|^2 \, dx \\ &+ \frac{c(\alpha+1)}{\varepsilon} \int_{\operatorname{supp}\eta} (\mu^2 + |\mathfrak{X}u|^2)^{\frac{p^2}{2(p-2)}} \sum_{s=1}^n |W'((X_s u)^2)|^{\frac{p}{p-2}} \, dx\end{aligned}$$

$$+\frac{c(\alpha+1)}{\varepsilon}\int_{\operatorname{supp}\eta}|Tu|^p\,dx.$$

Completely analogous considerations and estimates work for III_4. Making use of (62) we arrive at

$$\begin{aligned}
&|III_2|+|III_4|\\
&\le \varepsilon\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\cdot\\
&\qquad\cdot\sum_{s=1}^n\left[W((X_su)^2)|\mathfrak{X}X_su|^2+W((Y_su)^2)|\mathfrak{X}Y_su|^2\right]\,dx\\
&\quad+\frac{cC_\eta}{\varepsilon}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\sum_{s=1}^n\left[W((X_su)^2)^{\frac{p}{p-2}}+W((Y_su)^2)^{\frac{p}{p-2}}\right]\,dx\\
&\quad+\frac{c(\alpha+1)}{\varepsilon}\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p^2}{2(p-2)}}\cdot\\
&\qquad\cdot\sum_{s=1}^n\left[|W'((X_su)^2)|^{\frac{p}{p-2}}+|W'((Y_su)^2)|^{\frac{p}{p-2}}\right]\,dx\\
&\quad+\frac{c(\alpha+1)}{\varepsilon}\int_{\operatorname{supp}\eta}|Tu|^p\,dx.
\end{aligned}$$

Connecting this last inequality, (125), (126), taking ε small enough in the standard way, and re-absorbing terms we obtain the following analog of (112):

$$\begin{aligned}
&\int_\Omega \eta^2(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{2}}\sum_{s=1}^n\left[W((X_su)^2)|\mathfrak{X}X_su|^2+W((Y_su)^2)|\mathfrak{X}Y_su|^2\right]\,dx\\
&\le c\,C_\eta\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p}{2}}\sum_{s=1}^n\left[W((X_su)^2)^{\frac{p}{p-2}}+W((Y_su)^2)^{\frac{p}{p-2}}\right]\,dx\\
&\quad+c(\alpha+1)^2\int_{\operatorname{supp}\eta}(\mu^2+|\mathfrak{X}u|^2)^{\frac{p^2}{p-2}}\cdot\\
&\qquad\cdot\sum_{s=1}^n\left[\left|\frac{W((X_su)^2)}{1+(X_su)^2}\right|^{\frac{p}{p-2}}+\left|\frac{W((Y_su)^2)}{1+(Y_su)^2}\right|^{\frac{p}{p-2}}\right]\,dx\\
&\quad+c(\alpha+1)\int_{\operatorname{supp}\eta}|Tu|^p\,dx,
\end{aligned}$$

where again we made use of (62) to treat the next to last integral. The constant appearing in the last inequality is independent of k, and the dependence on α is the one explicitly computed. Once again we recall that we are taking $W = W_{\alpha,k}$; therefore, by virtue of (63), the monotone convergence theorem applied on the right hand side, and of Fatou's lemma applied on the left one, we let $k \to \infty$, getting, after an elementary estimation

$$\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^n \left[|X_s u|^{p-2+2\alpha}|\mathfrak{X}X_s u|^2 + |Y_s u|^{p-2+2\alpha}|\mathfrak{X}Y_s u|^2\right]\, dx \\
&\leq c(\alpha,\eta)\int_{\operatorname{supp}\eta} \sum_{s=1}^n 1 + \left[|X_s u|^{\frac{p}{p-2}(p-2+2\alpha)} + |Y_s u|^{\frac{p}{p-2}(p-2+2\alpha)}\right]\, dx \\
&\quad + c(\alpha)\int_{\operatorname{supp}\eta} |Tu|^p\, dx.
\end{aligned}$$

In the previous inequality, the constant c depends of course also on n, p and L (we recall that by the initial re-scaling we are assuming $\nu = 1$). In turn, applying Sobolev embedding theorem in the sub-elliptic version exactly in the case of Theorem 9, we end up with

$$\begin{aligned}
&\left(\int_\Omega \eta^{\frac{2Q}{Q-2}} \sum_{s=1}^n \left[|X_s u|^{p+2\alpha} + |Y_s u|^{p+2\alpha}\right]^{\frac{Q}{Q-2}}\, dx\right)^{\frac{Q-2}{Q}} \\
&\leq c(\alpha,\eta)\int_{\operatorname{supp}\eta} \sum_{s=1}^n 1 + \left[|X_s u|^{\frac{p}{p-2}(p-2+2\alpha)} + |Y_s u|^{\frac{p}{p-2}(p-2+2\alpha)}\right]\, dx \\
&\quad + c(\alpha)\int_{\operatorname{supp}\eta} |Tu|^p\, dx. \qquad (127)
\end{aligned}$$

We would like to iterate (127) as we did for (117). Unfortunately this is not possible, or, more precisely, the previous inequality would yield an improvement in the integrability of $\mathfrak{X}u$ up to L^s, for every $s < \infty$, provided

$$\frac{(p+2\alpha)Q}{Q-2} > \frac{p(p-2+2\alpha)}{p-2} \qquad \text{for } \alpha \geq 0, \qquad (128)$$

which is satisfied if and only if $p > Q$, a case excluded by (51). On the other hand (128) turns out to be true for certain, not arbitrarily large, values of $\alpha \geq 0$, as for instance $\alpha = 0$. We use this observation to implement a *non-divergent Moser-type iteration* that will eventually lead to establish the higher integrability of $\mathfrak{X}u$ in the range described in (122); in other words, we shall use exponents α as long as (128) is satisfied. The argument goes as follows: fix a ball $B(x_0, r) \subset\subset \Omega$ and we use the same sequences of cut-off functions $\{\eta_k\}$, and radii $\{\rho_k\}$, already emploied in Theorems 8 and 9. We define the sequence $\{\alpha_k\}_k$ according to

$$\begin{cases} \alpha_{k+1} = \bar{\chi}\alpha_k + \frac{p-2}{Q-2} \\ \alpha_0 \;=\; 0. \end{cases} \qquad (129)$$

This time we set

$$\bar{\chi} = \frac{Q}{Q-2}\frac{p-2}{p} < 1. \tag{130}$$

Notice that the previous inequality is obvious for $n = 1, 2$, and it is a consequence of (51) for $n \geq 3$, which implies in particular $p < Q$. Also, the fact that $\bar{\chi} < 1$ makes $\{\alpha_k\}_k$ a non-divergent sequence. Indeed, we have

$$\alpha_k = \frac{p-2}{Q-2}\sum_{j=0}^{k-1}\tilde{\chi}^j \nearrow \frac{p}{2}\frac{p-2}{Q-p}, \tag{131}$$

so that

$$\frac{(p+2\alpha_k)Q}{Q-2} \nearrow \frac{Qp}{Q-p}. \tag{132}$$

Clearly, giving α the value of the limit quantity (131), we obtain equality in (128), instead of strict inequality. Notice also that the choice in (129) implies

$$\frac{(p+2\alpha_k)Q}{Q-2} = \frac{p(p-2+2\alpha_{k+1})}{p-2} \qquad \text{for } \ k \in \mathbb{N}. \tag{133}$$

Taking advantage of (133), and iteratively applying (127) with the choice $\eta = \eta_k$, $\alpha = \alpha_k$, we obtain

$$\int_{B_{\rho_k}} \sum_{s=1}^{n} \left[|X_s u|^{p+2\alpha_k} + |Y_s u|^{p+2\alpha_k}\right]^{\frac{Q}{Q-2}} dx \leq C_k \qquad \text{for } \ k \in \mathbb{N}.$$

Here the constant C_k is finite and depends on k via the occurrence of the cut-off functions $\eta_1, \eta_2, \ldots, \eta_k$. It also depends on n, p, L and the norms $\|\mathfrak{X}u\|_{L^p(B_r)}$ and $\|Tu\|_{L^p(B_r)}$. As for the former norm, observe that it appears at the first step of the iteration, taking $\alpha = 0$, and it bounds, modulo increasing constants, all the higher power norms of $\mathfrak{X}u$; as for the latter, observe that $\|Tu\|_{L^p(B_r)}$ appears on the right hand side of (127), and therefore appears at each step of the iteration. Using at this point (132) we easily obtain that for every $s < Qp/(Q-p)$ there exists $\bar{k} \in \mathbb{N}$, depending on s but independent of the initial ball considered B_r, such that $\|\mathfrak{X}u\|_{L^s(B_{\rho_{\bar{k}}})} \leq C_{\bar{k}}$. From this fact, the local L^s integrability of $\mathfrak{X}u$ as claimed in (122) follows via a standard covering argument and the proof is finished.

7. Regularity of solutions

In this section we prove the main result, Theorem 1, and then also Theorem 2. First we prove that both $\mathfrak{X}u$ and Tu are locally bounded in Ω and obtain various a priori estimates.

Theorem 11. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(6). *Then $Du \in L^{\infty}_{\mathrm{loc}}(\Omega)$.*

Proof. We ask the reader to keep in mind the scheme in (11), and to be patient enough to keep track of the various bounds assumed on p through Sections 4-6. We shall first concentrate on the case $p \neq 2$. We start obtaining the higher integrability result for the horizontal gradient of Theorem 6. In turn this implies the higher integrability result of Theorem 10, which is an improvement of the previous one provided p is large enough (compare with Remark 9). In any case we conclude with $\mathfrak{X}u \in L^{\frac{Qp}{Q-p}-\varepsilon}_{\mathrm{loc}}(\Omega)$, for every $\varepsilon > 0$. Observe that the applications of these two theorems is possible since the bound (6) implies the one in (51), as noted in Remark 7. Again we observe that the assumption (6) implies

$$\frac{t}{t-p+2} < \frac{Q}{Q-2} \qquad \text{where} \qquad t = \frac{Qp}{Q-p}. \tag{134}$$

This is exactly the point where the quantity in (7) comes from. Consider also Remark 7 for the case $n = 3$, where the case $p = 10/3$ can be achieved. Therefore, by (134), using a continuity argument and Theorem 10, we can select $s < t$ such that $\mathfrak{X}u \in L^{s}_{\mathrm{loc}}(\Omega)$, with s close enough to t in order to have also (40) satisfied. At this point we apply Theorem 8 to obtain that Tu is (essentially) locally bounded in Ω. Then we apply Theorem 9 to deduce that also $\mathfrak{X}u$ is locally bounded in Ω and the proof is complete in the relevant case $p \neq 2$. The case $p = 2$ is already treated in the paper [2], but follows from our results as well. Indeed, following for instance Remark 4, we apply directly Theorem 8 with the choice $s = p = 2$, which obviously satisfies (40), and we get the local boundedness of Tu. This avoids the use of Theorem 10, which is the only point of the paper where we need to assume that $p \neq 2$. The rest of the proof follows as for the case $p \neq 2$.

Remark 10. The a priori estimate in (41) can be at this point improved since $\mathfrak{X}u \in L^{\infty}_{\mathrm{loc}}(\Omega)$, and we can obtain a “dual” version of the a priori estimate (52). Indeed, thanks to the fact that the constant c appearing in (41) is independent of s, we may let $s \to \infty$ in (41), and noticing that $\chi/(\chi-1) \to Q/2$, we obtain

$$\|Tu\|_{L^{\infty}(B_\rho)} \leq \left(\frac{c}{r-\rho}\right)^{\frac{Q}{2}} \left(\frac{\|\mu + |\mathfrak{X}u|\|_{L^{\infty}(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}} \|Tu\|_{L^{2}(B_r)} \tag{135}$$

where once again the constant c only depends on $n, L/\nu$ and p, being independent of the particular solution u, the constant μ, and the vector field (a_i). Taking $\rho = R/2$ and using Hölder's inequality we obtain

$$\sup_{B_{\frac{r}{2}}} |Tu| \le c \left(\frac{\|\mu + |\mathfrak{X}u|\|_{L^\infty(B_r)}}{\mu} \right)^{\frac{(p-2)Q}{4}} \left(\fint_{B_r} (1 + |Tu|^p)\, dx \right)^{\frac{1}{p}}. \quad (136)$$

Estimates (53) and (136) are dual each other, and reflect how, in our approach, the regularity for the horizontal gradient is controlled by the regularity of the vertical one and vice-versa. Keep in mind the diagram (11).

Once the boundedness of both $\mathfrak{X}u$ and Tu has been established, we can proceed to prove the local Hölder continuity of the full Euclidean gradient Du, by adapting the techniques from [2],[3],[4], that in turn extend to the setting of the Heisenberg group the classical work of Morrey [27],[28]. More precisely, while in the case $p = 2$ [2],[3],[4] the focal point for regularity is the existence of the second horizontal derivatives of solutions, here is their Lipschitz continuity (see Remark 11 below). Once this is done, the higher regularity is obtained by passing to the differentiated equation.

We prove the following result, from which Theorem 1 immediately follows.

Theorem 12. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(6). *Then for every open subset $A \subset\subset \Omega$ there exists $\alpha = \alpha(A) \in (0,1)$ such that $\mathfrak{X}u, Tu \in \Gamma^\alpha(A)$.*

Proof. Since the result of the theorem is of a local nature, in the following we shall assume without loss of generality that A is a smooth open subset, compactly contained in Ω. To begin with the proof we start proving that

$$Tu \in \Gamma^\alpha(A) \qquad\qquad \mathfrak{X}Tu \in M^{2,\alpha}(A) \quad (137)$$

for a certain $\alpha > 0$, depending on the open subset A, n, p, and L/ν; the definition of the spaces $\Gamma^\alpha(A)$, and $M^{2,\alpha}(A)$, can be found in Section 2, (17) and (18), respectively. Taking $T\varphi$ instead of φ, in (27), and integrating by parts, we have

$$\int_\Omega \sum_{i,j=1}^{2n} D_{z_j} a_i(\mathfrak{X}u) X_j Tu X_i \varphi \, dx = 0. \quad (138)$$

Here φ is of course taken with compact support in Ω. The previous integration by parts and differentiation are justified in view of the results of Sections 2 and 3. Equation (138) implies that setting

$$b_{i,j}(x) := D_{z_j} a_i(\mathfrak{X}u(x)), \quad (139)$$

we have that $w := Tu$ solves the following linear sub-elliptic equation in (horizontal) divergence form:

$$\sum_{i,j=1}^{2n} X_i^* \left(b_{i,j}(x) X_j w\right) = 0, \qquad \text{weakly in } A. \tag{140}$$

The coefficients are measurable, bounded and elliptic:

$$\mu^{p-2}|\lambda|^2 \leq \sum_{i,j=1}^{2n} b_{i,j}(x)\lambda_i\lambda_j \leq M(A)|\lambda|^2, \qquad \text{for } x \in A \quad \text{and } \lambda \in \mathbb{R}^{2n}, \tag{141}$$

where we have set

$$M(A) = L(\mu^2 + \|\mathfrak{X}u\|^2_{L^\infty(A)})^{\frac{p-2}{2}}. \tag{142}$$

Notice that $w \in HW^{1,2}(A)$ by virtue of Theorem 7, therefore w is a real energy solution to (138) in the same fashion of Definition 2. Therefore, from the Hölder regularity result of Theorem 4 we deduce the existence of

$$\alpha = \alpha\left(\frac{L}{\nu} \cdot \frac{M(A)}{\mu^{p-2}}\right) > 0 \tag{143}$$

such that $w \equiv Tu \in \Gamma^\alpha_{\text{loc}}(A)$. Since $A \subset\subset \Omega$ is arbitrary, by passing to a slight larger smooth open subset we conclude also the first inclusion in (137).

Remark 11. At this point we can see the real difference between the case $p = 2$, treated in [2],[3], and the one $p \neq 2$. Indeed when $p = 2$, referring to (142), we just have that $M = L$, and we do not need to use the boundedness of $\mathfrak{X}u$ to have (141). In the case $p \neq 2$ last information can be obtained only when $\mathfrak{X}u$ is locally bounded, which has to be independently proven.

In a standard way we can now obtain the usual Caccioppoli type inequality for elliptic equations by testing (140) with $\eta^2(w - (w)_{x_0,r})$, where $B(x_0, r) \subset\subset A$ and η is a suitable, smooth cut-off function between $B(x_0, r/2)$ and $B(x_0, r)$. After a few computations (see for instance [16], Chapter 6) we come up with

$$\begin{aligned} \fint_{B(x_0,\frac{r}{2})} |\mathfrak{X}Tu|^2\, dx &\leq \frac{c}{r^2} \fint_{B(x_0,r)} |Tu - (Tu)_{x_0,r}|^2\, dx \\ &\leq cr^{2(\alpha-1)}. \end{aligned}$$

The last estimate is obtained using the fact that $Tu \in \Gamma^\alpha(A)$, while the constant c depends on the ellipticity ratio $(L/\nu) \cdot M(A)/\mu^{p-2}$ of the equation (140) appearing in (143). This immediately implies that $\mathfrak{X}Tu \in M^{2,\alpha}_{\mathrm{loc}}(A)$. Actually, since $A \subset\subset \Omega$ is arbitrary, passing to a slightly larger subset, also this time we obtain that $\mathfrak{X}Tu \in M^{2,\alpha}(A)$, and (137) is completely proved. We are now ready to prove that also $\mathfrak{X}u$ is Hölder continuous in A, eventually with a different exponent $\beta > 0$, depending on α, and therefore ultimately on the quantity $(L/\nu)\cdot M(A)/\mu^{p-2}$ already in (143). We shall achieve this by proving that $X_s u$ is locally Hölder continuous in A for every $s \in \{1,\ldots,n\}$ - the same proof will work for $Y_s u$. Taking $X_s\varphi$, instead of φ in (27), commuting and integrating by parts, we obtain in a standard way that

$$\begin{aligned}&\int_\Omega \sum_{i,j=1}^{2n} D_{z_j} a_i(\mathfrak{X}u) X_j X_s u X_i \varphi \, dx \\ &= -\int_\Omega \sum_{i=1}^{2n} D_{z_{s+n}} a_i(\mathfrak{X}u) Tu X_i\varphi\, dx - \int_\Omega a_{s+n}(\mathfrak{X}u) T\varphi\, dx.\end{aligned}$$

Again integrating by parts the last integral we finally arrive at

$$\begin{aligned}&\int_\Omega \sum_{i,j=1}^{2n} D_{z_j} a_i(\mathfrak{X}u) X_j X_s u X_i \varphi \, dx \\ &= -\int_\Omega \sum_{i=1}^{2n} D_{z_{s+n}} a_i(\mathfrak{X}u) Tu X_i\varphi\, dx \\ &\qquad + \int_\Omega \sum_{j=1}^{2n} D_{z_j} a_{s+n}(\mathfrak{X}u) X_j Tu \varphi\, dx. \qquad (144)\end{aligned}$$

We remind the reader that the legality of the foregoing manipulations in ensured by the differentiability results of Section 2 and 3. Fix a smooth, open subset $A \subset\subset \Omega$ and this time set $w = X_s u$. We see that $w \in HW^{1,2}(A)$ by (26), and that w solves the following linear sub-elliptic equation with bounded and elliptic coefficients, and with lower order terms:

$$\sum_{i,j=1}^{2n} X_i^* \left(b_{i,j}(x) X_j w + c_i(x)\right) = d(x), \qquad \text{weakly in } A, \qquad (145)$$

where the matrix $\{b_{i,j}(x)\}$ as been already defined in (139), and we have set

$$\begin{cases} c_i(x) = D_{z_{s+n}} a_i(\mathfrak{X}u(x)) Tu \\ d(x) = -\sum_{i=1}^{2n} D_{z_j} a_{s+n}(\mathfrak{X}u(x)) X_j Tu(x). \end{cases} \qquad (146)$$

Theorem 11, and the first inclusion in (137), in particular the boundedness of Tu, immediately allows us to conclude that $c_i \in M^{2,\alpha}(A)$, since

$$\fint_{B(x,r)\cap A} |c_i|^2 \, dx \le c(M) \le cr^{2(\alpha-1)},$$

whenever $x \in A$, and r is suitably small. The second inclusion in (137), and again Theorem 11, implies that

$$\fint_{B(x,r)\cap A} |d| \, dx \le \left(\fint_{B(x,r)\cap A} |d|^2 \, dx \right)^{\frac{1}{2}} \le cr^{(\alpha-1)} \le cr^{(\alpha-2)}, \qquad (147)$$

whenever $x \in A$, and r is suitably small. At this point, by (146) and (147), we conclude by applying Theorem 6.32 from [3], or Theorem 4.6 from [2], that $X_s u \in M^{2,\beta}(A)$ for some $\beta > 0$ depending on α. This works for all $s \in \{1, \dots, n\}$, and also the Y_s: we get that $|\mathfrak{X} w| \in M^{2,\beta}(A)$, and therefore, using the inclusion (19), we obtain that $w \equiv X_s u \in \Gamma^{\beta}_{\mathrm{loc}}(A)$. Actually, as already noted in Remark 6.30 from [3], a careful re-adaptation of original Morrey's proofs lead to establish that in fact $\beta = \alpha$. The proof of Theorem 12 is concluded. The proof of Theorem 1 also follows.

Proof (of Theorem 2, and more). The higher regularity of solutions, according to the higher regularity of the vector field (a_i), is now a consequence of the sub-elliptic Schauder estimates of Xu [32]. We refer to [2] for a relevant discussion, in particular Section 4 of [2], from which we could also deduce Theorem 2.

8. Further a priori estimates

In this final section we are going to report a few additional results, which are not necessary to prove the regularity of solutions of (2), but that show how, assuming an additional restriction on the distance between p and 2, it is possible to prove some further a priori estimates allowing to locally bound the full Euclidean gradient norm $||Du||_{L^\infty}$, in terms of the horizontal norm $||u||_{HW^{1,p}}$, which is the datum given by the problem. In the whole section we shall argue under the additional bound

$$2 \le p < 2 + \frac{2}{(n+1)^2}, \qquad (148)$$

which is more restrictive than the one considered in (7), as a direct computation reveals.

Theorem 13. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to the equation* (2) *under the assumptions* (3)-(5), *and* (148). *If $B_r \subset\subset \Omega$ then*

$$\sup_{B_{\frac{r}{2}}} |Tu| \le \frac{c}{r^q}\Bigg[\left(\frac{\|1+|\mathfrak{X}u|\|_{L^p(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}} \cdot \cdot\left(\frac{\|1+|\mathfrak{X}u|\|_{L^p(B_r)}}{r^{\frac{2+Q}{2}}}+\frac{\|u\|_{L^p(B_r)}}{r^{\frac{4+Q}{2}}}\right)\Bigg]^{\frac{8}{8-Q^2(p-2)}} \tag{149}$$

and

$$\sup_{B_{\frac{r}{2}}} |\mathfrak{X}u| \le \frac{c}{r^{\frac{Q}{p}+\frac{Qq}{2}}}\|1+|\mathfrak{X}u|\|_{L^p(B_r)}\Bigg[\left(\frac{\|1+|\mathfrak{X}u|\|_{L^p(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}} \cdot \cdot\left(\frac{\|1+|\mathfrak{X}u|\|_{L^p(B_r)}}{r^{\frac{2+Q}{2}}}+\frac{\|u\|_{L^p(B_r)}}{r^{\frac{4+Q}{2}}}\right)\Bigg]^{\frac{8}{8-Q^2(p-2)}\frac{Q}{2}} \tag{150}$$

Here the constant c depends only on n, p, and L/ν, and

$$q = \left(\frac{Q}{2}+\frac{Q^2(p-2)}{4p}\right)\cdot\frac{8}{8-Q^2(p-2)} \ge 1. \tag{151}$$

Proof. Let us take $B_r \subset\subset \Omega$ as in the statement, and let us fix $r/2 < t < s < r$. We start by estimate (135). Taking $r \equiv (s+t)/2$ and $\varrho \equiv t$, and using Hölder's inequality and the fact that $p \ge 2$, we gain

$$\|Tu\|_{L^\infty(B_t)} \le \left(\frac{c}{s-t}\right)^{\frac{Q}{2}}\left(\frac{\|\mu+|\mathfrak{X}u|\|_{L^\infty(B_{(s+t)/2})}}{\mu}\right)^{\frac{(p-2)Q}{4}}\|Tu\|_{L^p(B_{(s+t)/2})}.$$

Now we use estimate (52) with the choice $r \equiv s$ and $\varrho \equiv (s+t)/2$, and obtain

$$\|\mathfrak{X}u\|_{L^\infty(B_{(s+t)/2})} \le \left(\frac{c}{s-t}\right)^{\frac{Q}{p}}\|1+|Tu|\|^{\frac{Q}{2}}_{L^\infty(B_s)}\|1+|\mathfrak{X}u|\|_{L^p(B_s)},$$

where the constant c only depends on n, p, and L/ν. Merging the previous two inequalities, and performing routine estimations, we have

$$\|1+|Tu|\|_{L^\infty(B_t)} \le \left(\frac{c}{s-t}\right)^{\frac{Q}{2}+\frac{Q^2(p-2)}{4p}}\|1+|Tu|\|^{\frac{Q^2(p-2)}{8}}_{L^\infty(B_s)}\cdot \cdot\left(\frac{\|1+|\mathfrak{X}u|\|_{L^p(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}}\|Tu\|_{L^p(B_r)}$$

$$+\left(\frac{c}{s-t}\right)^{\frac{Q}{2}}\|1+|Tu|\|_{L^p(B_r)}.$$

Now the point is that (148) allows us to conclude that $Q^2(p-2)/8<1$, and therefore we can use Young's inequality to have

$$\begin{aligned}\|1+|Tu|\|_{L^\infty(B_t)} \le \frac{\|1+|Tu|\|_{L^\infty(B_s)}}{2} \\ +\frac{c}{(s-t)^q}\left(\frac{\|\mu+|\mathfrak{X}u|\|_{L^p(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}\frac{8}{8-Q^2(p-2)}}\cdot \\ \cdot\|1+|Tu|\|_{L^p(B_r)}^{\frac{8}{8-Q^2(p-2)}},\end{aligned}$$

where q is as in (151). Now we let $I(s):=\|1+|Tu|\|_{L^\infty(B_s)}$ for $s\in(r/2,r)$, and apply Lemma 4 with $\theta=1/2$, $q\equiv q$, and the obvious choice of the constants A,B, to get

$$\begin{aligned}\|1+|Tu|\|_{L^\infty(B_{\frac{r}{2}})} \le \frac{c}{r^q}\left(\frac{\|\mu+|\mathfrak{X}u|\|_{L^p(B_r)}}{\mu}\right)^{\frac{(p-2)Q}{4}\frac{8}{8-Q^2(p-2)}}\cdot \\ \cdot\|1+|Tu|\|_{L^p(B_r)}^{\frac{8}{8-Q^2(p-2)}}.\end{aligned}$$

Note again that, with no surprise, when $p=2$, then $q=Q/2$ and we get back (43). Now observe that in Theorem 5, a simple covering argument allows to take $\alpha=1/2$, modulo enlarging the constants involved in the estimates. Therefore, merging estimate (23) with the previous one, and changing repeatedly the values of r ($3r/4$ instead of r and so on), we easily get (149). In order to get (150), it suffices to use (149) in (52), for a suitable choice of r and ϱ.

Acknowledgements. J.J.M. has been partially supported by NSF award DMS-0500983. G.M. has been partially supported by MIUR via the national project "Calcolo delle Variazioni", and by GNAMPA via the project "Studio delle singolarità in problemi geometrici e variazionali". Part of the work has been done while G.M. was visiting the Department of Mathematics of the University of Pittsburgh, and the Center for Nonlinear Analysis of the Carnegie Mellon University, in May 2005. He thanks both the institutions for the friendly support.

References

1. A. Bellaïche, The tangent space in sub-Riemannian geometry in *Sub-Riemannian geometry*, 1–78, Progr. Math., 144, Birkhäuser, Basel, 1996.
2. L. Capogna, Regularity of quasi-linear equations in the Heisenberg group. Comm. Pure Appl. Math. **50**, 867- 889 (1997)

3. L. Capogna, Regularity for quasilinear equation and 1-quasiconformal maps in Carnot groups. Mathematische Annalen **313**, 263-295 (1999)
4. L. Capogna, Doctoral dissertation. Purdue University, 1996.
5. L. Capogna & D. Danielli & N. Garofalo, An embedding theorem and the Harnack inequality for nonlinear subelliptic equations. Comm. P.D.E **18**, 1765-1794 (1993)
6. L. Capogna & N. Garofalo, Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type. J. Eur. Math. Soc. (JEMS) **5**, 1-40 (2003)
7. W.L Chow, Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. Mathematische Annalen **117**, 98-105 (1939)
8. A. Domokos, Differentiability of solutions for the non-degenerate p-Laplacian in the Heisenberg group. J. Differential Equations **204**, 439-470 (2004)
9. A. Domokos & J. J. Manfredi, Subelliptic Cordes estimates. Proc. Amer. Math. Soc. **133**, 1047-1056 (2005)
10. A. Domokos & J. J. Manfredi, $C^{1,\alpha}$-regularity for p-harmonic functions in the Heisenberg group for p near 2. In "The p-Harmonic Equation and Recent Advances in Analysis", ed. P. Poggi-Corradini, Contemporary Mathematics **370**, American Mathematical Society, 2005, pp. 17–23
11. J. J. Duistermaat & J. A. Kolk, Lie groups, Springer-Verlag Berlin Heidelberg, 2000
12. A. Föglein, Partial regularity results for systems in the Heisenberg group, Preprint (2006).
13. G. B. Folland, Applications of analysis on nilpotent groups to partial differential equations, Bulletin of AMS **83**, 912-930 (1977)
14. G. B. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups Ark. Mat. **13**, 161-207 (1975).
15. D. Gilbarg & N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Springer-Verlag, 1983.
16. E. Giusti, Direct methods in the Calculus of Variations, World Scientific, 2003.
17. C. Hamburger, Regularity of differential forms minimizing degenerate elliptic functionals J. reine angew. Math. (Crelles J.) **431**, 7-64 (1992)
18. L. Hörmander, Hypoelliptic second order differential equations. Acta Mathematica **119**, 147-171 (1967).
19. D. Jerison, The Poincaré inequality for vector fields satisfying Hörmander's condition. Duke Math. J. **53**, 503-523 (1986)
20. A. W. Knapp & E. M. Stein, Intertwining operators for semi-simple groups. Ann. of Math. **93**, 489-578 (1971)
21. J. J. Kohn, Pseudo-differential operators and hypoellipticity. *Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971),* pp. 61–69. Amer. Math. Soc., Providence, R.I., 1973.
22. A. Korányi & S. Vági, Singular integrals on homogeneous spaces and some problems of classical analysis. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **25**, 575-648 (1971)
23. G. Lu, Embedding theorems into Lipschitz and BMO spaces and applications to quasilinear subelliptic differential equations. Publ. Mat., 301-329 **40** (1996)
24. G. Lu, Embedding theorems on Campanato-Morrey spaces for vector fields and applications. C. R. Acad. Sci. Paris Sér. I Math. **320**, 429-434 (1995)
25. J. J. Manfredi, Regularity for minima of functionals with p-growth. J. Differential Equations **76**, 203-212 (1988)

26. S. Marchi, $C^{1,\alpha}$ local regularity for the solutions of the p-Laplacian on the Heisenberg group for $2 \leq p \leq 1+\sqrt{5}$. Z. Anal. Anwendungen **20**, 617-636 (2001). Erratum: Z. Anal. Anwendungen **22**, 471-472 (2003)
27. C. B. Morrey Jr., Multiple integrals in the calculus of variations. Die Grundlehren der mathematischen Wissenschaften, 130 Springer-Verlag, New York 1966
28. C. B. Morrey Jr., Second order elliptic equations in several variables and Hölder continuity, Math. Z. **72**, 146-164 (1959/1960).
29. A. Nagel & E.M. Stein & S. Wainger, Balls and metrics defined by vector fields. I. Basic properties. Acta Mathematica **155**, 103-147 (1985)
30. J. Peetre, New Thoughts on Besov Spaces, Duke Univ. Math. Series, Durham Univ. 1976
31. C. J. Xu, Subelliptic variational problems. Bull. Soc. Math. France **118**, 147–169 (1990)
32. C. J. Xu, Regularity for quasilinear second-order subelliptic equations. Comm. Pure Appl. Math. **45**, 77–96 (1992)