Available online at www.sciencedirect.com

ScienceDirect

Advances in Mathematics 222 (2009) 62–129

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

Gradient regularity for elliptic equations in the Heisenberg group

Giuseppe Mingione [a,*], Anna Zatorska-Goldstein [b], Xiao Zhong [c]

[a] *Dipartimento di Matematica, Università di Parma, Viale G.P. Usberti 53/A, I-43100 Parma, Italy*
[b] *Instytut Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski, Banacha 2, 02-097 Warszawa, Poland*
[c] *Department of Mathematics and Statistics, PO Box 35 (MaD) FI-40014, University of Jyväskylä, Finland*

Received 18 September 2007; accepted 4 March 2009

Available online 16 April 2009

Communicated by Luis Caffarelli

Abstract

We give dimension-free regularity conditions for a class of possibly degenerate sub-elliptic equations in the Heisenberg group exhibiting super-quadratic growth in the horizontal gradient; this solves an issue raised in [J.J. Manfredi, G. Mingione, Regularity results for quasilinear elliptic equations in the Heisenberg group, Math. Ann. 339 (2007) 485–544], where only dimension dependent bounds for the growth exponent are given. We also obtain explicit a priori local regularity estimates, and cover the case of the horizontal p-Laplacean operator, extending some regularity proven in [A. Domokos, J.J. Manfredi, $C^{1,\alpha}$-regularity for p-harmonic functions in the Heisenberg group for p near 2, in: Contemp. Math., vol. 370, 2005, pp. 17–23]. In turn, using some recent techniques of Caffarelli and Peral [L. Caffarelli, I. Peral, On $W^{1,p}$ estimates for elliptic equations in divergence form, Comm. Pure Appl. Math. 51 (1998) 1–21], the a priori estimates found are shown to imply the suitable local Calderón–Zygmund theory for the related class of non-homogeneous, possibly degenerate equations involving discontinuous coefficients. These last results extend to the sub-elliptic setting a few classical non-linear Euclidean results [T. Iwaniec, Projections onto gradient fields and L^p-estimates for degenerated elliptic operators, Studia Math. 75 (1983) 293–312; E. DiBenedetto, J.J. Manfredi, On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems, Amer. J. Math. 115 (1993) 1107–1134], and to the non-linear case estimates of the same nature that were available in the sub-elliptic setting only for solutions to linear equations.

* Corresponding author.
E-mail addresses: giuseppe.mingione@unipr.it (G. Mingione), azator@mimuw.edu.pl (A. Zatorska-Goldstein), zhong@maths.jyu.fi (X. Zhong).
URLs: http://www.unipr.it/~mingiu36 (G. Mingione), http://www.mimuw.edu.pl/~azator/ (A. Zatorska-Goldstein), http://www.math.jyu.fi/~zhong/ (X. Zhong).

0001-8708/$ – see front matter
doi:10.1016/j.aim.2009.03.016

MSC: primary 35H20, 35J70

Keywords: Heisenberg group; p-Laplacean; Weak solutions; Regularity

Contents

1. Introduction

The regularity in question concerns sub-elliptic equations of the type

$$\operatorname{div}_H a(\mathfrak{X}u) = \sum_{i=1}^{2n} X_i a_i(\mathfrak{X}u) = 0, \tag{1.1}$$

which are defined in a bounded, open sub-domain Ω of the Heisenberg group $\mathbb{H}^n$, $n \geqslant 1$. The vector field $a = (a_i)\colon \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ is assumed to be of class C^1 and satisfying the following growth and ellipticity conditions:

$$\big|Da(z)\big|\big(\mu^2 + |z|^2\big)^{\frac{1}{2}} + \big|a(z)\big| \leqslant L\big(\mu^2 + |z|^2\big)^{\frac{p-1}{2}}, \tag{1.2}$$

and

$$\nu\big(\mu^2 + |z|^2\big)^{\frac{p-2}{2}} |\lambda|^2 \leqslant \sum_{i,j=1}^{2n} D_{z_j} a_i(z)\lambda_i\lambda_j, \tag{1.3}$$

for every $z, \lambda \in \mathbb{R}^{2n}$, where

$$0 < \nu \leqslant 1 \leqslant L, \qquad \mu \in [0, 1], \qquad p \geqslant 2.$$

At certain stages we shall assume the (sub-elliptic) non-degeneracy condition

$$\mu > 0. \tag{1.4}$$

Assumptions (1.2)–(1.3) are standard when considering quasilinear equations, and their consideration traces back to the classical Euclidean work of Ladyzhenskaya and Ural'tseva [37]. Such assumptions are clearly tailored on the basic model equation

$$\operatorname{div}_H\left(\left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} \mathfrak{X}u\right) = 0, \tag{1.5}$$

whose left-hand side operator reduces to the Kohn–Laplacean for $p = 2$, while taking $\mu = 0$ we have the also familiar horizontal p-Laplacean operator on the left-hand side:

$$\operatorname{div}_H\left(|\mathfrak{X}u|^{p-2}\mathfrak{X}u\right) = 0. \tag{1.6}$$

In order to preliminarily fix some notation, let us recall that we are denoting points $x \in \mathbb{H}^n \equiv \mathbb{R}^{2n+1}$ by mean of the usual exponential coordinates

$$x = (x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{2n}, t), \tag{1.7}$$

while throughout the paper we are denoting

$$X_i \equiv X_i(x) = \partial_{x_i} - \frac{x_{n+i}}{2}\partial_t, \qquad X_{n+i} \equiv X_{n+i}(x) = \partial_{x_{n+i}} + \frac{x_i}{2}\partial_t, \tag{1.8}$$

and

$$T \equiv T(x) = \partial_t, \qquad \mathfrak{X}u = (X_1 u, X_2 u, \dots, X_{2n} u). \tag{1.9}$$

The functional ambient of the problem (1.1) is the sub-elliptic Sobolev space $HW^{1,p}(\Omega)$ (see Section 2.4 below), that is, solutions u are assumed to belong to $L^p(\Omega)$ and to satisfy

$$\mathfrak{X}u \in L^p\left(\Omega, \mathbb{R}^{2n}\right), \tag{1.10}$$

while nothing is assumed about Tu. We recall that if $F \equiv (F_i): \Omega \to \mathbb{R}^{2n}$ is an L^1 vector field in the following we shall denote the horizontal divergence operator by

$$\operatorname{div}_H F \equiv \sum_{i=1}^{2n} X_i F_i,$$

which is obviously defined in the distributional sense. We refer to Section 2 for more on the Heisenberg groups $\mathbb{H}^n$, $n = 1, 2, 3, \dots$, and for the related notation adopted in this paper.

1.1. Gradient regularity

The study of regularity properties of weak solutions to (1.1) started with the classical paper of Hörmander [29], which dealt with general vector fields and linear equations, and was later followed by other remarkable contributions devoted to the linear case, as for instance [22,23,35]. Capogna was the first to obtain Hölder continuity theorems for the gradient of solutions to quasi-linear sub-elliptic equations in divergence form: initially in the Heisenberg group [8], and then in more general Carnot groups [9]; see also his thesis [7]. The operators considered in [8,9] have quadratic growth, that is, they satisfy (1.2)–(1.3) for $p = 2$, so that equations as those in (1.5)–(1.6) are not covered by his theory unless a priori regularity assumptions are made on the gradient. The case $p > 2$ is another story; indeed while Hölder continuity of u has been obtained in [10,38], when considering the gradient of solutions only partial regularity results are available, that is, the regularity of the gradient outside a closed, negligible subset of the domain Ω; this fact has first been established by Capogna and Garofalo in [11]; another proof is given by Föglein [21]. When turning to everywhere continuity of Du, the regularity results obtained prescribe that the exponent p should not be "too far from 2," roughly meaning that the non-linearity of (1.1) is in some sense not too strong. In this respect, Domokos [16], extending earlier, pioneering results of Marchi [42], showed that $Tu \in L^p_{\mathrm{loc}}(\Omega)$ if $p < 4$. Proving that $Tu \in L^p_{\mathrm{loc}}(\Omega)$ is of course the first fundamental step towards the regularization of solutions u to (1.1), since for them the initial regularity information is just (1.10). As for the higher regularity of Du or $\mathfrak{X}u$, a few Hölder regularity results are available in [12,17,18,40]; a common feature of such papers is to prove regularity results for solutions assuming not only that $p < 4$, but also an additional dimensional bound of the type

$$2 \leqslant p < 2 + o_n \tag{1.11}$$

where $o_n > 0$ denotes a rather awkward, and only in principle explicitly computable quantity, such that $o_n \searrow 0$ when $n \nearrow \infty$. An unpleasant feature of an assumption such as (1.11) is that for a fixed p in the range $[2, 4)$ only low dimensional Heisenberg groups can be dealt with. For instance, considering the full range $[2, 4)$, the regularity results available in [40] only apply to $\mathbb{H}^1$ and $\mathbb{H}^2$; we note that the paper [40], where up to now the best bounds of the type (1.11) have been found, only regards the non-degenerate case $\mu > 0$. Indeed, we explicitly remark that only few regularity results are available in the (sub-elliptic) degenerate case $\mu = 0$, and therefore for solutions to (1.6). See [17]; moreover, in the degenerate case the quantity o_n in (1.11) is not explicitly computable.

In this respect, the aim of the present paper is now twofold: first we are giving *the first dimension-free pointwise regularity results for gradients of solutions*, therefore completely avoiding the use of any dimensional assumptions of the type (1.11). Second, and probably more interestingly, up to a certain extent we shall also treat the degenerate case $\mu = 0$, thereby *covering the sub-elliptic p-Laplacean equation* (1.6). For instance, we shall prove the local Lipschitz continuity of solutions with respect to the intrinsic Carnot–Carathèodory metric.

The first result we are presenting regards the non-degenerate case $\mu > 0$.

Theorem 1.1 *(The non-degenerate case). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4) *with* $2 \leqslant p < 4$. *Then the Euclidean gradient Du is locally Hölder continuous in Ω.*

See Section 2.4 below for the definition of the Horizontal Sobolev space $HW^{1,p}$. The previous result solves an issue raised in [40], where the authors were able to obtain the same degree of regularity only under an additional assumption of the type (1.11). As later described in Section 1.3, we shall adopt here different technical tricks from the ones used in [40]; these will allow us to develop more efficient bootstrap procedures.

Theorem 1.1 comes along with explicit a priori estimates:

Theorem 1.2 *(Non-degenerate estimates). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4) *with $2 \leqslant p < 4$. There exists a constant c, depending on n, p and L/ν, but otherwise independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that the following inequalities hold for any CC-ball $B_R \subset \Omega$:*

$$\sup_{B_{R/2}} |\mathfrak{X}u| \leqslant c\left(\fint_{B_R} \bigl(\mu + |\mathfrak{X}u|\bigr)^p\,dx\right)^{1/p}, \tag{1.12}$$

and

$$\sup_{B_{R/2}} R|Tu| \leqslant c\mu^{\frac{Q(2-p)}{4}}\left(\fint_{B_R} \bigl(\mu + |\mathfrak{X}u|\bigr)^p\,dx\right)^{\frac{1}{p}+\frac{Q(p-2)}{4p}}. \tag{1.13}$$

Finally, for every $1 < q < \infty$ there exists a constant $\tilde{c}$ depending only on n, p, L/ν, q such that

$$\left(\fint_{B_{R/2}} |Tu|^q\,dx\right)^{1/q} \leqslant \frac{\tilde{c}}{R}\left(\fint_{B_R} \bigl(\mu + |\mathfrak{X}u|\bigr)^p\,dx\right)^{1/p}. \tag{1.14}$$

For the definition of CC-balls and more notation see Section 2.3 below. See also (2.9) for more notation.

Let us remark that in the specific situation of Eq. (1.5), where the considered vector field is smooth, the previous theorem allows to prove the arbitrary smoothness of solutions via the boot-strap methods in [8].

Next we turn to the degenerate case $\mu = 0$, where the chief model example is (1.6).

Theorem 1.3 *(The degenerate case). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.3) *with $\mu = 0$, where $2 \leqslant p < 4$. Then*

$$\mathfrak{X}u \in L^{\infty}_{\mathrm{loc}}\bigl(\Omega, \mathbb{R}^{2n}\bigr), \quad \text{and} \quad Tu \in L^q_{\mathrm{loc}}(\Omega) \quad \text{for every } q < \infty. \tag{1.15}$$

Moreover there exists a constant c, depending on n, p, L/ν, but otherwise independent of the solution u, and of the vector field $a(\cdot)$, such that the following inequality holds for any CC-ball $B_R \subset \Omega$:

$$\sup_{B_{R/2}} |\mathfrak{X}u| \leqslant c\left(\fint_{B_R} |\mathfrak{X}u|^p\,dx\right)^{1/p}. \tag{1.16}$$

Finally, for every $q < \infty$ there exists a constant $\tilde{c}$ depending only on n, p, L/ν, q such that

$$\left(\fint_{B_{R/2}} |Tu|^q\,dx\right)^{1/q} \leqslant \frac{\tilde{c}}{R}\left(\fint_{B_R} |\mathfrak{X}u|^p\,dx\right)^{1/p}. \tag{1.17}$$

The previous theorem partially extends some regularity results proven in [18], where the authors work under an assumption of the type (1.11), this time o_n being a small, unspecified quantity coming from the application of abstract Cordes type condition methods. In turn, the boundedness of the horizontal gradient naturally yields a priori Lipschitz bounds:

Corollary 1.1 *(CC-Lipschitz regularity). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.3) *with $2 \leqslant p < 4$. Then u is locally Lipschitz continuous in Ω with respect to the CC-metric in $\mathbb{H}^n$. Moreover there exists a constant c, depending only on n, p, L/ν, but otherwise independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that*

$$\left|u(x) - u(y)\right| \leqslant c\left(\fint_{B_R} \left(\mu + |\mathfrak{X}u|\right)^p\,dx\right)^{1/p} d_{\mathrm{CC}}(x, y) \tag{1.18}$$

holds whenever $B_R \subset \Omega$, and $x, y \in B_{R/2}$.

See (2.4) below for the definition of the intrinsic distance $d_{\mathrm{CC}}(\cdot,\cdot)$. Another consequence of Theorem 1.3 and of the standard, Euclidean Sobolev–Morrey embedding theorem, is now the following:

Corollary 1.2 *(Almost Euclidean–Lipschitz regularity). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.3) *with $2 \leqslant p < 4$. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$ for every $\alpha < 1$.*

Needless to say, in the last result the Hölder continuity is referred to the standard Euclidean metric. We finally mention that the previous theorems are stated for $2 \leqslant p < 4$ for completeness, since in the automatically non-degenerate case $p = 2$ they are essentially due to Capogna [8].

1.2. Calderón–Zygmund type estimates

The estimate (1.16) found in Theorem 1.3 opens the way to a non-linear version of estimates of Calderón–Zygmund type in the Heisenberg group, up to now developed only in the case of linear sub-elliptic equations [4,5]. Here we shall deal with non-linear equations. Let us recall that in the Euclidean setting this is a classical result dating back to T. Iwaniec [30] in the scalar case, and later extended to systems of p-Laplacean type in [15] by DiBenedetto and Manfredi; see also [1,6] for different approaches. The equations considered by such authors are modeled by

$$\operatorname{div}\left(|Du|^{p-2}Du\right) = \operatorname{div}\left(|F|^{p-2}F\right), \tag{1.19}$$

in open subsets of $\mathbb{R}^n$, and the result asserts that $F \in L^q_{\mathrm{loc}}$ implies $Du \in L^q_{\mathrm{loc}}$ every $q \geqslant p$ – see also [31] for some results when $q < p$. More recently Calderón–Zygmund type estimates valid for solutions to general non-linear elliptic systems have been proposed by in [36], based on innovative work of Caffarelli and Peral [6]; following the techniques of these last papers, some estimates have been proposed in the Heisenberg group case in [25] for certain non-linear problems with quadratic growth, that is, when $p=2$. An extension for linear equations in CR manifolds has been obtained in [45]. In the following we shall give higher integrability results for problems with possibly super-quadratic growth $p \geqslant 2$. The equations we are considering are the natural horizontal version of (1.19), involving possibly discontinuous coefficients of VMO type; specifically

$$\operatorname{div}_H\big[b(x)a(\mathfrak{X}u)\big] = \operatorname{div}_H\big(|F|^{p-2}F\big), \tag{1.20}$$

with

$$b(\cdot) \in \mathrm{VMO}_{\mathrm{loc}}(\Omega) \quad \text{and} \quad \nu \leqslant b(x) \leqslant L. \tag{1.21}$$

See Section 2.5 for the precise definition of the space $\mathrm{VMO}_{\mathrm{loc}}(\Omega)$. The prototype of (1.20) is clearly the non-homogeneous p-Laplacean equation with VMO-coefficients, that is

$$\operatorname{div}_H\big(b(x)|\mathfrak{X}u|^{p-2}\mathfrak{X}u\big) = \operatorname{div}_H\big(|F|^{p-2}F\big), \tag{1.22}$$

where the coefficient function $b(\cdot)$ satisfies (1.21), and $F \in L^p(\Omega, \mathbb{R}^{2n})$. The main result is the following:

Theorem 1.4 *(of Calderón–Zygmund type). Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.20) *under the assumptions* (1.2)–(1.3) *with $2 \leqslant p < 4$, and* (1.21). *Then*

$$F \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n}) \quad \textit{implies that } \mathfrak{X}u \in L^q_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big),$$

whenever $p < q < \infty$. Moreover there exists a constant c, depending only on n, p, L/ν, q, and the function $b(\cdot)$, such that the following reverse-Hölder type inequality holds for any CC-ball $B_R \Subset \Omega$:

$$\left(\fint_{B_{R/2}} |\mathfrak{X}u|^q\,dx\right)^{1/q} \leqslant c\left(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p\,dx\right)^{1/p} + c\left(\fint_{B_R} |F|^q\,dx\right)^{1/q}. \tag{1.23}$$

For an alternative statement concerning the dependence of the constant in (1.23) see also Remark 10.1 below, while for a more precise dependence on the various constants see Remark 10.2. Let us recall that in the Euclidean case there is a wide literature on Calderón–Zygmund type estimates for linear problems with VMO-coefficients starting from the Euclidean work of Chiarenza, Frasca and Longo [13], dealing with linear problems. A non-linear approach has been proposed in [34]. As for the sub-elliptic setting, the theory is confined to the linear case [4], where the case of general Hörmander vector fields is considered. In this paper we give the first results for non-linear problems with VMO coefficients, allowing also for BMO coefficients with small BMO seminorm, see Remark 11.1 below. Anyway we remark that the integrability results obtained here are new already in the case $b(x) \equiv 1$ – that is, when no coefficients are actually involved.

Moreover, we remark that the result of Theorem 1.4 extends to a family of more general equations with continuous coefficients; the corresponding statements are presented at the end of the paper.

1.3. Technical approach, and novelties

The approach proposed in this paper strongly differs from those proposed in earlier ones. Indeed, a common strategy for attacking the regularity problem in the sub-elliptic setting, going back to Hörmander [29] and then followed in subsequent works [8,9,22,23], is to first obtain separately a certain maximal regularity for the vertical part of the gradient Tu, and then, using such an additional information, obtaining regularity results for the horizontal part $\mathfrak{X}u$. Such an approach is for instance followed also in the non-linear setting in [8,9], where it turns out to be successful since $p = 2$. On the other hand the same method does not seem to yield results when $p \neq 2$ and the equation becomes in a certain sense heavily non-linear. We take a different path, hereby proposing a double-bootstrap method: we shall obtain regularity for Tu using the one obtained for $\mathfrak{X}u$, and vice-versa. More precisely we shall prove that

$$Tu \in L^{q_k} \quad \Longrightarrow \quad \mathfrak{X}u \in L^{p_k} \quad \text{and} \quad Xu \in L^{p_k} \quad \Longrightarrow \quad Tu \in L^{q_{k+1}} \tag{1.24}$$

where $\{p_k\}$ and $\{q_k\}$ are two sequences diverging to infinity; in some sense we repeat Hörmander's original strategy breaking it in a countable number of pieces. As a first consequence we obtain that

$$\mathfrak{X}u, Tu \in L^q \quad \text{for every } q < \infty, \tag{1.25}$$

while we remark that all the foregoing inclusions are meant to be local since no boundary information is a priori given on solutions. The use of such a mixed iteration is a direct consequence of the non-linearity of the problem (1.1), since after a preliminary differentiation of the equation, Tu cannot be realized as a solution of a linear equation with bounded coefficients, and a deeper interaction between the horizontal and the vertical parts of the gradient must be exploited. The implementation of (1.24) requires a rather delicate interaction between: suitable Caccioppoli type estimates – also called energy estimates – for the horizontal and vertical gradients, see Section 5; interpolation inequalities of Gagliardo–Nirenberg type in the Heisenberg group, see Section 4; integration-by-parts methods, see Section 7; a certain kind of non-standard energy estimates of mixed type, see Section 6. In some sense, we shall replace the usual Moser's iteration by a different kind of iteration where, at each step, the gain in the integrability exponent is not achieved by Sobolev inequality, but, rather, by an interpolation estimate of Gagliardo–Nirenberg type via an integration-by-parts procedure. Once the integrability information in (1.25) is gained, a suitable variant of Moser's iteration technique will lead to $\mathfrak{X}u \in L^\infty$, see Section 8. Finally, in the non-degenerate case $\mu > 0$ this will lead to $Tu \in L^\infty$ via the results in [40], and eventually to the local Hölder continuity of the Euclidean gradient, which is a standard implication after the work in [7,8,40].

An important background of our technique is the observation of the natural analogy between sub-elliptic equations of the type (1.1), and the more classical Euclidean non-uniformly elliptic equations, or "equations with non-standard growth conditions," or with "(p,q)-growth conditions," as very often called in the setting of the Calculus of Variations [41]. In fact, our techniques are inspired by those developed for such situations, see for instance [3,19,20], although the implementation in the Heisenberg group requires a completely different technical approach. Problems with non-standard growth indeed involve equations featuring ellipticity properties which appear to be weaker in certain special spatial directions: this immediately reminds of the situation of horizontal quasilinear equations in the Heisenberg group as (1.1), where the vertical derivative Tu does not appear directly in the operator. It rather appears only in an intrinsic way, via the horizontal vector fields $\mathfrak{X}u$ and after commutation, see (2.1) below, and therefore the vertical direction is clearly playing a very special role. Such a lack of "vertical ellipticity" is in fact the basic source of problems in the theory of elliptic equations in the Heisenberg group.

As mentioned above, a key ingredient for the subsequent results are the explicit a priori estimates (1.12) and (1.16). Indeed, these will allow for a suitable application of recent non-linear techniques for obtaining higher integrability estimates for non-homogeneous equations [6,36]. Here, due to the presence of the VMO coefficients, we shall use these in combination with various maximal operators, and higher integrability estimates in the spirit of Gehring's lemma. Observe that, due to the non-linearity of the problems we are considering, the standard approaches based on harmonic analysis tools such as, singular integrals, commutators, and so forth, are not available in the present setting.

Finally, let us summarize the content of the paper. In Section 2 we shall collect preliminaries concerning the sub-elliptic setting, while in Section 3 we shall re-visit and re-state in a suitable way a few known regularity results for elliptic equations in the Heisenberg group. Sections 4–7 are devoted to the implementation of (1.24), in the way described a few lines above. Here we shall else re-visit some arguments from [40], and we shall use the a priori boundedness of the solution already obtained in [10]. In Section 8 we prove L^∞-estimates for the gradient and therefore Theorems 1.1, 1.2. Section 9 is devoted to the degenerate case: we prove Theorem 1.3, by combining Theorem 1.2 with a standard approximation method, and then we obtain Corollaries 1.1–1.2. The proof of Theorem 1.4 is in Section 10, while in Section 11 we give a few possible generalizations of Theorem 1.4.

1.4. On the exponent limitation $2 \leqslant p < 4$

In this paper, the main assumption we are considering is, as widely clarified above, the limitation on the growth exponent $2 \leqslant p < 4$. This bound is unnatural as the regularity of solutions is expected to hold for every $p > 1$, and it is linked to the peculiar techniques used in this paper. We shall now briefly explain the various points where such a limitations shows up, because at several stages we shall not need such a limitation. One of the ingredients used here is the initial vertical integrability $Tu \in L^p$ for $p < 4$; indeed many of the estimates here appear with a right-hand side which contains the term $\|Tu\|_{L^p}$ which therefore we have to take bounded. This is for instance the case of Lemma 5.1, where this limitation does not appear explicitly in the proof, but the final right-hand side in the resulting inequalities is finite – via Lemma 3.1 – when $Tu \in L^p$, which holds for $p < 4$. The same happens in other points as for instance in Lemma 5.2 or in Lemma 4.2; see for instance Theorem 4.1, which is a more abstract version of Lemma 4.2, and where the limitation $p < 4$ does not appear. A crucial point where such limitation appears is the integration by parts Lemma 7.2, which is essential to implement the double bootstrap procedure used in order to get the boundedness of the horizontal gradient; here we especially refer to the integration by parts in (7.12). Finally, we emphasize that the Calderón–Zygmund estimates from Section 10 are completely independent of the bound $p \in [2, 4)$; indeed they are essentially based in the peculiar form of the local estimates (1.12) and (1.16) the theorems from Sections 10–11 work with no further constraint on p.

2. Notation, preliminaries

2.1. Notations, conventions

In this paper we shall adopt the usual, but somehow arguable convention to denote by c a general constant, that may vary from line to line; peculiar dependence on parameters will be properly emphasized in parentheses when needed. More precisely we shall usually denote $c \equiv c(\alpha, \beta, \gamma, \ldots)$, meaning that c is actually an increasing (or decreasing) function of $\alpha, \beta, \gamma, \ldots$; in general $c \nearrow \infty$ when either one of the parameters goes to infinity or to zero. For this reason, when dealing with a constant potentially depending on several parameters, in the case when one of the parameters remains bounded, the constant is in fact independent on the parameter in question. Specific occurrences will be clarified by the context. Moreover, special occurrences will be denoted by c_*, c_1, c_2 or the like. In this paper all the constant named by c_*, c_1, c_2 and so on will be assumed without loss of generality to be larger than 1. The scalar product between elements z_1, z_2 of $\mathbb{R}^{2n}$ will be denoted by $\langle z_1, z_2\rangle$; very often, when no ambiguities will arise, we shall simply denote $\langle z_1, z_2\rangle \equiv z_1 z_2$. Finally $\{e_1, \ldots, e_{2n+1}\}$ denotes the standard basis of $\mathbb{R}^{2n+1}$.

In the following, several of the integral estimates for solutions to (1.1) will involve constants depending on the ellipticity/growth parameters μ and L, displayed in (1.2)–(1.3). Without loss of generality, eventually replacing the vector field $a(\cdot)$ by $a(\cdot)/\nu$ we may assume that $\nu = 1$. Therefore, scaling back, we see that all the constants depending on ν, L will actually depend on the unique quantity L/ν, and as such they will be denoted for the rest of the paper.

2.2. Heisenberg groups

We identify the Heisenberg group $\mathbb{H}^n$ with $\mathbb{R}^{2n+1}$, $n \geqslant 1$, via the exponential coordinates in (1.7), see also (2.3) below. The group multiplication is given by

$$(x_1, \ldots, x_{2n}, t) \cdot (y_1, \ldots, y_{2n}, s) = \left(x_1 + y_1, \ldots, x_{2n} + y_{2n}, t + s + \frac{1}{2}\sum_{i=1}^{n}(x_i y_{n+i} - x_{n+i} y_i)\right),$$

and makes $\mathbb{H}^n$ a non-commutative group. For $1 \leqslant i \leqslant n$ the canonical left invariant vector fields are those in (1.8)–(1.9). The only non-trivial commutator is

$$T = \partial_t = [X_i, X_{n+i}] \equiv X_i X_{n+i} - X_{n+i} X_i, \quad \text{for every } i = 1, \ldots, n. \tag{2.1}$$

The vector fields $X_1, X_2, \ldots, X_{2n}$ are called horizontal vector fields, while T is the vertical vector field. The horizontal gradient of a function $u : \mathbb{H}^n \mapsto \mathbb{R}$ is the vector $\mathfrak{X}u$ defined in (1.9). The vector fields $\{X_i\}_i$ enjoy the remarkable property of being opposite to their formal adjoint, that is

$$X_i^* = -X_i, \quad \text{for every } i = 1, \ldots, 2n. \tag{2.2}$$

The second horizontal derivatives are given by the $2n \times 2n$ matrix $\mathfrak{X}\mathfrak{X}u = \mathfrak{X}^2u$ with entries $(\mathfrak{X}(\mathfrak{X}u))_{i,j} = (\mathfrak{X}\mathfrak{X}u)_{i,j} = X_i(X_j(u))$. Note that such a matrix is not symmetric due to the non-commutativity of the horizontal vector fields X_i. We shall denote the standard Euclidean gradient of a function u as $Du = (D_1u, \dots, D_{2n+1}u)$. For notational convenience, when referring to the coordinates and vector fields in (1.7)–(1.8) we shall also denote $Y_s = X_{s+n}$ and $y_s = x_{s+n}$, for $s \in \{1, \dots, n\}$.

The Heisenberg Lie algebra $\mathfrak{h}^n$ is a Step 2 nilpotent Lie algebra. This means that $\mathfrak{h}^n$ admits a decomposition as a direct sum of vector spaces $\mathfrak{h}^n = \mathfrak{h}_0 \oplus \mathfrak{h}_1$ such that $[\mathfrak{h}_0, \mathfrak{h}_0] = \mathfrak{h}_1$. The horizontal part $\mathfrak{h}_0$ is generated by $\{X_1, \dots, X_n, Y_1, \dots, Y_n\}$ 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 $\exp : \mathfrak{h}^n \mapsto \mathbb{H}^n$ is a global diffeomorphism. A point $x \in \mathbb{H}^n$ has exponential coordinates $(x_1, \dots, x_n, y_1, \dots, y_n, t)$ if

$$x = \exp\left(\left(\sum_{j=1}^{n} x_i X_i + y_i Y_i\right) + tT\right). \tag{2.3}$$

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 (1.7); in the following we shall denote $\exp(Z) \equiv \mathrm{e}^Z$.

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

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

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 [14] implies that $\Gamma(x, y) \neq \emptyset$.

For convenience, we fix an ambient Riemannian metric in $\mathbb{H}^n$ so that the set $\mathfrak{h}_0 = \{X_1, \dots, X_n, Y_1, \dots, Y_n\}$ 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 (CC-distance) is then defined by

$$d_{\mathrm{CC}}(x, y) = \inf\bigl\{\text{length}(\gamma) \colon \gamma \in \Gamma(x, y)\bigr\}. \tag{2.4}$$

It depends only on the restriction of the ambient Riemannian metric to the horizontal distribution generated by the horizontal tangent space. In the following, with $A, B \subset \mathbb{H}^n$ being non-empty subsets, we denote $\text{dist}(A, B) := \inf\{d_{\mathrm{CC}}(x, y) \colon x \in A,\ y \in B\}$, the Carnot–Carathèodory distance between sets. For more on CC-distances and general properties of metrics related to vector fields we refer to the classical paper [44].

2.3. CC-balls, and the homogeneous dimension Q

The Carnot gauge is $|x|_{\mathrm{CC}} = d_{\mathrm{CC}}(x, 0)$. A few explicit formulas are available [2], but it is probably more convenient to work with an equivalent gauge [2], smooth away from the origin, called the Heisenberg gauge:

$$|x|_{\mathbb{H}^n} := \left(\left(\sum_{j=1}^{n} x_i^2 + y_i^2\right)^2 + t^2\right)^{1/4} \approx |x|_{\mathrm{CC}}. \tag{2.5}$$

In this paper all the balls, centered at $x_0 \in \mathbb{H}^n$ and with radius R, will be defined with respect to the CC-distance, that is $B(x_0, R) = \{y \in \mathbb{H}^n\colon d_{CC}(x_0, y) < R\}$. In view of (2.5) they are equivalent to the gauge balls obviously defined by $\{y \in \mathbb{H}^n\colon |y^{-1} \cdot x_0|_{\mathbb{H}^n} < R\}$. The non-isotropic dilations are the group homomorphisms given by

$$\delta_R(x_1, \ldots, x_n, y_1, \ldots, y_n, t) = \bigl(Rx_1, \ldots, Rx_n, Ry_1, \ldots, Ry_n, R^2 t\bigr), \tag{2.6}$$

where $R > 0$. The point is that we get the ball centered at the origin of radius $R > 0$ by applying the non-isotropic dilation δ_R to the unit ball centered at the origin, that is

$$B(0, R) = \delta_R B(0, 1). \tag{2.7}$$

The equivalence (2.5) and the natural scaling in (2.6) leads to define the number $Q = 2n + 2$ as the homogeneous dimension of $\mathbb{H}^n$. In particular, we have $|B(x_0, R)| \approx R^Q$, where $|B_R|$ denotes the Lebesgue measure of the ball $B(x_0, R)$. From such an estimate the doubling property of the CC-balls B_R easily follows; specifically, for any $B(x_0, R) \subset \mathbb{H}^n$, there holds

$$\bigl|B(x_0, 2R)\bigr| \leqslant C_d \bigl|B(x_0, R)\bigr|. \tag{2.8}$$

In the following, when clear, or not essential to 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. Finally, again when no ambiguity will arise, we shall also denote $\lambda B \equiv B(x_0, \lambda R)$, if $B \equiv B(x_0, R)$, and, when the center of the ball will not be important, we shall use the short-hand notation $B(x_0, R) \equiv B_R$. Moreover, when some constant will depend on the homogeneous dimension Q, such a dependence will be very often indicated as on the number n.

Let $B_R \subset \mathbb{R}^n$ be a ball, and $f : B_R \to \mathbb{R}^k$ be an integrable map; we define the average of f over the ball B_R as

$$(f)_R \equiv (f)_{B_R} := \fint_{B_R} f(x)\,dx = \frac{1}{|B_R|} \int_{B_R} f(x)\,dx \approx R^{-Q} \int_{B_R} f(x)\,dx. \tag{2.9}$$

The following Krylov–Safonov type covering lemma may be inferred from [25,33].

Lemma 2.1. *Let $B_R \subset \mathbb{H}^n$ be a ball with radius R, and let $\delta \in (0, 1)$. Assume that $E, G \subset B_R$ are measurable sets such that $|E| \leqslant \delta |B_R|$. Assume also that for any ball $B(x_0, \varrho)$ centered in B_R, with $\varrho \leqslant 2R$, and such that $|E \cap B(x_0, 5\varrho)| > \delta |B_R \cap B(x_0, \varrho)|$, there holds $E \cap B(x_0, 5\varrho) \subset G$. Then it follows that $|E| \leqslant \delta |G|$.*

2.4. Horizontal Sobolev spaces and weak solutions

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 $L^p(\Omega)$, that is $\mathfrak{X}u \in L^p(\Omega, \mathbb{R}^{2n})$. $HW^{1,p}(\Omega)$ is a Banach space when equipped with the norm defined by $\|u\|_{HW^{1,p}(\Omega)} := \|u\|_{L^p(\Omega)} + \|\mathfrak{X}u\|_{L^p(\Omega,\mathbb{R}^{2n})}$, for $p \geqslant 1$. The closure of $C_0^\infty(\Omega)$ in $HW^{1,p}(\Omega)$ is denoted by $HW_0^{1,p}(\Omega)$, while the local variant $HW^{1,p}_{\mathrm{loc}}(\Omega)$ is obviously 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 $\Omega' \Subset \Omega$. Now, keeping (2.2) in mind, a weak solution to Eq. (1.20) with $F \in L^p(\Omega, \mathbb{R}^{2n})$ is a function $u \in HW^{1,p}(\Omega)$ such that

$$\int_\Omega b(x) \sum_{i=1}^{2n} a_i(\mathfrak{X}u) X_i\varphi\,dx = \int_\Omega \sum_{i=1}^{2n} |F|^{p-2} F_i X_i\varphi\,dx, \quad \text{for all } \varphi \in HW_0^{1,p}(\Omega). \tag{2.10}$$

Therefore, when considering Eq. (1.1), this means to require that

$$\int_\Omega \sum_{i=1}^{2n} a_i(\mathfrak{X}u) X_i\varphi\,dx = 0, \quad \text{for all } \varphi \in HW_0^{1,p}(\Omega). \tag{2.11}$$

A crucial result concerning horizontal Sobolev spaces is the following Heisenberg group version of the Sobolev embedding theorem.

Theorem 2.1. *Let $w \in HW_0^{1,q}(B)$ with $1 < q < Q$, where $B \subset \mathbb{H}^n$ is a CC-ball. Then there exists a constant $c \equiv c(n,q)$ such that*

$$\left(\fint_B |w|^{\frac{Qq}{Q-q}}\,dx\right)^{\frac{Q-q}{Qq}} \leqslant c|B|^{\frac{1}{Q}} \left(\fint_B |\mathfrak{X}w|^q\,dx\right)^{\frac{1}{q}}. \tag{2.12}$$

A proof of the previous result can be found for instance in [10,38], where the statement is given in the case of balls with a suitably small radius $r \leqslant R_0$. The general case stated above easily follows by a standard scaling argument, using the dilation operator in (2.6) and (2.7). See also the proof of Proposition 7.1 below, end of Step 2.

2.5. *Vanishing mean oscillations*

Let $b : \Omega \to \mathbb{R}$ be a measurable function, and $\Omega' \Subset \Omega$; we define

$$[b]_{R_0} \equiv [b]_{R_0,\Omega'} := \sup_{B_R \subset \Omega',\, R \leqslant R_0} \fint_{B_R} \left|b(x) - (b)_{B_R}\right| dx, \tag{2.13}$$

where $R_0 > 0$, B_R is any CC-ball with radius R, and, accordingly to (2.9)

$$(b)_R \equiv (b)_{B_R} := \fint_{B_R} b(x)\,dx. \tag{2.14}$$

The function b is said to have (locally) vanishing mean oscillation, that is, to be a VMO-function iff, for every choice of the subset $\Omega' \Subset \Omega$ it holds that

$$\lim_{R \searrow 0} [b]_{R,\Omega'} = 0. \tag{2.15}$$

2.6. Difference quotients

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

Definition 2.1. Let Z be a vector field in $\mathbb{H}^n$. The difference quotient of the function w at the point x is

$$D_h^Z w(x) = \frac{w(x\mathrm{e}^{hZ}) - w(x)}{h}, \quad h \neq 0.$$

The latter definition will be always used whenever the function w in question is defined both at $x\mathrm{e}^{hZ}$ and at x. The following lemma collects a few standard properties of difference quotients that can be for instance inferred from [8,16,24,29,40].

Lemma 2.2. *Let $\Omega' \Subset \Omega$ be an open subset. Let Z being a left-invariant vector field, and $w \in L^p_{\mathrm{loc}}(\Omega)$ for $p > 1$. If there exist two positive constants $\sigma < \operatorname{dist}(\Omega', \partial\Omega)$ and C such that*

$$\sup_{0<|h|<\sigma} \int_{\Omega'} \left|D_h^Z w\right|^p dx \leqslant C^p$$

then $Zw \in L^p(\Omega')$ and $\|Zw\|_{L^p(\Omega')} \leqslant C$. Conversely, if $Zw \in L^p(\Omega')$ then for some $\sigma > 0$

$$\sup_{0<|h|<\sigma} \int_{\Omega'} \left|D_h^Z w\right|^p dx \leqslant c(p)\|Zw\|^p_{L^p(\Omega)}.$$

Moreover $D_h^Z w \to Zw$ strongly in $L^p(\Omega')$.

Finally a trivial lemma, which is basically a consequence of the Campbell–Hausdorff formula; the proof is left to the reader.

Lemma 2.3. *Let $\varphi \in HW^{1,t}(\Omega)$, and X, Z be smooth left-invariant vector fields such that $[X, Z]\varphi \in L^t_{\mathrm{loc}}(\Omega)$, with $t \geqslant 1$. If $\tilde{\varphi} := \varphi(x\mathrm{e}^Z)$ then $X\tilde{\varphi} \in L^t_{\mathrm{loc}}(\Omega)$ and*

$$X\left[\varphi\left(\cdot \mathrm{e}^Z\right)\right](x) = X\tilde{\varphi}(x) = X\varphi\left(x\mathrm{e}^Z\right) + [X, Z]\varphi\left(x\mathrm{e}^Z\right) \tag{2.16}$$

holds provided $x, x\mathrm{e}^Z \in \Omega$. As a consequence we have, for $h \neq 0$

$$X\left(D_h^Z\varphi\right)(x) = D_h^Z(X\varphi)(x) + [X, Z]\varphi\left(x\mathrm{e}^{hZ}\right). \tag{2.17}$$

Before going on, first two algebraic lemmata; see [28], for instance.

Lemma 2.4. *Let $1 < p < \infty$. There exists a constant $c = c(n, p) > 1$, independent of $\mu \in [0, 1]$, such that, for any $z_1, z_2 \in \mathbb{R}^{2n}$*

$$c^{-1}\big(\mu^2+|z_1|^2+|z_2|^2\big)^{\frac{p-2}{2}} \leqslant \int_0^1 \big(\mu^2+|z_2+\tau z_1|^2\big)^{\frac{p-2}{2}}\,d\tau$$
$$\leqslant c\big(\mu^2+|z_1|^2+|z_2|^2\big)^{\frac{p-2}{2}}. \tag{2.18}$$

Lemma 2.5. *Let $1<p<\infty$. There exists a constant $c\equiv c(n,p)>1$, independent of $\mu\in[0,1]$, such that, for any $z_1,z_2\in\mathbb{R}^{2n}$*

$$c^{-1}\big(\mu^2+|z_1|^2+|z_2|^2\big)^{\frac{p-2}{2}}|z_2-z_1|^2 \leqslant \big|\big(\mu^2+|z_2|^2\big)^{\frac{p-2}{4}}z_2-\big(\mu^2+|z_1|^2\big)^{\frac{p-2}{4}}z_1\big|^2$$
$$\leqslant c\big(\mu^2+|z_1|^2+|z_2|^2\big)^{\frac{p-2}{2}}|z_2-z_1|^2.$$

Finally a few general properties related to growth/ellipticity conditions (1.2)–(1.3).

Lemma 2.6. *The following equality holds:*

$$\big(D_h^Z a_i(\mathfrak{X}u)\big)(x)=\sum_{j=1}^{2n} a_{i,j}^Z(x)D_h^Z X_j u(x), \tag{2.19}$$

where

$$a_{i,j}^Z(x)=\int_0^1 D_{z_j}a_i\big(\mathfrak{X}u(x)+\tau h D_h^Z\mathfrak{X}u(x)\big)\,d\tau, \tag{2.20}$$

and $i,j\in\{1,\dots,2n\}$. Moreover there exists a constant $c\equiv c(n,p)\geqslant 1$ such that

$$\big|a_{i,j}^Z(x)\big|\leqslant c\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hZ}\big)\big|^2\big)^{\frac{p-2}{2}} \tag{2.21}$$

and

$$c^{-1}\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hZ}\big)\big|^2\big)^{\frac{p-2}{2}}|\lambda|^2\leqslant\sum_{i,j=1}^{2n}a_{i,j}^Z(x)\lambda_i\lambda_j, \tag{2.22}$$

hold for every $\lambda\in\mathbb{R}^{2n}$, whenever $x, x\mathrm{e}^{hZ}\in\Omega$.

Proof. The proof of (2.19) follows directly from the definition of $a_{i,j}^Z(x)$, while that of (2.22)–(2.21) follow from (1.2)–(1.3) and Lemma 2.4. □

Lemma 2.7. *Let $u\in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.3) *with $2\leqslant p<4$. Then for any $\varphi\in C_c^\infty(\Omega)$, left-invariant vector field Z and $h>0$ such that $|\mathrm{e}^{hZ}|_{CC}<\operatorname{dist}(\operatorname{supp}\varphi,\partial\Omega)$ we have*

$$\int_\Omega\sum_{i=1}^{2n}\big(D_h^Z a_i(\mathfrak{X}u)(x)X_i\varphi(x)+a_i(\mathfrak{X}u)\big(x\mathrm{e}^{hZ}\big)[Z,X_i]\varphi(x)\big)\,dx=0. \tag{2.23}$$

Proof. With $\tilde{\varphi}(x) := \varphi(x\mathrm{e}^{-hZ})$, using (2.16) we have that $X_i\tilde{\varphi}(x) = X_i\varphi(x\mathrm{e}^{-hZ}) + h[Z, X_i] \times \varphi(x\mathrm{e}^{-hZ})$. Testing (2.11) with $\tilde{\varphi}$ and changing variable $x \mapsto x\mathrm{e}^{hZ}$, we obtain

$$\int_{\Omega} \sum_{i=1}^{2n} a_i\big(\mathfrak{X}u\big(x\mathrm{e}^{hZ}\big)\big)\big(X_i\varphi(x) + h[Z, X_i]\varphi(x)\big)\,dx = 0.$$

Now we subtract (2.11) from the last identity and divide the resulting equation by h. This finally gives (2.23). $\square$

Finally, a standard property of weak derivatives in the Euclidean case, that holds in the present setting too. We give a sketchy proof for the sake of completeness.

Lemma 2.8. *Let $v, w \in L^1_{\mathrm{loc}}(\Omega)$ such that $vw, vX_sw, wX_sv \in L^1_{\mathrm{loc}}(\Omega)$ for some $s \in \{1, \dots, 2n\}$. Then $X_s(vw) \in L^1_{\mathrm{loc}}(\Omega)$ and $X_s(vw) = vX_sw + wX_sv$.*

Proof. We first assume that both the functions are locally essentially bounded. Then we mollify them using standard mollifiers φ_ε, obtaining $v_\varepsilon = v * \varphi_\varepsilon$, $w_\varepsilon = w * \varphi_\varepsilon$, so that $v_\varepsilon \to v$ and $w_\varepsilon \to w$ almost everywhere and $X_sv_\varepsilon \to X_sv$ and $X_sw_\varepsilon \to X_sw$ locally in $L^1(\Omega)$; see the formulas in the proof of [26, Theorem 11.9] for details. Therefore, using that $v_\varepsilon, w_\varepsilon$ are locally uniformly bounded we get that $v_\varepsilon X_s w_\varepsilon \to vX_sw$ and $w_\varepsilon X_s v_\varepsilon \to wX_sv$ locally in $L^1(\Omega)$; at this point using the definition of distributional derivative in the X_s-direction the assertion of the lemma follows in this first case. In a second case we consider the situation when only one function is bounded, say v. We can apply the result of the first case to v and to the truncated function $w_k := \max\{\min\{w, k\}, -k\}$, for $k \in \mathbb{N}$, and the assertion follows using Lebesgue's dominated convergence when letting $k \nearrow \infty$, and the fact that vX_sw, wX_sv are supposed to be locally in $L^1(\Omega)$. Finally, the general case follows by the second one applying the same truncation argument of the second case to one of the two functions. $\square$

2.7. Maximal operators

Here we present a miscellanea of various maximal operators and related inequalities. Let $B_0 \subset \mathbb{R}^n$ be a CC-ball. We shall consider, in the following, the Restricted Maximal Function Operator relative to B_0. This is defined as

$$M^*_{B_0}(f)(x) := \sup_{B \subseteq B_0,\, x \in B} \fint_B \big|f(y)\big|\,dy, \tag{2.24}$$

whenever $f \in L^1(B_0)$, where B denotes any CC-ball contained in B_0, not necessarily with the same center, as long as it contains the point x. More generally, if $s \geqslant 1$ we define

$$M^*_{s,B_0}(f)(x) := \sup_{B \subseteq B_0,\, x \in B} \left(\fint_B \big|f(y)\big|^s\,dy\right)^{1/s} \tag{2.25}$$

whenever $f \in L^s(B_0)$; of course $M^*_{1,B_0} \equiv M^*_{B_0}$. Another type of restricted – but "centered" – maximal operator is given by

$$M_R(f)(x) := \sup_{B(x,r),\, r \leqslant R} \fint_{B(x,r)} \big|f(y)\big|\, dy. \tag{2.26}$$

We recall the following weak type (1, 1) estimate for $M^*_{B_0}$:

$$\big|\big\{x \in B_0\colon\ M^*_{B_0}(f)(x) \geqslant \lambda\big\}\big| \leqslant \frac{c_W}{\lambda^\gamma} \int_{B_0} \big|f(y)\big|^\gamma\, dy, \quad \text{for every } \lambda > 0 \text{ and } \gamma \geqslant 1, \tag{2.27}$$

which is valid for any $f \in L^1(B_0)$; the constant c_W depends only on the homogeneous dimension Q via the doubling constant C_d in (2.8), and therefore ultimately on n; for this and related issues we refer to [46]. A standard consequence of (2.27) is then

$$\int_{B_0} \big|M^*_{B_0}(f)\big|^\gamma\, dx \leqslant \frac{c(Q,\gamma)}{\gamma - 1} \int_{B_0} |f|^\gamma\, dx, \quad \text{for every } \gamma > 1. \tag{2.28}$$

A straightforward consequence of (2.28) is the following similar estimate for M^*_{s,B_0}:

$$\int_{B_0} \big|M^*_{s,B_0}(f)\big|^\gamma\, dx \leqslant \frac{c(Q,\gamma)}{s(\gamma - s)} \int_{B_0} |f|^\gamma\, dx, \quad \text{for every } \gamma > s. \tag{2.29}$$

Finally, we report an inequality due to Hajlasz and Strzelecki [27], see also [26], Section 3, for related results.

Proposition 2.1. *Let $f \in HW^{1,1}(\Omega)$ and $R > 0$. Then there exists an absolute constant $c \equiv c(n)$ such that*

$$\big|f(x) - f(y)\big| \leqslant c\big[M_R\big(|\mathfrak{X}f|\big)(x) + M_R\big(|\mathfrak{X}f|\big)(y)\big] d_{\mathrm{CC}}(x,y)$$

whenever $d_{\mathrm{CC}}(x,y) \leqslant R/2 \leqslant \operatorname{dist}(\Omega', \partial\Omega)/2$ and $x, y \in \Omega' \Subset \Omega$.

3. Basic regularity

In this section we summarize and revisit a few regularity results known for solutions to (1.1), in order to get statements in a form tailored to our later needs.

3.1. Basic regularity results

The following is a basic result of Capogna, Danielli and Garofalo [10], and Lu [38].

Theorem 3.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.3) *with $p > 1$. Then there exists a positive Hölder exponent $\alpha \equiv \alpha(n, p, L/\nu)$ such that $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$. In particular, u is a locally bounded function, and for every open subset $\Omega' \Subset \Omega$*

there exists a constant c, depending only on $n, p, L/\nu$, *and* $\operatorname{dist}(\Omega', \partial\Omega)$ *but otherwise independent of* $\mu \in [0, 1]$, *of the solutions u and on the vector field* $a(\cdot)$, *such that*

$$\|u\|_{L^\infty(\Omega')} \leqslant c\big(\|u\|_{L^p(\Omega)} + \mu\big). \tag{3.1}$$

Just let us observe that the validity of (3.1) directly follows from the weak Harnack inequality of Theorem 3.2 in [10], via a standard covering argument. Now another basic result, due to Domokos [16], see also [42].

Theorem 3.2. *Let* $u \in HW^{1,p}(\Omega)$ *be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4), *with* $2 \leqslant p < 4$. *Then we have* $Tu \in L^p_{\mathrm{loc}}(\Omega)$. *Moreover, for every couple of open subsets* $\Omega' \Subset \Omega'' \subset \Omega$ *there exists a constant c depending only on* $\operatorname{dist}(\Omega', \partial\Omega'')$, *n, p,* L/ν, *but otherwise independent of* $\mu \in (0, 1]$, *of the solutions u and on the vector field* $a(\cdot)$, *such that*

$$\int_{\Omega'} |Tu|^p \, dx \leqslant c \int_{\Omega''} \big(\mu + |\mathfrak{X}u|\big)^p \, dx. \tag{3.2}$$

In the previous estimate $c \nearrow \infty$ *when* $p \nearrow 4$.

Proof. The proof of the fact that $Tu \in L^p_{\mathrm{loc}}(\Omega)$ is contained in Theorem 1.2 from [16]. In order to get estimate (3.2) we first use the estimate contained in Theorem 1.2 from [16], that gives

$$\int_{B_{\gamma R}} |Tu|^p \, dx \leqslant c \int_{B_R} \big(|\mathfrak{X}u|^p + |u|^p + \mu^p\big) \, dx,$$

whenever $B_R \Subset \Omega$ and where $\gamma \in (0, 1)$; the constant c here depends on $n, p, L/\nu, \gamma$ and R. Then we observe that if u weakly solves (1.1) then so does $u - (u)_{B_R}$ and therefore, applying the previous estimate to this new function we get

$$\int_{B_{\gamma R}} |Tu|^p \, dx \leqslant c \int_{B_R} \big(|\mathfrak{X}u|^p + \big|u - (u)_{B_R}\big|^p + \mu^p\big) \, dx. \tag{3.3}$$

Now, in order to get rid of the integrals involving u in the previous estimate, we use Jerison's Poincaré inequality [32], that is $\|u - (u)_{B_R}\|_{L^p(B_R)} \leqslant c(n, p) R \|\mathfrak{X}u\|_{L^p(B_R)}$. Now (3.2) follows by joining the previous inequality to (3.3) and finally using a standard covering argument. Note that the constant c in (3.2) critically depends on $\operatorname{dist}(\Omega', \partial\Omega)$ in the sense that $c \nearrow \infty$ when $\operatorname{dist}(\Omega', \partial\Omega) \searrow 0$. The constant c remains bounded when $\mu \searrow 0$ as a careful inspection of the proof of Theorem 1.2 from [16] reveals. □

The proof of the following result can be found in [40], Theorem 8.

Theorem 3.3. *Let* $u \in HW^{1,p}(\Omega)$ *be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4), *with* $2 \leqslant p < 4$. *Assume also that* $\mathfrak{X}u \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$, *where* $q \geqslant p$ *satisfies*

$$p < 2 + \frac{q}{n+1}. \tag{3.4}$$

Then we have $Tu \in L^{\infty}_{\mathrm{loc}}(\Omega)$. Moreover, let $B_r = B(x_0, r) \Subset \Omega$, then we have

$$\|Tu\|_{L^{\infty}(B_\rho)} \leqslant \left(\frac{c}{r-\rho}\right)^{\frac{\chi}{\chi-1}} \left(\frac{\|\mu + |\mathfrak{X}u|\|_{L^q(B_r)}}{\mu}\right)^{\frac{(p-2)\chi}{2(\chi-1)}} \|Tu\|_{L^{\frac{2q}{q-p+2}}(B_r)}, \tag{3.5}$$

for every $B_\rho = B(x_0, \rho) \subset B_r$, where

$$\chi = \frac{Q}{Q-2}\frac{q-p+2}{q} > 1. \tag{3.6}$$

The constant c only depends on n, p, L/ν, being otherwise independent of the particular solution u, the constant μ, and the vector field $a(\cdot)$, and q.

We just remark that conditions (3.4) and (3.6) are actually equivalent.

3.2. Difference quotients results

Before going on let us clarify a few conventions we shall adopt for the rest of the paper when dealing with difference quotients as defined in Lemma 2.2; such conventions should be kept in mind in the following especially when reading the proofs of Lemma 3.1 and Proposition 7.1 below. By the writing "$h \to 0$" we shall implicitly mean "$h_k \to 0$," since we shall actually have $h \equiv h_k$ where $\{h_k\}_k$ is a positive decreasing sequence such that $h_k \to 0$; we shall also eventually, and actually very often, pass to non-relabeled sub-sequences that will still be denoted by $\{h_k\}_k$. This will be useful since when letting $h \to 0$ we shall need to use certain real analysis convergence results, that are valid up to the passage to sub-sequences. With such a definition/use of $D^Z_h \equiv D^Z_{h_k}$, all the standard properties of difference quotients remain valid, and the final results are the same, since the point in the use of difference quotients is approximating real derivatives with discrete finite difference operators. Finally in the following we shall state convergence results such as "$G(x\mathrm{e}^{hZ}) \to G(x)$ in $L^t_{\mathrm{loc}}(\Omega)$" as $h \to 0$, for some $G \in L^t(\Omega)$, and a smooth vector field Z. This must be interpreted as follows: it is clear that it makes sense to consider $G(x\mathrm{e}^{hZ})$ only provided $x\mathrm{e}^{hZ} \in \Omega$; on the other hand, for each open subset $\Omega'' \Subset \Omega$ there exists a number $h_0 > 0$, depending on Ω'' and Z, such that $x\mathrm{e}^{hZ} \in \Omega$ provided $x \in \Omega''$ and $|h| \leqslant h_0$. Therefore by the previous convergence statement on $G(x\mathrm{e}^{hZ})$ we actually mean $G(x\mathrm{e}^{hZ}) \to G(x)$ in $L^t(\Omega'')$, where $0 \swarrow |h| \leqslant h_0$, for every possible choice of the open subset $\Omega'' \Subset \Omega$.

The next lemma summarizes and exploits various difference quotient arguments and results scattered in [16] and [40].

Lemma 3.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Then we have*

$$D^{e_i}_h \mathfrak{X}u \to D_i \mathfrak{X}u \quad \text{in } L^2_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \text{ for every } i = 1, \dots, 2n+1, \tag{3.7}$$

and therefore

$$|\mathfrak{X}\mathfrak{X}u|^2 + |T\mathfrak{X}u|^2 \in L^1_{\mathrm{loc}}(\Omega). \tag{3.8}$$

Moreover

$$\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big[|\mathfrak{X}\mathfrak{X}u|^2+|T\mathfrak{X}u|^2\big]\in L^1_{\mathrm{loc}}(\Omega), \tag{3.9}$$

and for every choice of open subset $\Omega'\Subset\Omega''\Subset\Omega$ there exists a constant c depending only on n, p, L/ν and $\operatorname{dist}(\Omega',\partial\Omega'')$ *such that*

$$\int_{\Omega'}\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big[|\mathfrak{X}\mathfrak{X}u|^2+|T\mathfrak{X}u|^2\big]\,dx\leqslant c\int_{\Omega''}\big(|\mathfrak{X}u|^p+|Tu|^p+\mu^p\big)\,dx. \tag{3.10}$$

In the last inequality the constant c is in particular independent of $\mu\in(0,1]$, of the solution u, and of the vector field $a(\cdot)$. Finally, we have

$$a(\mathfrak{X}u)\in W^{1,\frac{p}{p-1}}_{\mathrm{loc}}\big(\Omega,\mathbb{R}^{2n}\big). \tag{3.11}$$

Proof. We have to go back to the difference quotient arguments of [16] and [40] where the inclusions in (3.9) are proved; in particular we refer to Section 3 of [40]. Then, due to the non-degeneracy condition $\mu>0$, we have that $\mathfrak{X}u\in W^{1,2}_{\mathrm{loc}}(\Omega,\mathbb{R}^n)$, and this fact immediately implies (3.7) and (3.8) via Lemma 2.2. In order to establish the remaining implications we shall argue first to get differentiation assertions with respect to the horizontal directions X_i, $i=1,\dots,2n$; then, in view of [40, Theorem 7] the same arguments will apply when taking difference quotients with respect to the vertical direction T, that is D_h^T. By the proof of Theorem 1.3 in [16] we see that the quantity $(\mu^2+|\mathfrak{X}u(x)|^2+|\mathfrak{X}u(xe^{hX_i})|^2)^{\frac{p-2}{2}}|D_h^{X_i}\mathfrak{X}u(x)|^2$ remains locally bounded in $L^1(\Omega,\mathbb{R}^{2n})$, or more precisely, it stays bounded in $L^1(\Omega',\mathbb{R}^{2n})$ for every $\Omega'\Subset\Omega$, as long as h is suitably small, depending on Ω' – see the "conventions" immediately before the lemma. Therefore, we also see that the quantity $D_h^{X_i}[(\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{4}}\mathfrak{X}u]$ remains locally bounded in $L^2(\Omega,\mathbb{R}^{2n})$ since an application of Lemma 2.5 gives

$$\begin{aligned}&\int_{\Omega'}\big|D_h^{X_i}\big(\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{4}}\mathfrak{X}u\big)\big|^2\,dx\\&\quad\leqslant c(n,p)\int_{\Omega'}\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_i}\big)\big|^2\big)^{\frac{p-2}{2}}\big|D_h^{X_i}\mathfrak{X}u\big|^2\,dx.\end{aligned}$$

Therefore by Lemma 2.2 we have that $X_i((\mu^2+|\mathfrak{X}u|^2)^{\frac{p-2}{4}}\mathfrak{X}u)\in L^2_{\mathrm{loc}}(\Omega,\mathbb{R}^{2n})$ and

$$D_h^{X_i}\big(\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{4}}\mathfrak{X}u\big)\to X_i\big(\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{4}}\mathfrak{X}u\big)\quad\text{in } L^2_{\mathrm{loc}}\big(\Omega,\mathbb{R}^{2n}\big). \tag{3.12}$$

Moreover, as $\mathfrak{X}\mathfrak{X}u, T\mathfrak{X}u\in L^2_{\mathrm{loc}}(\Omega)$, we may assume that

$$\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_i}\big)\big|^2\big)^{\frac{p-2}{2}}\big|D_h^{X_i}\mathfrak{X}u(x)\big|^2\to\big(\mu^2+2\big|\mathfrak{X}u(x)\big|^2\big)^{\frac{p-2}{2}}\big|X_i\mathfrak{X}u(x)\big|^2,$$

and

$$D_h^{X_i}\mathfrak{X}u(x) \to X_i\mathfrak{X}u(x)$$

almost everywhere. In turn this last fact together with another application of Lemma 2.5, and the use of (3.12) allow to apply a well-known variant of Lebesgue's dominated convergence theorem, finally yielding

$$\begin{aligned}&\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hX_i}\big)\big|^2\big)^{\frac{p-2}{2}}\big|D_h^{X_i}\mathfrak{X}u(x)\big|^2\\&\quad\to\big(\mu^2+2\big|\mathfrak{X}u(x)\big|^2\big)^{\frac{p-2}{2}}\big|X_i\mathfrak{X}u(x)\big|^2\quad\text{in } L^1_{\mathrm{loc}}(\Omega).\end{aligned}\tag{3.13}$$

Now, according to the notation used Lemma 2.6, we write

$$D_h^{X_i}\big(a_i(\mathfrak{X}u)\big)(x)=\int_0^1 Da\big(\mathfrak{X}u(x)+\tau hD_h^{X_i}\mathfrak{X}u(x)\big)\,d\tau\,D_h^{X_i}\mathfrak{X}u(x),\tag{3.14}$$

so that $D_h^{X_i}a(\mathfrak{X}u)\to Da(\mathfrak{X}u)X_i\mathfrak{X}u$ almost everywhere. Using (3.14) and again Lemma 2.6, we have

$$\big|D_h^{X_i}\big(a_i(\mathfrak{X}u)\big)(x)\big|\leqslant c(n,p,L)\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hX_i}\big)\big|^2\big)^{\frac{p-2}{2}}\big|D_h^{X_i}\mathfrak{X}u(x)\big|.$$

Therefore, using Lemma 3.2 below with $\varepsilon=1$, we have

$$\begin{aligned}\big|D_h^{X_i}\big(a_i(\mathfrak{X}u)\big)(x)\big|^{\frac{p}{p-1}}&\leqslant c\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hX_i}\big)\big|^2\big)^{\frac{p-2}{2}}\big|D_h^{X_i}\mathfrak{X}u(x)\big|^2\\&\quad+c\big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(x\mathrm{e}^{hX_i}\big)\big|^2\big)^{\frac{p}{2}}.\end{aligned}$$

Therefore $D_h^{X_i}(a(\mathfrak{X}u))\to Da(\mathfrak{X}u)X_i\mathfrak{X}u$ in $L^{\frac{p}{p-1}}_{\mathrm{loc}}(\Omega,\mathbb{R}^{2n})$ follows applying Lemma 2.2 by (3.13) and again the well-known variant of Lebesgue's dominated convergence theorem, and in a similar way (3.11) also follows. Finally, as already mentioned above, the differentiability results involving $T\mathfrak{X}u$ follow exactly as those involving $\mathfrak{X}\mathfrak{X}u$; see for instance [40, Theorem 7]. In particular the local estimate thereby included implies the one in (3.10) via a standard covering argument. The peculiar dependence of the constant c comes from a straightforward analysis of the proofs in [16,40]. □

Lemma 3.2. *For every $a,b\geqslant 0$, $p\geqslant 2$, and $\varepsilon>0$ we have $(a^{p-2}b)^{\frac{p}{p-1}}\leqslant\varepsilon a^{p-2}b^2+c(p,\varepsilon)a^p$.*

Proof. When $p\neq 2$ – otherwise the statement is trivial – just write

$$\big(a^{p-2}b\big)^{\frac{p}{p-1}}=a^{\frac{p(p-2)}{2(p-1)}}a^{\frac{p(p-2)}{2(p-1)}}b^{\frac{p}{p-1}}$$

and then apply the standard Young's inequality with conjugate exponents $2(p-1)/p$ and $2(p-1)/(p-2)$. □

3.3. Higher integrability in Gehring's style

Let us first report a few trivial consequences of assumptions (1.2)–(1.3), see also [43], Section 2.2. Since $p \geqslant 2$, assumption (1.3) implies, for any $z_1, z_2 \in \mathbb{R}^{2n}$

$$c^{-1}|z_2 - z_1|^p \leqslant \langle a(z_2) - a(z_1), z_2 - z_1 \rangle. \tag{3.15}$$

Finally, inequality (1.2), together with a standard use of Young's inequality, yield for every $z \in \mathbb{R}^{2n}$

$$c^{-1}\left(\mu^2 + |z|^2\right)^{\frac{p-2}{2}}|z|^2 - c\mu^p \leqslant \langle a(z), z \rangle, \qquad c \equiv c(n, p, L/\nu) \geqslant 1. \tag{3.16}$$

Then a standard consequence of (1.2) and (3.16) follows in the next

Lemma 3.3. *Let $v \in u + HW_0^{1,p}(B_R)$ be the unique solution to the following Dirichlet problem:*

$$\begin{cases} \operatorname{div} a(\mathfrak{X}v) = 0 & \text{in } B_R \\ v = u & \text{on } \partial B_R, \end{cases} \tag{3.17}$$

where the vector field $a : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ satisfies (1.2)–(1.3) *for $p > 1$, and $B_R \Subset \Omega$ is a CC-ball. Then there exists a constant c depending only on n, p, L/ν, such that*

$$\int_{B_R} |\mathfrak{X}v|^p \, dx \leqslant c \int_{B_R} \left(\mu + |\mathfrak{X}u|\right)^p dx. \tag{3.18}$$

For a related proof using quasiminima see [24, Chapter 6], dealing with related, completely standard, Euclidean cases.

Next, a higher integrability result for solutions to (1.20), together with a first form of inequality (1.23). Note that here no upper bound on p is required.

Theorem 3.4. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.20) *under the assumptions* (1.2)–(1.3)*, with $p \geqslant 2$, and $F \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$ for some $q > p$. Then there exists $\tilde{q} > p$, depending only on n, p, L/ν, such that $\mathfrak{X}u \in L^{\tilde{q}}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$. Moreover, there exists a constant c depending only on n, p, L/ν such that for every CC-ball $B_{2R} \Subset \Omega$ the following reverse type inequality:*

$$\left(\fint_{B_R} |\mathfrak{X}u|^{q_0} \, dx\right)^{1/q_0} \leqslant c\left(\fint_{B_{2R}} \left(\mu + |\mathfrak{X}u|\right)^p dx\right)^{1/p} + c\left(\fint_{B_{2R}} |F|^{q_0} \, dx\right)^{1/q_0}, \tag{3.19}$$

holds whenever $p \leqslant q_0 \leqslant \tilde{q}$.

Proof. The proof more or less works as in the standard Euclidean setting, and we shall only give a sketch of it; see [24, Chapter 6] for the Euclidean case or directly [47]. Let $B_R \Subset \Omega$ be a CC-ball, and let us fix a cut-off function $\eta \in C_0^\infty(B_R)$ such that $0 \leqslant \eta \leqslant 1$, $\eta \equiv 1$ in $B_{R/2}$, and $|\mathfrak{X}\eta| \leqslant c/R$. The existence of such a function is as in [10], and in the specific setting of the

Heisenberg group it easily follows from (2.5) and the definition of CC-balls; see Section 2.3. Testing (2.10) by $\varphi = \eta^p(u-(u)_{B_R})$, and using (1.2) and (3.16) in a standard way together with Young's inequality, we get

$$\fint_{B_{R/2}} |\mathfrak{X}u|^p\,dx \leqslant cR^{-p}\fint_{B_R}\left|u-(u)_{B_R}\right|^p dx + c\fint_{B_R}\left(\mu^p+|F|^p\right)dx,$$

with $c \equiv c(n,p,L/\nu)$. See again [24, Chapter 6]. The intermediate integral in the last inequality can be estimated by using the Sobolev–Poincaré inequality in the Heisenberg group [32,38], that is

$$\fint_{B_R}\left|u-(u)_{B_R}\right|^p dx \leqslant cR^p\left(\fint_{B_R}|\mathfrak{X}u|^{p\sigma}\,dx\right)^{1/\sigma},$$

for some $\sigma \equiv \sigma(n,p)\in(0,1)$. Therefore, combining the last two inequalities we get

$$\fint_{B_{R/2}} |\mathfrak{X}u|^p\,dx \leqslant c\left(\fint_{B_R}|\mathfrak{X}u|^{p\sigma}\,dx\right)^{1/\sigma} + c\fint_{B_R}\left(\mu^p+|F|^p\right)dx.$$

This is a reverse-Hölder inequality with increasing support, in turn allowing to apply Gehring's lemma in the sub-elliptic setting – see for instance [47]. This finally yields the full statement and (3.19), after a few elementary manipulations. □

4. Interpolation and basic integrability

4.1. Interpolation inequalities

The following inequality is an end point instance of the general Gagliardo–Nirenberg inequality in the Euclidean spaces $\mathbb{R}^n$. For all $f\in C_0^\infty(\mathbb{R}^n)$, it holds that

$$\int_{\mathbb{R}^n}|\nabla f|^{\gamma+2}\,dx \leqslant c(n,\gamma)\|f\|^2_{L^\infty(\mathbb{R}^n)}\int_{\mathbb{R}^n}|\nabla f|^{\gamma-2}|\nabla^2 f|^2\,dx, \quad \gamma\geqslant 0. \tag{4.1}$$

The proof of the above inequality is elementary; indeed, it follows from integration by parts. In the rest of the section we shall give the analog of inequality (4.1) in the Heisenberg group; again, the proof involves only integration by parts. Actually, we shall first give a version of (4.1) for solutions to (1.1), that is the thing we are mainly interested in for the subsequent developments, and then, as a corollary of the proof given, a more general Heisenberg group version of (4.1) will follow in Theorem 4.1 below.

First a few technical preliminaries. Consider the following truncation operators:

$$\mathcal{T}_{\beta,k}(t) := \begin{cases}(\mu^2+t)^\beta & \text{if } t\in[0,k)\\ (\mu^2+k)^\beta & \text{if } t\in[k,\infty)\end{cases} \quad \text{for } t,\beta,k\geqslant 0,\ \mu>0. \tag{4.2}$$

To make the notation easier we shall also denote here $\mathcal{T}_\beta \equiv \mathcal{T}_{\beta,k}$, with the understanding that k is temporarily fixed.

Lemma 4.1. *For every choice of $\varepsilon \in (0,1)$, $\alpha, k \geqslant 0$, and $b \in \mathbb{R}$ it holds that*

$$2\mathcal{T}_{p/2+\alpha,k}\big(t^2\big)b \leqslant \varepsilon \mathcal{T}_{p/2+\alpha+1,k}\big(t^2\big) + \varepsilon^{-1}\mathcal{T}_{p/2+\alpha-1,k}\big(t^2\big)b^2. \tag{4.3}$$

Proof. First the case $t^2 < k$. Using the standard quadratic Young's inequality we have

$$\begin{aligned}\mathcal{T}_{p/2+\alpha,k}\big(t^2\big)b &= \sqrt{\varepsilon}\big(\mu^2+t^2\big)^{p/4+\alpha/2+1/2}(1/\sqrt{\varepsilon})\big(\mu^2+t^2\big)^{p/4+\alpha/2-1/2}b \\ &\leqslant (\varepsilon/2)\big(\mu^2+t^2\big)^{p/2+\alpha+1} + \big(\varepsilon^{-1}/2\big)\big(\mu^2+t^2\big)^{p/2+\alpha-1}b^2 \\ &= (\varepsilon/2)\mathcal{T}_{p/2+\alpha+1,k}\big(t^2\big) + \big(\varepsilon^{-1}/2\big)\mathcal{T}_{p/2+\alpha-1,k}\big(t^2\big)b^2, \end{aligned} \tag{4.4}$$

and (4.3) follows in this case. When $t^2 \geqslant k$ we write the previous chain of inequalities substituting μ^2+t^2 by μ^2+k everywhere in (4.4) and (4.3) follows in this case too. □

Lemma 4.2. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Then for all $\sigma \geqslant 0$ and $\eta \in C_c^\infty(\Omega)$, we have*

$$\begin{aligned}&\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p+2+\sigma}{2}}\,dx \\ &\quad\leqslant c\int_\Omega \big(\eta^2\mu^2+|\mathfrak{X}\eta|^2u^2\big)\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p+\sigma}{2}}\,dx \\ &\qquad + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_\Omega \eta^2\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p-2+\sigma}{2}}|X_sX_su|^2\,dx,\end{aligned} \tag{4.5}$$

where $c \equiv c(n,p,\sigma) > 0$.

Proof. For ease of notation in the following we let $\alpha := \sigma/2$. First let us observe that the very definition in (4.2) implies that the map $t \to \mathcal{T}_{p/2+\alpha}(t^2)t$ is globally Lipschitz continuous and therefore the chain rule in the Heisenberg group – see [10] – and the fact that $\mathfrak{X}u \in W^{1,2}_{\mathrm{loc}}(\Omega,\mathbb{R}^{2n})$ as given by Lemma 3.1, imply that

$$\eta^2\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)X_su \in W^{1,2}_{\mathrm{loc}}(\Omega), \tag{4.6}$$

holds for every $s \in \{1,\dots,2n\}$. Now, inclusion (4.6) allows for the following integration by parts:

$$\begin{aligned}P_0 &:= \int_\Omega \eta^2\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)(X_su)^2\,dx = \int_\Omega \eta^2\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)X_suX_su\,dx \\ &= -\int_\Omega u\eta^2\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)X_sX_su\,dx - 2\int_\Omega u\eta^2\mathcal{T}'_{p/2+\alpha}\big((X_su)^2\big)(X_su)^2X_sX_su\,dx \\ &\quad -2\int_\Omega u\eta X_s\eta\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)X_su\,dx =: P_1+P_2+P_3.\end{aligned} \tag{4.7}$$

Of course we used (2.2). Let us now estimate the three integrals defined in (4.7), that is P_1, P_2 and P_3. With $\varepsilon \in (0,1)$, by means of (4.3) we have

$$
\begin{aligned}
|P_1| &\leqslant \varepsilon \int_\Omega \eta^2 \mathcal{T}_{p/2+\alpha+1}\big((X_s u)^2\big)\,dx + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)} \int_\Omega \eta^2 \mathcal{T}_{p/2+\alpha-1}\big((X_s u)^2\big)|X_s X_s u|^2\,dx \\
&\leqslant \varepsilon P_0 + \int_\Omega \eta^2 \mu^2 \mathcal{T}_{p/2+\alpha}\big((X_s u)^2\big)\,dx \\
&\quad + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)} \int_\Omega \eta^2 \mathcal{T}_{p/2+\alpha-1}\big((X_s u)^2\big)|X_s X_s u|^2\,dx,
\end{aligned}
$$

as, obviously, $\mathcal{T}_{p/2+\alpha+1}((X_s u)^2) \leqslant \mathcal{T}_{p/2+\alpha}(X_s u)^2(\mu^2 + (X_s u)^2)$. In the previous inequality we have $c \equiv c(\varepsilon)$. The estimate of P_2 requires slightly more care; by Young's inequality and the definition in (4.2), we have

$$
\begin{aligned}
|P_2| &\leqslant (p+2\alpha)\|u\|_{L^\infty(\operatorname{supp}\eta)} \int_{\{(X_s u)^2 \leqslant k\}} \eta^2\big(\mu^2 + (X_s u)^2\big)^{\frac{p+2\alpha}{2}} |X_s X_s u|\,dx \\
&\leqslant \varepsilon \int_{\{(X_s u)^2 \leqslant k\}} \eta^2\big(\mu^2 + (X_s u)^2\big)^{\frac{p+2+2\alpha}{2}}\,dx \\
&\quad + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)} \int_{\{(X_s u)^2 \leqslant k\}} \eta^2\big(\mu^2 + (X_s u)^2\big)^{\frac{p-2+2\alpha}{2}} |X_s X_s u|^2\,dx \\
&\leqslant \varepsilon \int_\Omega \eta^2 \mathcal{T}_{p/2+\alpha}\big((X_s u)^2\big)\big(\mu^2 + (X_s u)^2\big)\,dx \\
&\quad + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)} \int_{\{(X_s u)^2 \leqslant k\}} \eta^2\big(\mu^2 + (X_s u)^2\big)^{\frac{p-2+2\alpha}{2}} |X_s X_s u|^2\,dx \\
&\leqslant \varepsilon P_0 + \int_\Omega \eta^2 \mu^2 \mathcal{T}_{p/2+\alpha}\big((X_s u)^2\big)\,dx \\
&\quad + c\|u\|^2_{L^\infty(\operatorname{supp}\eta)} \int_\Omega \eta^2 \mathcal{T}_{p/2+\alpha-1}\big((X_s u)^2\big)|X_s X_s u|^2\,dx,
\end{aligned}
$$

where again $c \equiv c(p, \varepsilon, \sigma)$. Finally, the estimation of P_3; again using standard Young's inequality

$$
\begin{aligned}
|P_3| &\leqslant \int_\Omega \eta|\mathfrak{X}\eta||u|\mathcal{T}_{p/2+\alpha}\big((X_s u)^2\big)|X_s u|\,dx \\
&\leqslant \varepsilon P_0 + c(\varepsilon) \int_\Omega |\mathfrak{X}\eta|^2 u^2 \mathcal{T}_{p/2+\alpha}\big((X_s u)^2\big)\,dx.
\end{aligned}
$$

Connecting the inequalities found for P_1, P_2, P_3 to (4.7) we have

$$P_0 \leqslant 3\varepsilon P_0 + c\int_\Omega \big(\eta^2\mu^2 + |\mathfrak{X}\eta|^2u^2\big)\mathcal{T}_{p/2+\alpha}\big((X_su)^2\big)\,dx$$
$$+ c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_\Omega \eta^2\mathcal{T}_{p/2+\alpha-1}\big((X_su)^2\big)|X_sX_su|^2\,dx,$$

where c depends on n, p, σ and ε. Observing that all the quantities involved in the previous inequality are finite as $\mathfrak{X}u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$, taking $\varepsilon = 1/6$, recalling that $\alpha = \sigma/2$, an easy manipulation now yields

$$\int_\Omega \eta^2\mathcal{T}_{p/2+\sigma/2,k}\big((X_su)^2\big)\big(\mu^2 + (X_su)^2\big)\,dx$$
$$\leqslant c\int_\Omega \big(\eta^2\mu^2 + |\mathfrak{X}\eta|^2u^2\big)\mathcal{T}_{p/2+\sigma/2,k}\big((X_su)^2\big)\,dx$$
$$+ c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_\Omega \eta^2\mathcal{T}_{p/2+\sigma/2-1,k}\big((X_su)^2\big)|X_sX_su|^2\,dx, \tag{4.8}$$

for any $s \in \{1,\dots,2n\}$, where c depends only on n, p and σ. At this point (4.5) follows summing up inequalities (4.8) for $s \in \{1,\dots,2n\}$ and eventually letting $k \nearrow \infty$, using the monotone convergence theorem. □

Remark 4.1. In the previous proof we never used that u is a solution of (1.1) but only that $\mathfrak{X}u$ locally belongs to $HW^{1,2}(\Omega, \mathbb{R}^{2n})$, and that u is locally bounded. Therefore neither the ellipticity ratio L/ν, nor the degeneracy parameter μ, appear in the dependence on the parameters of the constant c involved in (4.5).

We conclude with a more general statement extending the Euclidean one in (4.1), which is at this stage an obvious consequence of the proof of Lemma 4.2, and of the previous remark.

Theorem 4.1. *Let σ be a non-negative number and $p \geqslant 2$. Then for all $u \in C^\infty(\Omega)$ and $\eta \in C^\infty_c(\Omega)$, we have*

$$\int_\Omega \eta^2|\mathfrak{X}u|^{p+2+\sigma}\,dx \leqslant c\int_\Omega |\mathfrak{X}\eta|^2u^2|\mathfrak{X}u|^{p+\sigma}\,dx + c\int_\Omega \eta^2u^2\sum_{s=1}^{2n}|X_su|^{p-2+\sigma}|X_sX_su|^2\,dx,$$

where $c \equiv c(n, p, \sigma) > 0$.

4.2. Basic higher integrability

As an immediate corollary of Lemma 4.3 applied with $\sigma = 0$, and of Lemma 3.1, we gain a first higher integrability property of solutions to (1.1):

Lemma 4.3. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Then*

$$\mathfrak{X}u \in L^{p+2}_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big). \tag{4.9}$$

Moreover, for every couple of open subsets $\Omega' \Subset \Omega'' \Subset \Omega$ there exists a constant c depending only on n, p, L/ν, $\operatorname{dist}(\Omega', \partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that

$$\int_{\Omega'} |\mathfrak{X}u|^{p+2}\, dx \leqslant c \int_{\Omega''} \big(|\mathfrak{X}u|^p + |Tu|^p + \mu^p\big)\, dx. \tag{4.10}$$

Observe that (4.10) immediately follows by (4.5) with $\sigma = 0$, and by (3.10) via a standard covering argument – note that the choice of η, Ω' and Ω'' in (4.5) and (4.10) is arbitrary.

5. Caccioppoli type inequalities

In this section we shall derive a few preliminary energy estimates, or so-called Caccioppoli type inequalities, for the horizontal and vertical gradients $\mathfrak{X}u$ and Tu respectively. We shall modify some of the arguments introduced in [40] in order to find new types of Caccioppoli inequalities – that is, energy estimates. In turn these will be at the core of the main iteration in Section 7.

5.1. Smooth truncation operators

We shall start defining certain “smooth truncation operators” which are already used, in a slightly different from, in [40]. We define

$$g_{\alpha,k}(t) = \frac{k(\mu^2 + t)^\alpha}{k + (\mu^2 + t)^\alpha}, \quad t, \alpha \geqslant 0,\ \mu > 0,\ k \in \mathbb{N}. \tag{5.1}$$

We have that

$$0 \leqslant g_{\alpha,k}(t) \leqslant \min\big\{k, \big(\mu^2 + t\big)^\alpha\big\}, \quad \text{and} \quad 0 \leqslant g_{\alpha,k}(t) \leqslant g_{\alpha,k+1}(t) \tag{5.2}$$

hold for every $k \in \mathbb{N}$, and moreover

$$\lim_{k \to \infty} g_{\alpha,k}(t) = \big(\mu^2 + t\big)^\alpha. \tag{5.3}$$

A few elementary computations, actually a variant of the ones already presented in [40], Section 5.2, give that

$$g'_{\alpha,k}(t)\big(\mu^2 + t\big) \leqslant \alpha g_{\alpha,k}(t), \qquad \big|g''_{\alpha,k}(t)\big|\big(\mu^2 + t\big) \leqslant 3(\alpha + 1) g'_{\alpha,k}(t). \tag{5.4}$$

We shall also deal with the following family of functions:

$$W_{\alpha,k}(t) := 2g'_{\alpha,k}(t)t + g_{\alpha,k}(t), \quad t, \alpha \geqslant 0,\ k \in \mathbb{N}. \tag{5.5}$$

Using the first inequality in (5.4) and then the first in (5.2), together with the fact that $g'_{\alpha,k}(t) \geqslant 0$, we find

$$g_{\alpha,k}(t) \leqslant W_{\alpha,k}(t) \leqslant (2\alpha+1)g_{\alpha,k}(t) \leqslant (2\alpha+1)k. \tag{5.6}$$

Moreover, taking the second estimate in (5.4) into account, and then again the first estimate in (5.4), we also find

$$\left|W'_{\alpha,k}(t)\right|t \leqslant \left|W'_{\alpha,k}(t)\right|\left(\mu^2+t\right) \leqslant 3(\alpha+1)W_{\alpha,k}(t). \tag{5.7}$$

Using that $g'_{\alpha,k}(t) \leqslant g'_{\alpha,k+1}(t)$ for every k, α and t, taking the second inequality in (5.2) into account we have

$$W_{\alpha,k}(t) \leqslant W_{\alpha,k+1}(t) \quad \text{for all } k \in \mathbb{N}. \tag{5.8}$$

Finally, by (5.3) it follows that

$$\begin{aligned}\left(\mu^2+t\right)^{\alpha} \leqslant \lim_{k\to\infty} W_{\alpha,k}(t) &= \left(\mu^2+t\right)^{\alpha-1}\left[2\alpha t + \left(\mu^2+t\right)\right] \\ &\leqslant 3(\alpha+1)\left(\mu^2+t\right)^{\alpha}.\end{aligned} \tag{5.9}$$

5.2. The horizontal Caccioppoli inequality

Here we prove a suitable energy estimate involving powers of the natural quantity $(\mu^2+|\mathfrak{X}u|^2)^{1/2}$, that is "the weight" of Eq. (1.5).

Lemma 5.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Let $\sigma \geqslant 2$ and assume that*

$$\mathfrak{X}u \in L^{p+\sigma}_{\mathrm{loc}}\left(\Omega, \mathbb{R}^{2n}\right), \quad \textit{and} \quad |\mathfrak{X}u|^{p-2+\sigma}|Tu|^2 \in L^1_{\mathrm{loc}}(\Omega). \tag{5.10}$$

Then for all $\eta \in C^\infty_c(\Omega)$, we have

$$\begin{aligned}&\int_\Omega \eta^2\left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p-2}{2}}\sum_{s=1}^{2n}\left(\mu^2+|X_su|^2\right)^{\frac{\sigma}{2}}|\mathfrak{X}X_su|^2\,dx \\ &\quad\leqslant c(\sigma+1)\int_\Omega\left(|\mathfrak{X}\eta|^2+\eta|T\eta|\right)\left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p+\sigma}{2}}\,dx \\ &\qquad + c(\sigma+1)^3\int_\Omega \eta^2\left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p-2+\sigma}{2}}|Tu|^2\,dx,\end{aligned} \tag{5.11}$$

and moreover

$$\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{\sigma}{2}}|\mathfrak{X}X_su|^2\,dx$$
$$\leqslant c(\sigma+1)\int_\Omega\big(|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p+\sigma}{2}}\,dx$$
$$+\,c(\sigma+1)^3\int_\Omega\eta^2\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p-2+\sigma}{2}}|Tu|^2\,dx. \tag{5.12}$$

Both in (5.11) *and in* (5.12) *we have* $c\equiv c(n,p,L/\nu)>1$, *and in particular the constant* c *does not depend on* μ, u, *and on the vector field* $a(\cdot)$.

Proof. With the definition in (5.1), in the following we shall abbreviate $g(\cdot)\equiv g_{\sigma/2,k}$, for a fixed $k\in\mathbb{N}$, while, according to (5.5), we shall denote $W(\cdot):=2g'(\cdot)t+g(\cdot)$. For the rest of the proof all the constants denoted by c or the like will depend only on n, p, L/ν, and will be independent of μ, u, k and σ. Any dependence on σ in the following inequalities will be explicitly displayed. We start by applying Lemma 2.7 with the choice $Z=X_s$ for $s\in\{1,\dots,n\}$; for every $\varphi\in C_c^\infty(\Omega)$, and $h\neq 0$ accordingly small, we arrive at

$$\int_\Omega\big\langle D_h^{X_s}a(\mathfrak{X}u),\mathfrak{X}\varphi\big\rangle\,dx=-\int_\Omega a_{n+s}(\mathfrak{X}u)\big(x\mathrm{e}^{hX_s}\big)T\varphi\,dx. \tag{5.13}$$

We test (5.13) with $\varphi\equiv\phi_1:=\eta^2g(|D_h^{X_s}u|^2)D_h^{X_s}u$, for $s\in\{1,\dots,n\}$. By a simple density argument this is an admissible test function in (5.13), since g is bounded, and moreover $Tu\in L^p_{\mathrm{loc}}(\Omega)$. We obtain, for every $i\in\{1,\dots,2n\}$

$$X_i\phi_1=2\eta X_i\eta g\big(\big|D_h^{X_s}u\big|^2\big)D_h^{X_s}u+\eta^2W\big(\big|D_h^{X_s}u\big|^2\big)X_iD_h^{X_s}u$$

and

$$T\phi_1=2\eta T\eta g\big(\big|D_h^{X_s}u\big|^2\big)D_h^{X_s}u+\eta^2W\big(\big|D_h^{X_s}u\big|^2\big)TD_h^{X_s}u.$$

Inserting the last two equalities into (5.13) yields

$$\int_\Omega\eta^2\sum_{i=1}^{2n}D_h^{X_s}a_i(\mathfrak{X}u)X_iD_h^{X_s}uW\big(\big|D_h^{X_s}u\big|^2\big)\,dx$$
$$=-2\int_\Omega\eta\sum_{i=1}^{2n}D_h^{X_s}a_i(\mathfrak{X}u)X_i\eta g\big(\big|D_h^{X_s}u\big|^2\big)D_h^{X_s}u\,dx$$
$$-\,2\int_\Omega\eta T\eta a_{n+s}(\mathfrak{X}u)\big(x\mathrm{e}^{hX_s}\big)g\big(\big|D_h^{X_s}u\big|^2\big)D_h^{X_s}u\,dx$$
$$-\int_\Omega\eta^2a_{n+s}(\mathfrak{X}u)\big(x\mathrm{e}^{hX_s}\big)W\big(\big|D_h^{X_s}u\big|^2\big)TD_h^{X_s}u\,dx. \tag{5.14}$$

As we are dealing with difference quotients in the horizontal directions, the operators $\mathfrak{X}$ and $D_h^{X_s}$ do not commute. Therefore we need to use identity (2.17); this gives, for every $j \in \{1, \dots, 2n\}$

$$\big(D_h^{X_s} X_j u\big)(x) = X_j\big(D_h^{X_s} u\big)(x) + [X_s, X_j]u\big(x\mathrm{e}^{hX_s}\big).$$

Now use Lemma 2.6 with $Z \equiv X_s$, and adopting the related notation in (2.20), we have

$$\begin{aligned} D_h^{X_s} a_i(\mathfrak{X}u)(x) &= \sum_{j=1}^{2n} a_{i,j}^{X_s}(x) D_h^{X_s} X_j u(x) \\ &= \sum_{j=1}^{2n} a_{i,j}^{X_s}(x)\big[X_j D_h^{X_s} u(x) + [X_s, X_j]u\big(x\mathrm{e}^{hX_s}\big)\big] \\ &= \sum_{j=1}^{2n} a_{i,j}^{X_s}(x) X_j D_h^{X_s} u(x) + a_{i,n+s}^{X_s}(x) T u\big(x\mathrm{e}^{hX_s}\big). \end{aligned} \tag{5.15}$$

From now on in every occurrence of the symbol $\sum$ the indexes i, j will run from 1 to $2n$. Joining (5.14) and (5.15) we obtain

$$\begin{aligned} &\int_\Omega \eta^2 \sum_{i,j} a_{i,j}^{X_s}(x) X_j D_h^{X_s} u X_i D_h^{X_s} u W\big(\big|D_h^{X_s} u\big|^2\big)\,dx \\ &\quad = -\int_\Omega \eta^2 \sum_i a_{i,n+s}^{X_s}(x) X_i D_h^{X_s} u T u\big(x\mathrm{e}^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big)\,dx \\ &\qquad - 2\int_\Omega \eta \sum_{i,j} a_{i,j}^{X_s}(x) X_i \eta X_j D_h^{X_s} u g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u\,dx \\ &\qquad - 2\int_\Omega \eta \sum_i a_{i,n+s}^{X_s}(x) X_i \eta T u\big(x\mathrm{e}^{hX_s}\big) g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u\,dx \\ &\qquad -2\int_\Omega \eta T \eta a_{n+s}(\mathfrak{X}u)\big(x\mathrm{e}^{hX_s}\big) g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u\,dx \\ &\qquad - \int_\Omega \eta^2 a_{n+s}(\mathfrak{X}u)\big(x\mathrm{e}^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big) T D_h^{X_s} u\,dx. \end{aligned} \tag{5.16}$$

A completely similar equation, with $Y_s = X_{n+s}$ replacing X_s everywhere in (5.16), can be obtained by testing (5.13) with $\varphi \equiv \phi_2 := \eta^2 g(|D_h^{Y_s} u|^2) D_h^{Y_s} u$. We finally sum up the resulting two equalities over $s = 1, 2, \dots, n$, thereby obtaining

$$
\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^{2n} \sum_{i,j} a_{i,j}^{X_s}(x) X_j D_h^{X_s} u X_i D_h^{X_s} u W\big(\big|D_h^{X_s} u\big|^2\big)\, dx \\
&= -2 \int_\Omega \eta \sum_{s=1}^{2n} \sum_{i,j} a_{i,j}^{X_s}(x) X_i \eta\, X_j D_h^{X_s} u g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u\, dx \\
&\quad - \int_\Omega \eta^2 \sum_{s=1}^{n} \sum_{i} \big(a_{i,n+s}^{X_s}(x) T u\big(x e^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big) X_i D_h^{X_s} u \\
&\quad - a_{i,s}^{Y_s}(x) T u\big(x e^{hY_s}\big) W\big(\big|D_h^{Y_s} u\big|^2\big) X_i D_h^{Y_s} u\big)\, dx \\
&\quad - 2 \int_\Omega \eta \sum_{s=1}^{n} \sum_{i} X_i \eta \big(a_{i,n+s}^{X_s}(x) T u\big(x e^{hX_s}\big) g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u \\
&\quad - a_{i,s}^{Y_s}(x) T u\big(x e^{hY_s}\big) g\big(\big|D_h^{Y_s} u\big|^2\big) D_h^{Y_s} u\big)\, dx \\
&\quad - 2 \int_\Omega \eta T \eta \sum_{s=1}^{n} \big(a_{n+s}(\mathfrak{X} u)\big(x e^{hX_s}\big) g\big(\big|D_h^{X_s} u\big|^2\big) D_h^{X_s} u \\
&\quad - a_s(\mathfrak{X} u)\big(x e^{hY_s}\big) g\big(\big|D_h^{Y_s} u\big|^2\big) D_h^{Y_s} u\big)\, dx \\
&\quad - \int_\Omega \eta^2 \sum_{s=1}^{n} \big(a_{n+s}(\mathfrak{X} u)\big(x e^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big) T D_h^{X_s} u \\
&\quad - a_s(\mathfrak{X} u)\big(x e^{hY_s}\big) W\big(\big|D_h^{Y_s} u\big|^2\big) T D_h^{Y_s} u\big)\, dx \\
&=: I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned} \tag{5.17}
$$

We now proceed estimating the various terms spreading-up from (5.17). To estimate the left-hand side from below we use (2.22) obtaining

$$
\begin{aligned}
&\text{l.h.s. of (5.17)} \\
&\quad \geqslant c^{-1} \int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2 + \big|\mathfrak{X} u(x)\big|^2 + \big|\mathfrak{X} u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s} u\big|^2\big) \big|\mathfrak{X} D_h^{X_s} u\big|^2\, dx,
\end{aligned} \tag{5.18}
$$

with $c \equiv c(n, p, L/\nu) \geqslant 1$. In order to estimate the integrals $I_1, \ldots, I_4$ we use (2.19), (2.21) and Young's inequality, obtaining for $\varepsilon \in (0,1)$ that

$$
\begin{aligned}
|I_1| &\leqslant c \int_\Omega \eta |\mathfrak{X}\eta| \sum_{s=1}^{2n} \big(\mu^2 + \big|\mathfrak{X} u(x)\big|^2 + \big|\mathfrak{X} u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} g\big(\big|D_h^{X_s} u\big|^2\big) \big|D_h^{X_s} u\big| \big|\mathfrak{X} D_h^{X_s} u\big|\, dx \\
&\leqslant \varepsilon \int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2 + \big|\mathfrak{X} u(x)\big|^2 + \big|\mathfrak{X} u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s} u\big|^2\big) \big|\mathfrak{X} D_h^{X_s} u\big|^2\, dx \\
&\quad + c(\varepsilon) \int_\Omega |\mathfrak{X}\eta|^2 \sum_{s=1}^{2n} \big(\mu^2 + \big|\mathfrak{X} u(x)\big|^2 + \big|\mathfrak{X} u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s} u\big|^2\big) \big|D_h^{X_s} u\big|^2\, dx,
\end{aligned}
$$

and, in a similar way

$$
\begin{aligned}
|I_2| &\leqslant c\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2+2\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s}u\big|^2\big) \\
&\quad \times \big|Tu\big(xe^{hX_s}\big)\big|\big|\mathfrak{X}D_h^{X_s}u\big|\,dx \\
&\leqslant \varepsilon\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s}u\big|^2\big)\big|\mathfrak{X}D_h^{X_s}u\big|^2\,dx \\
&\quad + c(\varepsilon)\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} \\
&\quad \times W\big(\big|D_h^{X_s}u\big|^2\big)\big|Tu\big(xe^{hX_s}\big)\big|^2\,dx,
\end{aligned}
$$

$$
\begin{aligned}
|I_3| &\leqslant c\int_\Omega \eta|\mathfrak{X}\eta| \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} \big|Tu\big(xe^{hX_s}\big)\big| \\
&\quad \times g\big(\big|D_h^{X_s}u\big|^2\big)\big|D_h^{X_s}u\big|\,dx \\
&\leqslant c\int_\Omega |\mathfrak{X}\eta|^2 \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s}u\big|^2\big)\big|D_h^{X_s}u\big|^2\,dx \\
&\quad + c\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u(x)\big|^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-2}{2}} W\big(\big|D_h^{X_s}u\big|^2\big)\big|Tu\big(xe^{hX_s}\big)\big|^2\,dx,
\end{aligned}
$$

and finally

$$
|I_4| \leqslant c\int_\Omega \eta|T\eta| \sum_{s=1}^{2n} \big(\mu^2+\big|\mathfrak{X}u\big(xe^{hX_s}\big)\big|^2\big)^{\frac{p-1}{2}} W\big(\big|D_h^{X_s}u\big|^2\big)\big|D_h^{X_s}u\big|\,dx.
$$

The estimation of the last integral I_5 in (5.17) needs slightly more care, and will be done later. We have that $\mathfrak{X}u \in L^p(\Omega, \mathbb{R}^{2n})$ and, by Theorem 3.2 we also have $Tu \in L^p_{\mathrm{loc}}(\Omega)$, while Lemma 3.1 gives $\mathfrak{X}\mathfrak{X}u \in L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n\times 2n})$, therefore, using also (3.7), up to passing to non-relabeled subsequences, we may assume for every $s = 1, \dots, 2n$ that

$$
\begin{aligned}
&\mathfrak{X}u\big(xe^{hX_s}\big) \to \mathfrak{X}u(x) \qquad \text{in } L^p_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \text{ and a.e.,} \\
&Tu\big(xe^{hX_s}\big) \to Tu(x) \qquad \text{in } L^p_{\mathrm{loc}}(\Omega) \text{ and a.e.,} \qquad (5.19) \\
&\mathfrak{X}D_h^{X_s}u(x) \to \mathfrak{X}X_s u(x) \quad \text{in } L^2_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \text{ and a.e.} \qquad (5.20)
\end{aligned}
$$

See also Section 3.2. The convergence statement in (5.20) needs perhaps an explanation; for $i = 1, \dots, 2n$, write $X_i D_h^{X_s}u(x) = D_h^{X_s}X_i u(x) + [X_i, X_s]u(xe^{hX_s})$, according to (2.17). Then, using the result of Lemma 2.2, the fact that $p \geqslant 2$, and the convergence in (5.19), we have that $X_i D_h^{X_s} u(x) \to X_s X_i u(x) + [X_i, X_s]u(x) = X_i X_s u(x)$ locally in $L^2(\Omega)$, and, up to a subsequence, almost everywhere. Therefore (5.20) is completely proved.

Now we want to pass to the limit with $h \to 0$ in (5.17) taking into account the estimates for the integrals $I_1, \dots, I_4$. Absorbing the terms with ε in the l.h.s., applying Fatou's lemma for the resulting l.h.s., and Lebesgue's dominated convergence theorem for the r.h.s. – keep in mind that $W(\cdot)$ is bounded by (5.6) – we obtain

$$\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2 + |\mathfrak{X}u|^2\big)^{\frac{p-2}{2}} W\big(|X_s u|^2\big) |\mathfrak{X} X_s u|^2 \, dx \\
&\quad \leqslant c \int_\Omega \big(|\mathfrak{X}\eta|^2 + \eta |T\eta|\big)\big(\mu^2 + |\mathfrak{X}u|^2\big)^{\frac{p}{2}} \sum_{s=1}^{2n} W\big(|X_s u|^2\big) \, dx \\
&\qquad + c \int_\Omega \eta^2 \big(\mu^2 + |\mathfrak{X}u|^2\big)^{\frac{p-2}{2}} |Tu|^2 \sum_{s=1}^{2n} W\big(|X_s u|^2\big) \, dx \\
&\qquad + \limsup_{h \to 0} \left| \int_\Omega \eta^2 \sum_{s=1}^{n} a_{n+s}(\mathfrak{X}u)\big(x e^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big) T D_h^{X_s} u \, dx \right| \\
&\qquad + \limsup_{h \to 0} \left| \int_\Omega \eta^2 \sum_{s=1}^{n} a_s(\mathfrak{X}u)\big(x e^{hY_s}\big) W\big(\big|D_h^{Y_s} u\big|^2\big) T D_h^{Y_s} u \, dx \right|.
\end{aligned} \tag{5.21}$$

Now we compute and estimate the last two limits, that actually exist, in the previous inequality; we shall concentrate on the second-last one, similar arguments working for the last one. By Lemma 3.1 we know that $\mathfrak{X}Tu \in L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$. Therefore, for every $s \in \{1, \dots, 2n\}$ we have that

$$D_h^{X_s} Tu \to X_s Tu \quad \text{in } L^2_{\mathrm{loc}}(\Omega) \text{ as } h \to 0. \tag{5.22}$$

Using Young's inequality we can bound the term under the integral sign as follows:

$$\begin{aligned}
&\big|a_{n+s}(\mathfrak{X}u)\big(x e^{hX_s}\big) W\big(\big|D_h^{X_s} u\big|^2\big) T D_h^{X_s} u\big| \\
&\quad \leqslant c(\sigma, k)\big(\mu^2 + \big|\mathfrak{X}u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{p-1}{2}} \big|D_h^{X_s} Tu\big| \\
&\quad \leqslant c(\sigma, k)\Big[\big(\mu^2 + \big|\mathfrak{X}u\big(x e^{hX_s}\big)\big|^2\big)^{\frac{2p-2}{2}} + \big|D_h^{X_s} Tu\big|^2\Big],
\end{aligned} \tag{5.23}$$

where we used (5.6) and that $\alpha = \sigma/2$. Since $\sigma \geqslant 2$ then (5.10) implies that $\mathfrak{X}u \in L^{p+2}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$ and moreover $p < 4$ implies that we can use the fact that $2p - 2 < p + 2$. Therefore $\mathfrak{X}u \in L^{2p-2}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$ and hence

$$\mathfrak{X}u\big(x e^{hX_s}\big) \to \mathfrak{X}u(x) \quad \text{in } L^{2p-2}_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \text{ and a.e. as } h \to 0.$$

Thus, thanks to (5.22)–(5.23), we can let $h \to 0$ using a well-known variant of Lebesgue's dominated convergence theorem; therefore we obtain

$$\begin{aligned}
&\lim_{h\to 0}\int_\Omega \eta^2 \sum_{s=1}^n a_{n+s}(\mathfrak{X}u)\big(x e^{hX_s}\big) W\big(\big|D_h^{X_s}u\big|^2\big) T D_h^{X_s} u\, dx \\
&\quad = \int_\Omega \eta^2 \sum_{s=1}^n a_{n+s}(\mathfrak{X}u) W\big(|X_s u|^2\big) X_s T u\, dx. \qquad (5.24)
\end{aligned}$$

In a completely similar manner, we also have

$$\begin{aligned}
&\lim_{h\to 0}\int_\Omega \eta^2 \sum_{s=1}^n a_{s}(\mathfrak{X}u)\big(x e^{hY_s}\big) W\big(\big|D_h^{Y_s}u\big|^2\big) T D_h^{Y_s} u\, dx \\
&\quad = \int_\Omega \eta^2 \sum_{s=1}^n a_{s}(\mathfrak{X}u) W\big(|Y_s u|^2\big) Y_s T u\, dx. \qquad (5.25)
\end{aligned}$$

Connecting (5.24) and (5.25) to (5.21) we get

$$\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^{2n} \big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}} W\big(|X_s u|^2\big)|\mathfrak{X}X_s u|^2\, dx \\
&\quad \leqslant c\int_\Omega \big(|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}} \sum_{s=1}^{2n} W\big(|X_s u|^2\big)\, dx \\
&\qquad + c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}} \sum_{s=1}^{2n} W\big(|X_s u|^2\big)|Tu|^2\, dx \\
&\qquad + \left|\int_\Omega \eta^2 \sum_{s=1}^n a_{n+s}(\mathfrak{X}u) W\big(|X_s u|^2\big) X_s T u\, dx\right| \\
&\qquad + \left|\int_\Omega \eta^2 \sum_{s=1}^n a_{s}(\mathfrak{X}u) W\big(|Y_s u|^2\big) Y_s T u\, dx\right|, \qquad (5.26)
\end{aligned}$$

with $c \equiv c(n, p, L/\nu)$. We continue estimating the last two integrals; we shall estimate the first one, the estimation of the latter being completely analogous. We integrate by parts as follows:

$$\begin{aligned}
&\int_\Omega \eta^2 \sum_{s=1}^n a_{n+s}(\mathfrak{X}u) W\big(|X_s u|^2\big) X_s T u\, dx \\
&\quad = -2\int_\Omega \eta T u \sum_{s=1}^n X_s\eta\, a_{n+s}(\mathfrak{X}u) W\big(|X_s u|^2\big)\, dx
\end{aligned}$$

$$
\begin{aligned}
&-\int_\Omega \eta^2 Tu \sum_{s=1}^{n}\sum_{\alpha=1}^{2n} D_{z_\alpha} a_{n+s}(\mathfrak{X}u) X_s X_\alpha u W\big(|X_s u|^2\big)\,dx \\
&-2\int_\Omega \eta^2 Tu \sum_{s=1}^{n} a_{n+s}(\mathfrak{X}u) W'\big(|X_s u|^2\big) X_s u X_s X_s u\,dx \\
&=: A + B + C. \qquad (5.27)
\end{aligned}
$$

The previous integration by parts needs of course to be justified; we postpone its verification to the very end of the proof. The estimates for A, B, C follow again by (2.21), (2.22) and Young's inequality; indeed, as for A we have

$$
\begin{aligned}
|A| &\leqslant 2\int_\Omega \eta|\mathfrak{X}\eta|\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-1}{2}}|Tu|\sum_{s=1}^{n} W\big(|X_s u|^2\big)\,dx \\
&\leqslant c\int_\Omega |\mathfrak{X}\eta|^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^{n} W\big(|X_s u|^2\big)\,dx \\
&\quad + c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{n} W\big(|X_s u|^2\big)|Tu|^2\,dx.
\end{aligned}
$$

Using that $X_s X_\alpha = X_\alpha X_s + [X_s, X_\alpha]$, we have, with $\varepsilon \in (0,1)$

$$
\begin{aligned}
|B| &\leqslant \left|\int_\Omega \eta^2 Tu \sum_{s=1}^{n}\sum_{\alpha=1}^{2n} D_{z_\alpha} a_{n+s}(\mathfrak{X}u) X_\alpha X_s u W\big(|X_s u|^2\big)\,dx\right| \\
&\quad + \left|\int_\Omega \eta^2 |Tu|^2 \sum_{s=1}^{n} D_{z_{n+s}} a_{n+s}(\mathfrak{X}u) W\big(|X_s u|^2\big)\,dx\right| \\
&\leqslant c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{n} W\big(|X_s u|^2\big)\big(|Tu||\mathfrak{X}X_s u|+|Tu|^2\big)\,dx \\
&\leqslant \varepsilon\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{n} W\big(|X_s u|^2\big)|\mathfrak{X}X_s u|^2\,dx \\
&\quad + c(\varepsilon)\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{n} W\big(|X_s u|^2\big)|Tu|^2\,dx.
\end{aligned}
$$

Finally, using (5.7) we have

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

Joining together the estimates for A, B, C, we obtain

$$
\begin{aligned}
&\left|\int_\Omega \eta^2\sum_{s=1}^n a_{n+s}(\mathfrak{X}u)W\big(|X_su|^2\big)X_sTu\,dx\right|\\
&\quad\leqslant c\varepsilon\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^n W\big(|X_su|^2\big)|\mathfrak{X}X_su|^2\,dx\\
&\qquad+c\int_\Omega |\mathfrak{X}\eta|^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^n W\big(|X_su|^2\big)\,dx\\
&\qquad+c(\varepsilon)\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^n W\big(|X_su|^2\big)|Tu|^2\,dx\\
&\qquad+c(\varepsilon)(\sigma+1)^2\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^n \frac{W(|X_su|^2)}{(\mu^2+|X_su|^2)}|Tu|^2\,dx,
\end{aligned}
\tag{5.28}
$$

where $c\equiv c(n,p,L/\nu)$. A completely analogous estimate, replacing on the right-hand side of (5.28) X_s by Y_s, holds also for the term

$$
\int_\Omega \eta^2\sum_{s=1}^n a_s(\mathfrak{X}u)W\big(|Y_su|^2\big)Y_sTu\,dx,
$$

appearing in (5.26). Therefore using (5.28), and its Y_s-analog, to estimate (5.26), absorbing terms with ε on the left-hand side, we finally obtain

$$\begin{aligned}
&\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{2n} W\big(|X_su|^2\big)|\mathfrak{X}X_su|^2\,dx\\
&\quad\leqslant c\int_\Omega \big(|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^{2n} W\big(|X_su|^2\big)\,dx\\
&\qquad + c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{2n} W\big(|X_su|^2\big)|Tu|^2\,dx\\
&\qquad + c(\sigma+1)^2\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^{2n} \frac{W(|X_su|^2)}{(\mu^2+|X_su|^2)}|Tu|^2\,dx,
\end{aligned}$$

where c only depends on n, p, L/ν, but is otherwise independent of μ, σ, k, of the solution u, and of the vector field $a(\cdot)$. Letting $k \nearrow \infty$ in the previous inequality, using (5.8)–(5.9) to apply the monotone convergence theorem, and finally using the elementary inequalities

$$\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{\sigma}{2}} \leqslant c(n,p)\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p+\sigma}{2}},$$

$$\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{\sigma}{2}} \leqslant c(n,p)\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p-2+\sigma}{2}},$$

and, since $\sigma \geqslant 2$ by assumption,

$$\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{\sigma-2}{2}} \leqslant c(n,p)\sum_{s=1}^{2n}\big(\mu^2+|X_su|^2\big)^{\frac{p-2+\sigma}{2}},$$

we get (5.12), from which also (5.11) immediately follows. It remains to give the

Justification of (5.27). Fix $s \in \{1,\dots,n\}$; assume that

$$X_s\big(\eta^2 a_{n+s}(\mathfrak{X}u)W\big(|X_su|^2\big)Tu\big) \in L^1_{\mathrm{loc}}(\Omega) \tag{5.29}$$

and that the identity

$$\begin{aligned}
X_s\big(\eta^2 a_{n+s}(\mathfrak{X}u)W\big(|X_su|^2\big)Tu\big) &= \big(X_s\eta^2\big)a_{n+s}(\mathfrak{X}u)W\big(|X_su|^2\big)Tu\\
&\quad + \eta^2\sum_{j=1}^{2n} D_{z_j}a_{n+s}(\mathfrak{X}u)X_sX_juW\big(|X_su|^2\big)Tu\\
&\quad + 2\eta^2 a_{n+s}(\mathfrak{X}u)W'\big(|X_su|^2\big)X_suX_sX_suTu\\
&\quad + \eta^2 a_{n+s}(\mathfrak{X}u)W\big(|X_su|^2\big)X_sTu\\
&=: B_1+B_2+B_3+B_4,
\end{aligned} \tag{5.30}$$

holds in the distributional sense, with $B_1, \dots, B_4 \in L^1_{\mathrm{loc}}(\Omega)$. Then, since η has compact support in Ω, we have that

$$\int_\Omega X_s\big(\eta^2 a_{n+s}(\mathfrak{X}u)W\big(|X_s u|^2\big)Tu\big)\,dx = 0,$$

from which (5.27) follows via (5.30). In turn it remains to establish the validity of (5.29)–(5.30). We shall repeatedly use Lemma 2.8; we start observing that by (1.2) and $Tu \in L^p_{\mathrm{loc}}(\Omega)$, Young's inequality gives that $a_{n+s}(\mathfrak{X}u)W(|X_s u|^2)Tu \in L^1_{\mathrm{loc}}(\Omega)$. We are of course using that $W(\cdot)$ is bounded. The same argument gives that $B_1 \in L^1_{\mathrm{loc}}(\Omega)$. Next we have

$$|B_2| \leqslant c(k,\sigma)\big[\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}\mathfrak{X}u|^2 + \mu^p + |\mathfrak{X}u|^p + |Tu|^p\big],$$

and observe that the right-hand side belongs to $L^1_{\mathrm{loc}}(\Omega)$ by (3.9), therefore $B_2 \in L^1_{\mathrm{loc}}(\Omega)$. Then, by (1.2), (5.7) and Young's inequality we have

$$\begin{aligned} |B_3| &\leqslant \frac{c(k,\sigma)|X_s u|}{\mu^2+|X_s u|^2}\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-1}{2}}|\mathfrak{X}\mathfrak{X}u||Tu| \\ &\leqslant c(k,\sigma,\mu)\big[\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}\mathfrak{X}u|^2 + \big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}|Tu|^2\big] \end{aligned}$$

and observe that all the quantities in the right hand side belong to $L^1_{\mathrm{loc}}(\Omega)$ by (3.9) and (5.10), since here we are assuming $\sigma \geqslant 2$. We again conclude that $B_3 \in L^1_{\mathrm{loc}}(\Omega)$. Finally, again by (1.2) we have that

$$|B_4| \leqslant c(k,\sigma)\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big[|\mathfrak{X}Tu|^2 + \big(\mu^2+|\mathfrak{X}u|^2\big)\big],$$

and again, $B_4 \in L^1_{\mathrm{loc}}(\Omega)$ follows from (3.9). At this stage we can apply Lemma 2.8 to the product $a_{n+s}(\mathfrak{X}u)W(|X_s u|^2)Tu \in L^1_{\mathrm{loc}}(\Omega)$ concluding that (5.29)–(5.30) hold. □

5.3. The vertical Caccioppoli inequality

We now state the energy estimate involving Tu. Its proof is considerably simpler and it is close to similar estimates in the Euclidean case, since the operators T and $\mathfrak{X}$ commute. We report the full proof for the sake of completeness.

Lemma 5.2. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Let $\sigma \geqslant 0$ and assume that*

$$\mathfrak{X}u \in L^{p+2+\sigma}_{\mathrm{loc}}\big(\Omega,\mathbb{R}^{2n}\big), \quad \textit{and} \quad Tu \in L^{\frac{p+2+\sigma}{2}}_{\mathrm{loc}}(\Omega). \tag{5.31}$$

Then we have

$$\int_\Omega \big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^{\frac{\sigma}{2}}|\mathfrak{X}Tu|^2\eta^2\,dx \leqslant c\int_\Omega \big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^{\frac{\sigma+4}{2}}|\mathfrak{X}\eta|^2\,dx, \tag{5.32}$$

for all $\eta \in C_c^\infty(\Omega)$, where the constant $c \equiv c(n, p, L/\nu)$, is independent of μ, of the solution u, and of the vector field $a(\cdot)$.

Proof. We again start by applying Lemma 2.7, this time with the choice $Z = T$; for every $\varphi \in C_c^\infty(\Omega)$, and $h \neq 0$ accordingly small, we arrive at

$$\int_\Omega \langle D_h^T a(\mathfrak{X}u), \mathfrak{X}\varphi\rangle\, dx = 0. \tag{5.33}$$

Observe that we have used that $[T, X_i] = 0$ for every $i = 1, \dots, 2n$. As a test function in (5.33) we choose $\varphi = \eta^2 |D_h^T u|^{\frac{\sigma}{2}} D_h^T u$. Note that this is an admissible test function in view of the fact that u is locally bounded, see Theorem 3.1. Since $[T, X_s] = 0$ for any $s = 1, \dots, 2n$, we have $\mathfrak{X}(D_h^T u) = D_h^T(\mathfrak{X}u)$ by Lemma 2.3. Inserting φ into (5.33) we find

$$\begin{aligned}(1+\sigma/2)\int_\Omega \eta^2 \sum_{i=1}^{2n} D_h^T a_i(\mathfrak{X}u) X_i D_h^T u \big|D_h^T u\big|^{\frac{\sigma}{2}}\, dx \\ = -2\int_\Omega \eta \sum_{i=1}^{2n} D_h^T a_i(\mathfrak{X}u) X_i \eta \big|D_h^T u\big|^{\frac{\sigma}{2}} D_h^T u\, dx. \end{aligned}\tag{5.34}$$

Using (2.19) and (2.22) with $Z \equiv X$, we can estimate the l.h.s. of (5.34) from below

$$\text{l.h.s. of (5.34)} \geqslant c^{-1}\int_\Omega \big(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2\big)^{\frac{p-2}{2}} \big|D_h^T u\big|^{\frac{\sigma}{2}} \big|\mathfrak{X}D_h^T u\big|^2 \eta^2\, dx,$$

where $c \equiv c(n, p, L/\nu) \geqslant 1$. For the r.h.s. of (5.34) we use again (2.19) together with (2.21) and Young's inequality obtaining, with $\varepsilon \in (0, 1)$

$$\begin{aligned}\big|\text{r.h.s. of (5.34)}\big| \leqslant{}& \varepsilon \int_\Omega \big(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2\big)^{\frac{p-2}{2}} \big|D_h^T u\big|^{\frac{\sigma}{2}} \big|\mathfrak{X}D_h^T u\big|^2 \eta^2\, dx \\ &+ c(\varepsilon)\int_\Omega \big(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2\big)^{\frac{p-2}{2}} \big|D_h^T u\big|^{\frac{\sigma}{2}+2} |\mathfrak{X}\eta|^2\, dx.\end{aligned}$$

Combining these estimates and choosing ε suitably small as usual, we arrive at the following Caccioppoli-type estimate:

$$\begin{aligned} I_h &:= \int_\Omega \big(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2\big)^{\frac{p-2}{2}} \big|D_h^T u\big|^{\frac{\sigma}{2}} \big|\mathfrak{X}D_h^T u\big|^2 \eta^2\, dx \\ &\leqslant \tilde{c}\int_\Omega \big(\mu^2 + |\mathfrak{X}u(x)|^2 + |\mathfrak{X}u(xe^{hT})|^2\big)^{\frac{p-2}{2}} \big|D_h^T u\big|^{\frac{\sigma}{2}+2} |\mathfrak{X}\eta|^2\, dx =: II_h \end{aligned}\tag{5.35}$$

which is obviously valid for any $h>0$ such that $\sqrt{h}=|\mathrm{e}^{hT}|_{\mathrm{CC}}<\operatorname{dist}(\operatorname{supp}\eta,\partial\Omega)$; here $\tilde{c}$ depends on n, p, L/ν. Using Young's inequality to estimate the r.h.s. of (5.35) we finally obtain

$$II_h \leqslant c \int\limits_{\operatorname{supp}\eta} \left(\mu^2+\left|\mathfrak{X}u(x)\right|^2+\left|\mathfrak{X}u\left(x\mathrm{e}^{hT}\right)\right|^2\right)^{\frac{p+2+\sigma}{2}} dx + c \int\limits_{\operatorname{supp}\eta} \left|D_h^T u\right|^{\frac{p+2+\sigma}{2}} dx, \tag{5.36}$$

with $c\equiv c(\|\mathfrak{X}\eta\|_{L^\infty})$. Since both Tu and $\mathfrak{X}u$ exist and satisfy (5.31), by Lemma 2.2, (5.36), and a well-known variant of Lebesgue's dominated convergence theorem, we obtain that

$$\lim_{h\to 0} II_h = \tilde{c}\int\limits_{\Omega} \left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} |Tu|^{\frac{\sigma+4}{2}} |\mathfrak{X}\eta|^2\, dx. \tag{5.37}$$

On the other hand, by Lemma 3.1 and using and Fatou's lemma we have that

$$\int\limits_{\Omega} \left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} |Tu|^{\frac{\sigma}{2}} |\mathfrak{X}Tu|^2\eta^2\, dx \leqslant \liminf_{h\to 0} I_h. \tag{5.38}$$

The proof of (5.32) now follows combining (5.37)–(5.38) with (5.35). □

6. Intermediate integrability

The aim of this section is to improve the already found higher integrability result in (4.9). Indeed the main result here is

Lemma 6.1. *Let $u\in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2\leqslant p<4$. Then*

$$\mathfrak{X}u\in L^{p+4}_{\mathrm{loc}}\left(\Omega,\mathbb{R}^{2n}\right). \tag{6.1}$$

Moreover, for every couple of open subsets $\Omega'\Subset\Omega''\Subset\Omega$ there exists a constant c depending only on n, p, L/ν, $\operatorname{dist}(\Omega',\partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that

$$\int\limits_{\Omega'} |\mathfrak{X}u|^{p+4}\, dx \leqslant c\int\limits_{\Omega''} \left(|\mathfrak{X}u|^p+|Tu|^p+\mu^p\right) dx. \tag{6.2}$$

The key to the previous lemma is in fact the following one, whose proof features a rather unorthodox choice of the test function φ in (2.11) – see (6.4) below.

Lemma 6.2. *Let $u\in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2\leqslant p<4$. Then*

$$\left(\mu^2+|\mathfrak{X}u|^2\right)^{\frac{p}{2}}|Tu|^2\in L^1_{\mathrm{loc}}(\Omega).$$

Moreover, for all $\eta\in C_c^\infty(\Omega)$, we have

$$\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p}{2}}|Tu|^2\,dx$$
$$\leqslant c\big(1+\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\big)\int_\Omega \big(\eta^2+|\mathfrak{X}\eta|^2\big)\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^2\,dx, \tag{6.3}$$

where $c\equiv c(n,p)>0$.

Proof. In the following we shall denote $\mathcal{T}_k(t):=\min\{t,k\}$, for $t\geqslant 0$ and $k\in\mathbb{N}$, slightly adjusting the definition already given in (4.2). Set

$$\varphi := \big(\mathcal{T}_k\big(|Tu|\big)\big)^2\eta^2 u, \tag{6.4}$$

for $k>0$; we wish to take φ as a test function in (2.11). We first observe that the function $t\mapsto(\mathcal{T}_k(|t|))^2$ is Lipschitz continuous and therefore, since $Tu\in HW^{1,2}(\Omega)$ then by the chain rule in the Heisenberg group – see [10] – it also follows that $(\mathcal{T}_k(|Tu|))^2\in HW^{1,2}(\Omega)$. Then, since $u\in HW^{1,p}(\Omega)\cap L^\infty_{\mathrm{loc}}(\Omega)$ a standard difference quotients argument, as for instance the one in Lemma 2.8, finally gives that $\varphi\in HW^{1,2}_0(\Omega)$. Now recall that in Lemma 4.3, we already showed that $\mathfrak{X}u\in L^{p+2}_{\mathrm{loc}}(\Omega,\mathbb{R}^{2n})$. So by a standard approximation argument, we can easily show that any function from $HW^{1,(p+2)/3}_0(\Omega)$ is an admissible in (2.11). Thus φ as defined in (6.4) is admissible test function, since $(p+2)/3<2$. Recall here that we are assuming $p<4$. Therefore, using φ in (2.11), we obtain

$$\int_\Omega \eta^2\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\big\langle a(\mathfrak{X}u),\mathfrak{X}u\big\rangle\,dx = -2\int_\Omega \eta u\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\big\langle a(\mathfrak{X}u),\mathfrak{X}\eta\big\rangle\,dx$$
$$-\int_\Omega \eta^2 u\big\langle a(\mathfrak{X}u),\mathfrak{X}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\big\rangle\,dx.$$

In turn, using (1.2) and (3.16) the previous equality yields

$$\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}u|^2\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx$$
$$\leqslant c\int_\Omega \eta|\mathfrak{X}\eta||u|\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-1}{2}}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx$$
$$+c\int_\Omega \eta^2|u|\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-1}{2}}\big|\mathfrak{X}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\big|\,dx$$
$$+c\int_\Omega \eta^2\mu^p\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx =: D+E+F, \tag{6.5}$$

with $c\equiv c(n,p,L/\nu)$. We use Young's inequality to estimate D as follows:

$$
\begin{aligned}
D \leqslant{}& \frac{1}{4}\int_{\Omega} \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}u|^2\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx \\
&+ c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_{\Omega} |\mathfrak{X}\eta|^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^2\,dx \\
&+ \int_{\Omega} \eta^2\mu^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx.
\end{aligned}
$$

We estimate E by Young's inequality and Lemma 5.2 with $\sigma = 0$, that is

$$
\begin{aligned}
E \leqslant{}& \frac{1}{4}\int_{\Omega} \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}u|^2\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx \\
&+ c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_{\Omega} \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}Tu|^2\,dx \\
&+ \int_{\Omega} \eta^2\mu^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx \\
\overset{(5.32)}{\leqslant}{}& \frac{1}{4}\int_{\Omega} \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|\mathfrak{X}u|^2\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx \\
&+ c\|u\|^2_{L^\infty(\operatorname{supp}\eta)}\int_{\Omega} |\mathfrak{X}\eta|^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^2\,dx \\
&+ \int_{\Omega} \eta^2\mu^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}\big(\mathcal{T}_k\big(|Tu|\big)\big)^2\,dx.
\end{aligned}
$$

Finally, since $\mu \leqslant 1$ we have

$$
F \leqslant c\int_{\Omega} \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^2\,dx.
$$

Plugging the above estimates for D, E and F into (6.5), and eventually letting $k \nearrow \infty$, we obtain (6.3), using that $\mu \leqslant 1$. This completes the proof of the lemma. $\square$

Proof. The proof of (6.1) follows combining Lemma 6.2, Lemma 5.1 with $\sigma = 2$, Lemma 4.3, and finally Lemma 4.2 again with $\sigma = 2$. Accordingly, the proof of (6.2) follows combining all the a priori estimates of the used lemmata, taking into account the fact that everywhere Ω', Ω'' and η can be chosen arbitrarily. Moreover, the right-hand side of (6.3) has to be estimated by means of Young's inequality, as follows:

$$
\int_{\operatorname{supp}\eta} \big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^2\,dx \leqslant c\int_{\operatorname{supp}\eta} \big(|\mathfrak{X}u|^p+|Tu|^p+\mu^p\big)\,dx. \qquad \square
$$

7. Iteration and higher integrability

The main result of this section is the following:

Proposition 7.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Then it holds that*

$$\mathfrak{X}u \in L^q_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \quad \text{and} \quad Tu \in L^q_{\mathrm{loc}}(\Omega) \quad \text{for every } q < \infty. \tag{7.1}$$

Moreover, for every $q < \infty$ there exists a constant c, depending on n, p, L/ν, and q, but otherwise independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that the following reverse-Hölder type inequalities hold for any CC-ball $B_R \subset \Omega$:

$$\bigg(\fint_{B_{R/2}} |\mathfrak{X}u|^q \, dx\bigg)^{1/q} \leqslant c\bigg(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p \, dx\bigg)^{1/p}, \tag{7.2}$$

and

$$\bigg(\fint_{B_{R/2}} |Tu|^q \, dx\bigg)^{1/q} \leqslant \frac{c}{R}\bigg(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p \, dx\bigg)^{1/p}. \tag{7.3}$$

In order to prove the previous result we need a few preliminary lemmata. Their iterated use will finally lead to the proof of Proposition 7.1.

Lemma 7.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Assume that*

$$\mathfrak{X}u \in L^{p+\sigma}_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big), \qquad |\mathfrak{X}u|^{p-2+\sigma}|Tu|^2 \in L^1_{\mathrm{loc}}(\Omega), \quad \text{and} \quad Tu \in L^{\frac{p+2+\sigma}{2}}_{\mathrm{loc}}(\Omega), \tag{7.4}$$

for some $\sigma \geqslant 2$. Then

$$\mathfrak{X}u \in L^{p+2+\sigma}_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big). \tag{7.5}$$

Moreover, for every couple of open subsets $\Omega' \Subset \Omega'' \Subset \Omega$ there exists a constant c depending only on $n, p, L/\nu, \sigma$, $\mathrm{dist}(\Omega', \partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent on μ, such that

$$\int_{\Omega'} |\mathfrak{X}u|^{p+2+\sigma} \, dx \leqslant c \int_{\Omega''} \big(|\mathfrak{X}u|^{p+\sigma} + |Tu|^{\frac{p+2+\sigma}{2}} + \mu^p\big) \, dx. \tag{7.6}$$

Proof. By (7.4) we can use Lemma 5.1; therefore combining (5.11) with (4.5), by means of a standard covering argument we deduce the validity of (7.5). Once (7.5) holds we use Young's inequality to estimate the last integral in the right hand side of (5.11) as follows:

$$\int_\Omega \eta^2\bigl(\mu^2+|\mathfrak{X}u|^2\bigr)^{\frac{p-2+\sigma}{2}}|Tu|^2\,dx \leqslant \varepsilon\int_\Omega \eta^2\bigl(\mu^2+|\mathfrak{X}u|^2\bigr)^{\frac{p+2+\sigma}{2}}\,dx + c(\varepsilon)\int_\Omega \eta^2|Tu|^{\frac{p+2+\sigma}{2}}\,dx, \tag{7.7}$$

where $\varepsilon \in (0,1)$; note that the intermediate integral in (7.7) is now finite. Connecting the previous inequality to (5.11) and eventually to (4.5), and choosing ε small enough, but depending only on n, p, L/ν, σ and $\|u\|_{L^\infty(\operatorname{supp}\eta)}$, in order to re-absorb the intermediate integral appearing in (7.7) in the left-hand side of (4.5), we gain, after a few elementary manipulations

$$\int_\Omega \eta^2\bigl(\mu^2+|\mathfrak{X}u|^2\bigr)^{\frac{p+2+\sigma}{2}}\,dx \leqslant c\int_{\operatorname{supp}\eta} \bigl(\mu^2+|\mathfrak{X}u|^2\bigr)^{\frac{p+\sigma}{2}}\,dx + c\int_{\operatorname{supp}\eta} |Tu|^{\frac{p+2+\sigma}{2}}\,dx.$$

The constant c in the last inequality depends only on the data n, p, L/ν, σ, and on the norms $\|\mathfrak{X}\eta\|_{L^\infty}$, $\|T\eta\|_{L^\infty}$, $\|u\|_{L^\infty(\operatorname{supp}\eta)}$, but is otherwise independent of the solution u, of the vector field $a(\cdot)$, and of μ. Note that we have used that $\mu \leqslant 1$. At this stage the inequality in (7.6) follows by the previous inequality via a standard covering argument involving a suitable choice of the cut-off function η; again we are using that $\mu \leqslant 1$. □

Lemma 7.2. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with $2 \leqslant p < 4$. Assume that*

$$\mathfrak{X}u \in L^{p+2+\sigma}_{\mathrm{loc}}\bigl(\Omega,\mathbb{R}^{2n}\bigr) \quad \textit{and} \quad Tu \in L^{\frac{p+2+\sigma}{2}}_{\mathrm{loc}}(\Omega),$$

for some $\sigma \geqslant 0$, then

$$Tu \in L^{\frac{p+3+\sigma}{2}}_{\mathrm{loc}}(\Omega). \tag{7.8}$$

Moreover, for every couple of open subsets $\Omega' \Subset \Omega'' \Subset \Omega$ there exists a constant c depending only on n, p, L/ν, σ, $\operatorname{dist}(\Omega', \partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent on μ, of the solution u, and of the vector field $a(\cdot)$, such that

$$\int_{\Omega'} |Tu|^{\frac{p+3+\sigma}{2}}\,dx \leqslant c\int_{\Omega''} \bigl(|\mathfrak{X}u|^{p+2+\sigma} + |Tu|^{\frac{p+2+\sigma}{2}} + \mu^p\bigr)\,dx \tag{7.9}$$

holds.

Proof. In the following we shall again denote $\mathcal{T}_k(t) := \min\{t,k\}$ for $t \geqslant 0$ and $k \in \mathbb{N}$. Let $\eta \in C_c^\infty(\Omega)$ be as usual a cut-off function such that $0 \leqslant \eta \leqslant 1$. Using that $T = [X_i, Y_i] = X_iY_i - Y_iX_i$, we start by integrating by parts as follows:

$$
\begin{aligned}
&\int_\Omega \eta^2 |Tu|^2 \mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big)\,dx \\
&\quad = \int_\Omega \eta^2 (X_1Y_1 - Y_1X_1)u\,Tu\,\mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big)\,dx \\
&\quad \leqslant 4\int_\Omega \eta|\mathfrak{X}\eta||\mathfrak{X}u||Tu|\mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big)\,dx + c\int_\Omega \eta^2|\mathfrak{X}u||\mathfrak{X}Tu|\mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big)\,dx \\
&\quad =: P_4 + P_5,
\end{aligned}
\tag{7.10}
$$

where $c = c(p,\sigma) > 0$. Note that the previous integration by parts is legal since

$$
Tu\,\mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big) \in HW^{1,2}_{\mathrm{loc}}(\Omega). \tag{7.11}
$$

This fact follows by chain rule in the Heisenberg group – see [10] – since by the very definition of $\mathcal{T}_k$ it follows that the function $t \mapsto t\mathcal{T}_k(|t|^{\frac{p-1+\sigma}{2}})$ is globally Lipschitz continuous on $\mathbb{R}$, together with the fact that $Tu \in HW^{1,2}_{\mathrm{loc}}(\Omega)$ – see (3.9).

Now, by Young's inequality, we have for the integral P_4

$$
P_4 \leqslant 4\int_\Omega |\mathfrak{X}\eta||\mathfrak{X}u|^{p+2+\sigma}\,dx + 4\int_\Omega |\mathfrak{X}\eta||Tu|^{\frac{p+2+\sigma}{2}}\,dx.
$$

We now come to P_5; using repeatedly Young's inequality and once inequality (5.32) from Lemma 5.2 we have

$$
\begin{aligned}
P_5 &\leqslant c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac12}|Tu|^{\frac{p-1+\sigma}{2}}|\mathfrak{X}Tu|\,dx \\
&\leqslant c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^{\frac{\sigma}{2}}|\mathfrak{X}Tu|^2\,dx \\
&\qquad\quad + c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{4-p}{2}}|Tu|^{\frac{2p-2+\sigma}{2}}\,dx \\
&\overset{(5.32)}{\leqslant} c\int_\Omega |\mathfrak{X}\eta|^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{p-2}{2}}|Tu|^{\frac{\sigma+4}{2}}\,dx \\
&\qquad + c\int_\Omega \eta^2\big(\mu^2+|\mathfrak{X}u|^2\big)^{\frac{4-p}{2}}|Tu|^{\frac{2p-2+\sigma}{2}}\,dx \\
&\leqslant c\int_\Omega \big(\eta^2+|\mathfrak{X}\eta|^2\big)\big(\mu^{p+2+\sigma}+|\mathfrak{X}u|^{p+2+\sigma}+|Tu|^{\frac{p+2+\sigma}{2}}\big)\,dx.
\end{aligned}
\tag{7.12}
$$

Note how the crucial assumption $p < 4$ hereby comes into the play once again. Using the estimates found for P_4, P_5, inequality (7.10) becomes

$$\int_\Omega \eta^2|Tu|^2\mathcal{T}_k\big(|Tu|^{\frac{p-1+\sigma}{2}}\big)\,dx$$

$$\leqslant c\int_\Omega \big(\eta^2+|\mathfrak{X}\eta|+|\mathfrak{X}\eta|^2\big)\big(\mu^{p+2+\sigma}+|\mathfrak{X}u|^{p+2+\sigma}+|Tu|^{\frac{p+2+\sigma}{2}}\big)\,dx.$$

The constant c in the last inequality depends only on n, p, σ. Letting $k \nearrow \infty$ and using the fact that $\mu \leqslant 1$, we have

$$\int_\Omega \eta^2|Tu|^{\frac{p+3+\sigma}{2}}\,dx \leqslant c\int_\Omega \big(\eta^2+|\mathfrak{X}\eta|+|\mathfrak{X}\eta|^2\big)\big(\mu^{p}+|\mathfrak{X}u|^{p+2+\sigma}+|Tu|^{\frac{p+2+\sigma}{2}}\big)\,dx.$$

Then (7.8) follows by a standard covering argument since the choice of η is arbitrary in the previous inequality. In the same way, (7.9) follows via a standard covering argument involving a suitable choice of η. □

Proof of Proposition 7.1. The proof is divided in two steps: first we prove the qualitative result in (7.1) with a first form of the main priori estimates, that is (7.13) below. Then, in a second step, we show how to get the explicit form of the a priori estimates in (7.2)–(7.3) from (7.13) by means of a “blow-up” argument.

Step 1: Iteration and higher integrability. Here we prove (7.1) and that, for every couple of open subsets $\Omega' \Subset \Omega'' \Subset \Omega$, and $q < \infty$, there exists a constant c depending only on n, p, L/ν, q, $\operatorname{dist}(\Omega', \partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that

$$\int_{\Omega'} \big(|\mathfrak{X}u|^q+|Tu|^q\big)\,dx \leqslant c\int_{\Omega''}\big(|\mathfrak{X}u|^p+1\big)\,dx. \tag{7.13}$$

For this, let us define the sequence

$$\begin{cases}\sigma_{k+1} := \sigma_k + \frac{4}{p+3+\sigma_k},\\ \sigma_0 := 2.\end{cases} \tag{7.14}$$

It is easy to see that $\{\sigma_k\}$ is a strictly increasing sequence such that $\sigma_k \nearrow \infty$. We shall prove by induction that

$$\mathfrak{X}u \in L^{p+2+\sigma_k}_{\mathrm{loc}}\big(\Omega,\mathbb{R}^{2n}\big) \quad \text{and} \quad Tu \in L^{\frac{p+2+\sigma_k}{2}}_{\mathrm{loc}}(\Omega), \tag*{$(\mathbf{A})_k$}$$

holds every $k \in \mathbb{N}$, and moreover that, for every couple of open subset $\Omega' \Subset \Omega'' \Subset \Omega$ and $k \in \mathbb{N}$ there exists a constant c depending only on $n, p, L/\nu, k$, $\operatorname{dist}(\Omega', \partial\Omega'')$, and $\|u\|_{L^\infty(\Omega'')}$, but independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that

$$\int_{\Omega'} \big(|\mathfrak{X}u|^{p+2+\sigma_k}+|Tu|^{\frac{p+2+\sigma_k}{2}}\big)\,dx \leqslant c\int_{\Omega''}\big(|\mathfrak{X}u|^p+|Tu|^p+1\big)\,dx. \tag*{$(\mathbf{B})_k$}$$

We shall eventually show that this will suffice to prove (7.1) and (7.13). Before going on let us point out that when proving estimates like $(\mathbf{B})_k$ we shall deal with similar estimates where Ω', Ω'' vary in an arbitrary way. Each time we shall implicitly pass to different open subsets, since every time the open subsets involved in the inequalities will be arbitrary.

Let us first prove the validity of $(\mathbf{A})_0$ and $(\mathbf{B})_0$. The parts of the statements concerning $\mathfrak{X}u$ directly come from Lemma 6.1, therefore we concentrate on Tu. To this aim we apply Lemma 7.2 twice. First we choose $\sigma = 0$, recalling that $(p+2)/2 \leqslant p$ in turn implies $Tu \in L^{(p+2)/2}_{\mathrm{loc}}(\Omega)$; at this point we get that $Tu \in L^{(p+3)/2}_{\mathrm{loc}}(\Omega)$ with a first corresponding estimate, that is

$$\int_{\Omega'} |Tu|^{\frac{p+3}{2}}\,dx \leqslant c \int_{\Omega''} \big(|\mathfrak{X}u|^{p+2} + |Tu|^{\frac{p+2}{2}} + 1\big)\,dx.$$

Then we are able to apply again Lemma 7.2, this time with $\sigma = 1$, getting that $Tu \in L^{(p+4)/2}_{\mathrm{loc}}(\Omega)$ and, in view of (7.9), also that

$$\int_{\Omega'} |Tu|^{\frac{p+4}{2}}\,dx \leqslant c \int_{\Omega''} \big(|\mathfrak{X}u|^{p+3} + |Tu|^{\frac{p+3}{2}} + 1\big)\,dx.$$

Joining the last two estimates to (6.2), passing each time to different open subsets, which are not renamed, we easily get the also the part of $(\mathbf{B})_0$ concerned with Tu.

Let us now assume the validity of $(\mathbf{A})_k$ and $(\mathbf{B})_k$ for some $k \geqslant 0$, and let us prove that of $(\mathbf{A})_{k+1}$ and $(\mathbf{B})_{k+1}$. By $(\mathbf{A})_k$ we may apply Lemma 7.2 with the choice $\sigma \equiv \sigma_k$ in order to get that

$$Tu \in L^{\frac{p+3+\sigma_k}{2}}_{\mathrm{loc}}(\Omega). \tag{7.15}$$

Observe that by the very definition of σ_k we have that

$$\sigma_{k+1} < \sigma_k + 1, \tag{7.16}$$

and therefore from (7.15) we immediately get that

$$Tu \in L^{\frac{p+2+\sigma_{k+1}}{2}}_{\mathrm{loc}}(\Omega). \tag{7.17}$$

We also observe that using $(\mathbf{B})_k$ and the estimate (7.9) for $\sigma \equiv \sigma_k$, since in every occurrence the open subsets $\Omega' \Subset \Omega''$ are arbitrary, we easily gain

$$\int_{\Omega'} |Tu|^{\frac{p+2+\sigma_{k+1}}{2}}\,dx \leqslant \int_{\Omega'} \big(|Tu|^{\frac{p+3+\sigma_k}{2}} + 1\big)\,dx \leqslant c \int_{\Omega''} \big(|\mathfrak{X}u|^p + |Tu|^p + 1\big)\,dx, \tag{7.18}$$

that in turn holds for every couple of $\Omega' \Subset \Omega''$ where c depends as in $(\mathbf{B})_{k+1}$. Here we used again (7.16) and an elementary estimation. We have indeed proved one part of $(\mathbf{B})_{k+1}$ too. Therefore it only remains to prove that $\mathfrak{X}u \in L^{p+2+\sigma_{k+1}}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$, that will complete the proof of $(\mathbf{A})_{k+1}$, and the corresponding remaining part of $(\mathbf{B})_{k+1}$ with the estimation of $\mathfrak{X}u$. For this we wish to use

Lemma 7.1 with the choice $\sigma \equiv \sigma_{k+1}$, therefore let us check its applicability; estimate (7.16), assumption $(\mathbf{A})_k$ and (7.17) imply that we actually just have to check the second inclusion in (7.4). To do this we apply Young's inequality as follows:

$$|\mathfrak{X}u|^{p-2+\sigma_{k+1}}|Tu|^2 \leqslant |\mathfrak{X}u|^{\frac{(p-2+\sigma_{k+1})(p+3+\sigma_k)}{p-1+\sigma_k}} + |Tu|^{\frac{p+3+\sigma_k}{2}}.$$

By the definition in (7.14) we have that

$$\frac{(p-2+\sigma_{k+1})(p+3+\sigma_k)}{p-1+\sigma_k} = p+2+\sigma_k,$$

and hence the second inclusion in (7.4) follows with $\sigma \equiv \sigma_{k+1}$ by the first inclusion in $(\mathbf{A})_k$ and (7.15). Therefore Lemma 7.1 and (7.5) with $\sigma \equiv \sigma_{k+1}$ finally imply that $\mathfrak{X}u \in L^{p+2+\sigma_{k+1}}_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$. Concerning the remaining part of the proof of $(\mathbf{B})_{k+1}$ observe that (7.16) allows for applying the elementary inequality $|\mathfrak{X}u|^{p+\sigma_{k+1}} \leqslant |\mathfrak{X}u|^{p+2+\sigma_k} + 1$; this, together with (7.18) and (7.6), since the open subsets involved everywhere are arbitrary, allows in turn to conclude that

$$\int_{\Omega'} |\mathfrak{X}u|^{p+2+\sigma_{k+1}}\,dx \leqslant c\int_{\Omega''}\big(|\mathfrak{X}u|^{p+2+\sigma_k} + |Tu|^{\frac{p+2+\sigma_k}{2}} + 1\big)\,dx.$$

At this point the full inequality in $(\mathbf{B})_{k+1}$ follows by the previous one together with (7.18) and $(\mathbf{B})_k$, after changing, accordingly, the open subsets Ω', Ω'' involved.

In this way both $(\mathbf{A})_k$ and $(\mathbf{B})_k$ hold for every $k \in \mathbb{N}$.

Now we prove the validity of (7.1) and (7.13). The assertions in (7.1) are immediate, while to prove (7.13) with a fixed q, take k large enough such that $(p+2+\sigma_k)/2 \geqslant q$, in order to estimate $|\mathfrak{X}u|^q + |Tu|^q \leqslant |\mathfrak{X}u|^{p+2+\sigma_k} + |Tu|^{\frac{p+2+\sigma_k}{2}} + 2$, and then apply $(\mathbf{B})_k$ in order to get

$$\int_{\Omega'}\big(|\mathfrak{X}u|^q + |Tu|^q\big)\,dx \leqslant c\int_{\Omega''}\big(|\mathfrak{X}u|^p + |Tu|^p + 1\big)\,dx.$$

Finally, changing again the subsets, the final form of (7.13) follows by Theorem 3.2.

Step 2: Blow-up and local estimates. Now, by means of scaling arguments, we shall see how to get the precise form of the a priori estimates in (7.2)–(7.3) from the rough one in (7.13); of course we shall assume that $q > p$. First, let us consider the case of a solution $v \in HW^{1,p}(B(0,1))$ to (1.1), that is, when $\Omega \equiv B(0,1) \equiv B_1$. In the following γ will denote a number such that $\gamma \in (0,1)$, and the constants in the subsequent estimates will deteriorate when $\gamma \nearrow 1$. Applying Theorem 3.1 we find

$$\|v\|_{L^\infty(B_\gamma)} \leqslant c_1\big(\|v\|_{L^p(B_1)} + \mu\big), \tag{7.19}$$

where $c_1 \equiv c_1(n, p, L/\nu, \gamma)$. Now let us define, for every $z \in \mathbb{R}^{2n}$

$$w := \frac{v}{A} \quad \text{and} \quad \tilde{a}(z) := \frac{a(Az)}{A^{p-1}}, \tag{7.20}$$

where

$$A := c_1\big(\|v\|_{L^p(B_1)} + \mu\big). \tag{7.21}$$

Obviously $A > 0$ and moreover

$$\mu/A \leqslant 1. \tag{7.22}$$

The new scaled function w weakly solves the equation

$$\operatorname{div}_H \tilde{a}(\mathfrak{X}w) = 0, \tag{7.23}$$

and, as a consequence of (7.19), it is such that

$$\|w\|_{L^\infty(B_\gamma)} \leqslant 1. \tag{7.24}$$

Moreover an easy computation reveals that the new vector field $\tilde{a}(z)$ defined in (7.20) satisfies assumptions (1.2)–(1.3) with μ replaced by μ/A. Therefore, keeping again (7.22) in mind, applying estimate (7.13) to w with the choice $\Omega' = B_{\gamma^2}$ and $\Omega'' = B_\gamma$, yields

$$\int_{B_{\gamma^2}} \big(|\mathfrak{X}w|^q + |Tw|^q\big)\,dx \leqslant c_2 \int_{B_1} \big(|\mathfrak{X}w|^p + 1\big)\,dx, \tag{7.25}$$

and the constant c_2 depends now only on $n, p, L/\nu, q, \gamma$ by the inequality in (7.24). Scaling back to v, that is taking (7.20) into account, (7.25) gives

$$\begin{aligned} \int_{B_{\gamma^2}} \big(|\mathfrak{X}v|^q + |Tv|^q\big)\,dx &\leqslant c_2\big[c_1\big(\|v\|_{L^p(B_1)} + \mu\big)\big]^{q-p} \int_{B_1} |\mathfrak{X}v|^p\,dx \\ &\quad + |B_1| c_2\big[c_1\big(\|v\|_{L^p(B_1)} + \mu\big)\big]^q. \end{aligned} \tag{7.26}$$

Applying Young's inequality with conjugate exponents q/p and $q/(q-p)$ to estimate the first quantity in the right-hand side of (7.26) easily gives

$$\|\mathfrak{X}v\|_{L^q(B_{\gamma^2})} + \|Tv\|_{L^q(B_{\gamma^2})} \leqslant c\big(\|\mathfrak{X}v\|_{L^p(B_1)} + \|v\|_{L^p(B_1)} + \mu\big), \tag{7.27}$$

where $c \equiv c(n, p, L/\nu, q, \gamma)$. Now we observe that if v solves (1.1) then $v - \xi$ also solves (1.1) whenever $\xi \in \mathbb{R}$. Therefore we apply estimate (7.27) to $v - (v)_{B_1}$, and using it together with Jerison's Poincaré's inequality – see [32,39] – that is $\|v - (v)_{B_1}\|_{L^p(B_1)} \leqslant c(n,p)\|\mathfrak{X}v\|_{L^p(B_1)}$, we finally get

$$\|\mathfrak{X}v\|_{L^q(B_{\gamma^2})} + \|Tv\|_{L^q(B_{\gamma^2})} \leqslant c\big\||\mathfrak{X}v| + \mu\big\|_{L^p(B_1)}, \tag{7.28}$$

where $c \equiv c(n, p, L/\nu, q, \gamma)$; observe that the constant c blows-up whenever: $\gamma \nearrow 1$, $q \nearrow \infty$, $p \nearrow 4$. Choosing $\gamma = 1/\sqrt{2}$ in (7.28), we immediately get that

$$\left(\fint_{B_{1/2}} \left(|\mathfrak{X}v|^q + |Tv|^q\right) dx \right)^{1/q} \leqslant c \left(\fint_{B_1} (\mu + |\mathfrak{X}v|)^p \, dx \right)^{1/p}, \tag{7.29}$$

with $c \equiv c(n, p, L/\nu, q)$, and this means that we have proved (7.2)–(7.3) in the case $R = 1$.

Now we can go back to the original solution u, taking a CC-ball $B_R \equiv B(x_0, R) \subset \Omega$, and defining

$$v(x) := \frac{u(x_0 \cdot \delta_R(x))}{R}, \quad \text{for every } x \in B(0, 1), \tag{7.30}$$

where the dilation operator δ_R has been defined in (2.6). Now observe that for every $i = 1, \dots, 2n$

$$X_i v(x) = X_i u\big(x_0 \cdot \delta_R(x)\big) \quad \text{and} \quad Tv(x) = RTu\big(x_0 \cdot \delta_R(x)\big). \tag{7.31}$$

Using this fact, and again the left invariance of the vector fields $\{X_i\}$, it is easy to see that the function v defined in (7.31) solves Eq. (1.1) in $B(0, 1)$, and therefore (7.29) is applicable. In fact, using (7.29) for v, re-scaling back to u in $B(x_0, R)$, and using (7.31) we get (7.2)–(7.3). Observe that in such a re-scaling procedure the appearance of the integral averages in (7.2)–(7.3) is essentially due to the change-of-variable formula together with the fact that det $(x \mapsto x_0\delta_R(x)) \approx R^Q \approx |B(x_0, R)|$. This is basically a consequence of (2.7). □

8. Non-degenerate equations

Proposition 8.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.1) *under the assumptions* (1.2)–(1.4)*, with* $2 \leqslant p < 4$. *Then it holds that*

$$\mathfrak{X}u \in L^\infty_{\mathrm{loc}}\big(\Omega, \mathbb{R}^{2n}\big) \quad \textit{and} \quad Tu \in L^\infty_{\mathrm{loc}}(\Omega). \tag{8.1}$$

Moreover there exists a constant c, depending on n, p and L/ν, but otherwise independent of μ, of the solution u, and of the vector field $a(\cdot)$, such that (1.12)–(1.13) *hold for any CC-ball $B_R \subset \Omega$.*

Proof. The proof is again divided in two steps. First we treat a special case; then we reduce to such a special case by a blow-up argument.

Step 1: Universal estimates. Here we assume that

$$\Omega \equiv B_1 \quad \text{and} \quad \|\mathfrak{X}u\|_{L^p(B_1, \mathbb{R}^{2n})} \leqslant 1, \tag{8.2}$$

and we shall prove that there exist absolute constants $c_3, c_4 \equiv c_3, c_4(n, p, L/\nu)$ such that

$$\sup_{B_{1/2}} |\mathfrak{X}u| \leqslant c_3, \quad \text{and} \quad \sup_{B_{1/2}} |Tu| \leqslant c_4 \mu^{\frac{Q(2-p)}{4}}. \tag{8.3}$$

With $\gamma = 99/100$, a simple covering argument and (7.2)–(7.3), gives that

$$\int_{B_\gamma} \left(|\mathfrak{X}u|^{\frac{Q(p+2)}{Q-1}} + |Tu|^{2Q} + |Tu|^2\right) dx \leqslant c, \tag{8.4}$$

where c is a constant depending only on the quantities $n,\ p,\ L/\nu$. Note that we have used (8.2) to get rid of the dependence on the norms of $\mathfrak{X}u, Tu$ in the constant c. Now we start from (5.12), which we shall employ to implement a suitable variant of Moser's iteration scheme. With $\eta \in C_0^\infty(B_\gamma)$ being non-negative and such that $\eta \leqslant 1$ we immediately have that for any $\sigma \geqslant 2$ it does hold that

$$\begin{aligned}
&\int_\Omega \eta^2\left(\mu^2 + |\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{\sigma}{2}} |\mathfrak{X}X_s u|^2\, dx \\
&\quad \leqslant c(\sigma+1)C_\eta \int_{\operatorname{supp}\eta} \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{p+\sigma}{2}} dx \\
&\qquad + c(\sigma+1)^3 \int_\Omega \eta^2 |Tu|^2 \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{p-2+\sigma}{2}} dx,
\end{aligned} \tag{8.5}$$

where we have set

$$C_\eta := \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|T\eta\|_{L^\infty} + 1. \tag{8.6}$$

To estimate the last term appearing in (8.5) we use Hölder's inequality and then (8.4), thereby gaining

$$\begin{aligned}
&\int_\Omega \eta^2 |Tu|^2 \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{p-2+\sigma}{2}} dx \\
&\quad \leqslant c(n)\left(\int_{B_\gamma} |Tu|^{2Q}\, dx\right)^{\frac{1}{Q}} \left(\int_{\operatorname{supp}\eta} \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{Q(p-2+\sigma)}{2(Q-1)}} dx\right)^{\frac{Q-1}{Q}} \\
&\quad \leqslant c\left(\int_{\operatorname{supp}\eta} 1 + \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{Q(p+\sigma)}{2(Q-1)}} dx\right)^{\frac{Q-1}{Q}},
\end{aligned}$$

where, as we used (8.4), the constant c in the last line depends on $n,\ p,\ L/\nu$. Moreover, again by Hölder's inequality, it trivially follows that

$$\begin{aligned}
\int_{\operatorname{supp}\eta} \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{p+\sigma}{2}} dx &\leqslant c(n)\left(\int_{\operatorname{supp}\eta} \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{Q(p+\sigma)}{2(Q-1)}} dx\right)^{\frac{Q-1}{Q}} \\
&\leqslant c(n)\left(\int_{\operatorname{supp}\eta} 1 + \sum_{s=1}^{2n} \left(\mu^2 + |X_s u|^2\right)^{\frac{Q(p+\sigma)}{2(Q-1)}} dx\right)^{\frac{Q-1}{Q}}.
\end{aligned} \tag{8.7}$$

The last two estimates together with (8.5), and again Hölder's inequality, give

$$\int_{\Omega} \eta^2 \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{p-2+\sigma}{2}} |\mathfrak{X} X_s u|^2 \, dx$$
$$\leqslant c(\sigma+1)^3 C_\eta \bigg(\int_{\operatorname{supp}\eta} 1 + \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{Q(p+\sigma)}{2(Q-1)}} \, dx \bigg)^{\frac{Q-1}{Q}}, \tag{8.8}$$

where $c \equiv c(n, p, L/\nu)$ and C_η is defined in (8.6). Now we observe that

$$\bigg| \mathfrak{X} \sum_{s=1}^{2n} \eta \big(\mu^2 + |X_s u|^2\big)^{\frac{p+\sigma}{4}} \bigg|^2 \leqslant c(n) C_\eta \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{p+\sigma}{2}}$$
$$+ c(n)(p+\sigma)^2 \eta^2 \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{p-2+\sigma}{2}} |\mathfrak{X} X_s u|^2.$$

Therefore, using again (8.7), the last estimate and (8.8) give

$$\int_{B_\gamma} \bigg| \mathfrak{X} \sum_{s=1}^{2n} \eta \big(\mu^2 + |X_s u|^2\big)^{\frac{p+\sigma}{4}} \bigg|^2 \, dx$$
$$\leqslant c(p+\sigma)^5 C_\eta \bigg(\int_{\operatorname{supp}\eta} 1 + \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{Q(p+\sigma)}{2(Q-1)}} \, dx \bigg)^{\frac{Q-1}{Q}}. \tag{8.9}$$

Applying Sobolev embedding theorem in the Heisenberg group, that is Theorem 2.1 with $q = 2$, in turn yields

$$\bigg(\int_{B_\gamma} \eta^{\frac{2Q}{Q-2}} \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{Q(p+\sigma)}{2(Q-2)}} \, dx \bigg)^{\frac{Q-2}{Q}}$$
$$\leqslant c(p+\sigma)^5 C_\eta \bigg(\int_{\operatorname{supp}\eta} 1 + \sum_{s=1}^{2n} \big(\mu^2 + |X_s u|^2\big)^{\frac{Q(p+\sigma)}{2(Q-1)}} \, dx \bigg)^{\frac{Q-1}{Q}}, \tag{8.10}$$

where the constant c depends only on $n, p, L/\nu$. Observe that here we are using that $\operatorname{supp}\eta \subset B_\gamma$. Now we choose the cut-off functions in the framework of Moser's iteration technique. We take a family of concentric interpolating balls $B_{3/4} \subset B_{\varrho_{k+1}} \subset B_{\varrho_k}$ such that $B_{\varrho_0} = B_{7/8} \subset B_\gamma$, $\varrho_{k+1} - \varrho_k \approx 2^{-k}$ and $\varrho_k \searrow 3/4$. Accordingly we select $\eta_k \in C_c^\infty(B_{\varrho_k})$ such that $\eta_k \equiv 1$ on $B_{\varrho_{k+1}}$, and $C_\eta \leqslant c^k$; the existence of such cut-off functions can be inferred as in [10, Lemma 3.2]. Setting

$$\tilde{\chi} := \frac{Q-1}{Q-2} > 1, \tag{8.11}$$

we recursively define the sequence $\{\sigma_k\}$ as follows:

$$\begin{cases} \sigma_{k+1} := \tilde{\chi}\sigma_k + \frac{p}{Q-2}, \\ \sigma_0 := 2, \end{cases}$$

so that

$$\frac{(p+\sigma_{k+1})Q}{Q-1} = \frac{(p+\sigma_k)Q}{Q-2} \tag{8.12}$$

holds for every $k \geqslant 0$. Observe that

$$p+\sigma_k \approx \tilde{\chi}^k, \quad \text{and} \quad |B_{\varrho_k}| \approx c(n) > 0. \tag{8.13}$$

Taking $\sigma \equiv \sigma_k$ and $\eta \equiv \eta_k$ in (8.10), and observing that $\eta_k \equiv 1$ on $B_{\varrho_{k+1}}$ and $\operatorname{supp}\eta_k \subset B_{\varrho_k}$, easily gives

$$\begin{aligned} &\left(\int_{B_{\varrho_{k+1}}} 1+\sum_{s=1}^{2n}\left(\mu^2+|X_s u|^2\right)^{\frac{Q(p+\sigma_k)}{2(Q-2)}}\,dx\right)^{\frac{Q-2}{Q}} \\ &\quad\leqslant c^{k+1}\left(\int_{B_{\varrho_k}} 1+\sum_{s=1}^{2n}\left(\mu^2+|X_s u|^2\right)^{\frac{Q(p+\sigma_k)}{2(Q-1)}}\,dx\right)^{\frac{Q-1}{Q}}, \end{aligned} \tag{8.14}$$

where $c \equiv c(n,p,L/\nu) \geqslant 1$ is a constant independent of k, and we used (8.13). Now, setting for every $k \geqslant 0$

$$A_k := \left(\int_{B_{\varrho_k}} 1+\sum_{s=1}^{2n}\left(\mu^2+|X_s u|^2\right)^{\frac{Q(p+\sigma_k)}{2(Q-1)}}\,dx\right)^{\frac{Q-1}{Q(p+\sigma_k)}},$$

using (8.12)–(8.14), an elementary manipulation gives that

$$A_{k+1} \leqslant c_0^{(k+1)\tilde{\chi}^{-k}} A_k,$$

for a new constant c_0 depending only on $n, p, L/\nu$. Keeping (8.11) in mind, iterating the previous inequality easily gives

$$A_{k+1} \leqslant \exp\left[(\log c_0)\sum_{i=0}^{\infty}\frac{i+1}{\tilde{\chi}^i}\right]A_0.$$

Letting $k \nearrow \infty$ in the previous inequality – note that the series in the last line converges by (8.11) – now gives

$$\sup_{B_{3/4}} |\mathfrak{X}u| \leqslant c(n,p,L/\nu)A_0, \tag{8.15}$$

while taking (8.4) and the fact that $\mu \leqslant 1$ into account we obtain the first inequality appearing in (8.3). As for the second inequality in (8.3), we observe that since $\mathfrak{X}u$ is bounded we may apply Theorem 3.3 with any q satisfying (3.4). Noting that this implies $2q/(q-p+2) \leqslant 2Q$, we may use (8.4); therefore taking $R = 3/4$ and $\varrho = 1/2$ in (3.5) yields

$$\|Tu\|_{L^\infty(B_{1/2})} \leqslant \tilde{c}c^{\frac{\chi}{\chi-1}}\mu^{\frac{(2-p)\chi}{2(\chi-1)}}, \tag{8.16}$$

where we also used (8.15), $\tilde{c} \equiv \tilde{c}(n, p, L/\nu)$, and where χ appears in (3.6). All the constants in the above inequality only depend on n, p, L/ν and are actually independent of q. Therefore letting $q \nearrow \infty$ in (8.16), and keeping (3.6) in mind, we obtain the second inequality in (8.3) with the specified dependence of the constant c_4.

Step 2: The general case. First we observe that we may reduce to the case $B_R \equiv B_1$ by performing the blow-up scaling (7.30). Indeed once estimates (1.12)–(1.13) hold for v on $B_R \equiv B_1$, then scaling back, and using (7.31), they also hold on general balls B_R as required in the statement. Therefore we just need to prove the result for a solution v in the ball B_1. In order to reduce to the assumptions in (8.2) we pass to the function w defined in (7.20) where this time we choose $A := (\|\mathfrak{X}v\|_{L^p(B_1)} + \mu)$, so that both $\|\mathfrak{X}w\|_{L^p(B_1,\mathbb{R}^{2n})} \leqslant 1$ and (7.22) hold. As noted in the proof of Proposition 7.1, Step 2, the function w is a solution of Eq. (7.23), while the new vector field $\tilde{a}(z)$ defined in (7.20) satisfies assumptions (1.2)–(1.3) with μ replaced by $\mu/A \leqslant 1$. Therefore, thanks to (7.22) we may apply the result of Step 1 to w, thereby obtaining

$$\sup_{B_{1/2}} |\mathfrak{X}w| \leqslant c_3, \quad \text{and} \quad \sup_{B_{1/2}} |Tw| \leqslant c_4\mu^{\frac{Q(2-p)}{4}} A^{\frac{Q(p-2)}{4}}. \tag{8.17}$$

Going back to $v = w/A$, and keeping in mind the current definition of A, we obtain the validity of (1.12)–(1.13) for v on B_1, and the proof is finally complete by the argument outlined at the beginning of Step 2. □

Proof of Theorems 1.1–1.2. The proof of the a priori estimates of Theorem 1.2 is a direct consequence of Proposition 8.1. As far as the Hölder continuity of the gradient is concerned, the focal point of the regularity theory for quasilinear elliptic equations with p-growth is the local Lipschitz regularity of solutions, as already explained in [8,9,40]. From this point on the proof of the local Hölder continuity of Du proceeds as in [40]; see also [7,9] for detailed explanations. □

9. The degenerate case

Proof of Theorem 1.3. Of course in the following we shall restrict to the case $p > 2$; indeed, as the reader will soon recognize, in the case $p = 2$ the role of μ is immaterial in (1.2)–(1.3), and the results of Theorems 1.1 and 1.2 still hold when $\mu = 0$. When $p > 2$ the case $\mu = 0$ is now a consequence of Proposition 7.1 when combined with a suitable approximation argument we are going to report in some detail. Let us consider the regularized vector fields

$$a_k(z) := a(z) + \varepsilon_k^{p-2} z, \quad \text{for every } z \in \mathbb{R}^{2n} \text{ and } k \in \mathbb{N},$$

where $\{\varepsilon_k\}_k$ is a sequence of positive numbers such that $\varepsilon_k \searrow 0$ and $\varepsilon_k \leqslant 1$. By using (1.2)–(1.3) it is easy to see that each vector field $a_k(z)$ satisfies the following growth and ellipticity conditions:

$$\left|Da_k(z)\right|\left(\varepsilon_k^2+|z|^2\right)^{\frac{1}{2}}+\left|a_k(z)\right| \leqslant c\left(\varepsilon_k^2+|z|^2\right)^{\frac{p-1}{2}}, \tag{9.1}$$

and

$$c^{-1}\left(\varepsilon_k^2+|z|^2\right)^{\frac{p-2}{2}}|\lambda|^2 \leqslant \sum_{i,j=1}^{2n} D_{z_j}(a_k)_i(z)\lambda_i\lambda_j, \tag{9.2}$$

for a constant $c>0$ depending only on $n, p, L/\nu$ but independent of $k \in \mathbb{N}$. Moreover, since $p \geqslant 2$, assumption (9.2) also implies, for a possibly different constant c still depending on $n, p, L/\nu$, but otherwise independent of $k \in \mathbb{N}$, that whenever $z, z_1, z_2 \in \mathbb{R}^{2n}$ the following inequalities hold:

$$c^{-1}|z_2-z_1|^p \leqslant \left\langle a_k(z_2)-a_k(z_1), z_2-z_1\right\rangle, \qquad c^{-1}|z|^p - c\varepsilon_k^p \leqslant \left\langle a_k(x,z), z\right\rangle. \tag{9.3}$$

Compare with (3.15) and (3.16). Now, let us consider a CC-ball $B_R \subset \Omega$ and let us define $u_k \in u + HW_0^{1,p}(B_R)$ as the unique solution to the Dirichlet problem (3.17) with $a_k(\cdot) \equiv a(\cdot)$; therefore, for the present application we have $v \equiv u_k$ in (3.17). Accordingly, by virtue of (9.3) we may apply Lemma 3.3 so that (3.18) used for $v \equiv u_k$ gives

$$\int_{B_R} |\mathfrak{X}u_k|^p\,dx \leqslant c\int_{B_R}\left(\varepsilon_k+|\mathfrak{X}u|\right)^p dx, \tag{9.4}$$

where $c \equiv c(n, p, L/\nu)$ is independent of k. Next, using (9.3), the fact that both u and u_k are solutions, and then applying the definition of $a_k(\cdot)$ together with Young's and Hölder's inequalities, we have

$$\begin{aligned}
\int_{B_R} |\mathfrak{X}u_k-\mathfrak{X}u|^p\,dx &\leqslant c\int_{B_R}\left\langle a(\mathfrak{X}u_k)-a(\mathfrak{X}u), \mathfrak{X}u_k-\mathfrak{X}u\right\rangle dx\\
&= c\int_{B_R}\left\langle a(\mathfrak{X}u_k)-a_k(\mathfrak{X}u_k), \mathfrak{X}u_k-\mathfrak{X}u\right\rangle dx\\
&\leqslant c\int_{B_R}\varepsilon_k^{p-2}|\mathfrak{X}u_k||\mathfrak{X}u_k-\mathfrak{X}u|\,dx\\
&\leqslant \frac{1}{2}\int_{B_R}|\mathfrak{X}u_k-\mathfrak{X}u|^p\,dx + c\varepsilon_k^{\frac{p(p-2)}{p-1}}\int_{B_R}|\mathfrak{X}u_k|^{\frac{p}{p-1}}\,dx\\
&\leqslant \frac{1}{2}\int_{B_R}|\mathfrak{X}u_k-\mathfrak{X}u|^p\,dx + c\varepsilon_k^{\frac{p(p-2)}{p-1}}\left(\int_{B_R}|\mathfrak{X}u_k|^p\,dx\right)^{\frac{1}{p-1}}.
\end{aligned}$$

Re-absorbing in the l.h.s. the first integral in the last line, eventually letting $k \nearrow \infty$, and keeping (9.4) in mind, we get

$$\mathfrak{X}u_k \to \mathfrak{X}u \quad \text{strongly in } L^p\big(B_R, \mathbb{R}^{2n}\big). \tag{9.5}$$

Now, using estimates (1.12) and (1.13) for u_k, and therefore considering the case $\mu \equiv \varepsilon_k > 0$, we get

$$\sup_{B_{R/2}} |\mathfrak{X}u_k| \leqslant c^* \bigg(\fint_{B_R} \big(\varepsilon_k + |\mathfrak{X}u_k|\big)^p \, dx \bigg)^{1/p}, \tag{9.6}$$

and

$$\bigg(\fint_{B_{R/2}} |Tu_k|^q \, dx \bigg)^{1/q} \leqslant \frac{c_*}{R} \bigg(\fint_{B_R} \big(\varepsilon_k + |\mathfrak{X}u_k|\big)^p \, dx \bigg)^{1/p}, \tag{9.7}$$

which hold uniformly with respect to k; in fact the constants c^*, c_* ultimately depend on n, p, L/ν, and also q as far as the latter is concerned, but are otherwise independent of k. This follows directly from the statement of Proposition 7.1. Letting $k \nearrow \infty$ in (9.6)–(9.7), standard lower semicontinuity arguments to deal with the left-hand sides of (9.6)–(9.7), and (9.5) to deal with right-hand ones, finally give (1.16)–(1.17). Since the ball considered $B_R \subset \Omega$ is arbitrary, this finally implies (7.1) via a standard covering argument and the proof of Theorem 1.3 is complete. □

Proof of Corollaries 1.1–1.2. Corollary 1.2 is immediate since from Theorem 1.3 we obtain higher integrability for the Euclidean gradient: $Du \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n+1})$ for every $q < \infty$. As for Corollary 1.1, it suffices to prove estimate (1.18). With $B_R \subset \Omega$ as in the statement, by (1.12)–(1.16) it immediately follows that

$$M_{R/4}\big(|\mathfrak{X}u|\big)(x) \leqslant \sup_{B(x,R/4)} |\mathfrak{X}u| \leqslant c \bigg(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p \, dz \bigg)^{1/p},$$

whenever $x \in B_{R/2}$, where c depends only on $n, p, L/\nu$. The operator $M_{R/4}$ is the one defined in (2.26). Therefore, using Proposition 2.1 we obtain

$$\big|u(x) - u(y)\big| \leqslant c \bigg(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p \, dz \bigg)^{1/p} d_{\mathrm{CC}}(x, y) \tag{9.8}$$

as soon as $x, y \in B_{R/2}$ are such that $d_{\mathrm{CC}}(x, y) \leqslant R/8$. At this stage estimate (1.18) follows from the last one, applied to suitable smaller balls, just magnifying the constant in (9.8) of a finite factor, say 16. □

10. Horizontal Calderón–Zygmund estimates

In this section we are going to prove Theorem 1.4; the use of various types of restricted maximal operator will be essential here. In the following, when dealing with (1.20) we shall always assume that $F \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$, for some $q > p$. Now, let us fix an arbitrarily fixed open subset $\Omega' \Subset \Omega$; for the rest of the section all balls the considered B will be such that $B \Subset \Omega'$ unless otherwise specified, and in the following all the regularity results we are going to prove are in Ω'. Since the choice of Ω' is arbitrary the corresponding local regularity of $\mathfrak{X}u$ in Ω will also follow. With $\tilde{q} \equiv \tilde{q}(n, p, L/\nu) > p$ being the higher integrability exponent identified in Theorem 3.4, let us define

$$q_0 := \frac{p + \tilde{q}}{2} \tag{10.1}$$

which is such that $q_0 \in (p, \tilde{q})$ and can be therefore used in (3.19). Moreover, for later use we observe that

$$q > \tilde{q} \quad \Longrightarrow \quad \frac{1}{q - q_0} < \frac{2}{\tilde{q} - p} \equiv c(n, p, L/\nu) \tag{10.2}$$

and the last dependence on the parameters follows from the one specified in Theorem 3.4. Accordingly, with $R_0 > 0$ being fixed, and eventually specified later, and with $\Omega' \Subset \Omega$ chosen as described above, we let

$$[b]^*_{R_0} \equiv [b]^*_{R_0, \Omega'} := \sup_{B_R \subseteq \Omega', R \leqslant R_0} \left(\fint_{B_R} \left| b(x) - (b)_{B_R} \right|^{(\frac{p}{p-1})(\frac{q_0}{q_0 - p})} dx \right)^{\frac{q_0 - p}{q_0}}, \tag{10.3}$$

where $(b)_{B_R}$ is the average in (2.14). Let us observe that

$$\lim_{R \searrow 0} [b]^*_{R, \Omega'} = 0, \tag{10.4}$$

for every choice of the open subset $\Omega' \Subset \Omega$, and this strengthens (2.15). Indeed since $|b(x)| \leqslant L$ by (1.21) we have

$$[b]^*_{R, \Omega'} \leqslant (2L)^{\frac{p}{p-1} - \frac{q_0 - p}{q_0}} \left([b]_{R, \Omega'} \right)^{\frac{q_0 - p}{q_0}}$$

and (10.4) immediately follows by (2.15).

Lemma 10.1. *Let $u \in HW^{1,p}(B_R)$ be a weak solution to* (1.20)*, and let $v \in u + HW^{1,p}_0(B_R)$ be a weak solution to the Dirichlet problem* (3.17) *under the assumptions* (1.2)–(1.3) *for $p \geqslant 2$, where $B_{2R} \Subset \Omega'$ and $R \leqslant R_0$, for a certain $R_0 > 0$.*

(1) *For any $p \geqslant 2$ it holds that*

$$\fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx \leqslant c_5 [b]^*_{R_0} \fint_{B_{2R}} \left(\mu + |\mathfrak{X}u| \right)^p dx + c_5 \left(1 + [b]^*_{R_0} \right) \left(\fint_{B_{2R}} |F|^{q_0} \, dx \right)^{p/q_0}, \tag{10.5}$$

where the constant c_5 depends only on n, p, L/ν.

(2) *Assuming* $p \in [2,4)$ *we have that for any* $p \leqslant s < \infty$ *there exists a constant* $c_6 \equiv c_6(n,p,L/\nu)$ *such that*

$$\left(\fint_{B_{R/2}} \big(\mu + |\mathfrak{X}v|\big)^s \, dx \right)^{\frac{1}{s}} \leqslant c_6 \left(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^p \, dx \right)^{\frac{1}{p}}, \tag{10.6}$$

and $c_6 \nearrow \infty$ *when* $p \nearrow 4$.

Proof. (1) Using that u and v are solutions to (1.20) and (3.17) respectively, testing (1.20) and (3.17) by $u - v \in HW_0^{1,p}(B_R)$ and summing up, with $(b)_{B_R}$ as in (2.14) we have

$$\begin{aligned}
\fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx &\overset{(3.15)}{\leqslant} c \fint_{B_R} (b)_{B_R} \big\langle a(\mathfrak{X}u) - a(\mathfrak{X}v), \mathfrak{X}u - \mathfrak{X}v \big\rangle \, dx \\
&= c \fint_{B_R} \big\langle (b)_{B_R} a(\mathfrak{X}u) - b(x) a(\mathfrak{X}u), \mathfrak{X}u - \mathfrak{X}v \big\rangle \, dx \\
&\quad + c \fint_{B_R} \big\langle |F|^{p-2} F, \mathfrak{X}u - \mathfrak{X}v \big\rangle \, dx =: I + II,
\end{aligned} \tag{10.7}$$

where $c \equiv c(n,p,L/\nu)$. In a standard way, via Young's inequality we have in turn

$$II \leqslant \frac{1}{4} \fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx + c \fint_{B_R} |F|^p \, dx, \tag{10.8}$$

while, taking (1.2) into account and using Hölder's inequality we have

$$\begin{aligned}
I &\leqslant \frac{1}{4} \fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx + c \fint_{B_R} \big| b(x) - (b)_{B_R} \big|^{\frac{p}{p-1}} \big(\mu + |\mathfrak{X}u|\big)^p \, dx \\
&\leqslant \frac{1}{4} \fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx + c[b]^*_{R_0} \left(\fint_{B_R} \big(\mu + |\mathfrak{X}u|\big)^{q_0} \, dx \right)^{p/q_0} \\
&\overset{(3.19)}{\leqslant} \frac{1}{4} \fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx \\
&\quad + c[b]^*_{R_0} \fint_{B_{2R}} \big(\mu + |\mathfrak{X}u|\big)^p \, dx + c[b]^*_{R_0} \left(\fint_{B_{2R}} |F|^{q_0} \, dx \right)^{p/q_0}.
\end{aligned}$$

Estimate (10.5) now follows combining the estimates found for I and II to (10.7).

(2) When $p \in [2,4)$ estimate (10.6) just follows applying (1.12)–(1.16) to the function v, and then applying (3.18). $\square$

In the following we shall concentrate on a ball B_{R_0}, such that $B_{100R_0} \subset \Omega'$. The symbol M^* will denote the restricted maximal operator relative to the ball B_{100R_0} in the sense of (2.24): $M^* \equiv M^*_{B_{100R_0}}$; accordingly we shall denote by $M^*_{q_0/p}$ the restricted maximal operator in the sense of (2.25), again relative to B_{100R_0}, that is, $M^*_{q_0/p} \equiv M^*_{q_0/p, B_{100R_0}}$. We recall that $q_0 > p$ has been defined in (10.1).

Lemma 10.2. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (1.20) *under assumptions* (1.2)–(1.3) *with $2 \leqslant p < 4$, and let $K \geqslant 1$ and $s > p$. There exist numbers $\varepsilon \equiv \varepsilon(n, p, L/\nu, K, s) \in (0,1)$ and $A \equiv A(n, p, L/\nu) \geqslant 1$ such that if $[b]^*_{100R_0} \leqslant \varepsilon$ then the following holds:*

If B is a CC-ball centered in B_{R_0} and with radius less than $2R_0$ satisfying

$$|E \cap 5B| > K^{-s/p}|B \cap B_{R_0}| \tag{10.9}$$

then

$$5B \cap B_{R_0} \subset G, \tag{10.10}$$

where

$$E := \left\{x \in B_{R_0}\colon\ M^*\left(\mu^p + |\mathfrak{X}u|^p\right)(x) > AK\lambda,\ \textit{and}\ M^*_{q_0/p}\left(|F|^p\right)(x) \leqslant \varepsilon\lambda\right\},$$

and

$$G := \left\{x \in B_{R_0}\colon\ M^*\left(\mu^p + |\mathfrak{X}u|^p\right)(x) > \lambda\right\},$$

while $\lambda > 0$.

Proof. We proceed by contradiction, therefore assuming that (10.10) fails, and showing that, choosing ε and A appropriately, but with the dependence on the constants as in the statement of the lemma, also (10.9) fails. Indeed, assume that (10.10) fails but (10.9) does not; then there exists $z_1 \in 5B \cap B_{R_0}$ such that $M^*(\mu^p + |\mathfrak{X}u|^p)(z_1) \leqslant \lambda$; moreover $E \cap 5B$ is non-empty and therefore there exists $z_2 \in 5B \cap B_{R_0}$ such that $M^*_{q_0/p}(|F|^p)(z_2) < \varepsilon\lambda$. All in all we have that

$$\fint_{40B} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \leqslant \lambda, \quad \text{and} \quad \fint_{40B} |F|^{q_0}\, dx \leqslant (\varepsilon\lambda)^{q_0/p}. \tag{10.11}$$

Now define $v \in u + HW^{1,p}_0(20B)$ as the unique solution to the Dirichlet problem (3.17) with $B_R \equiv 20B$. Therefore applying (10.5) in this context, and using (10.11) with $[b]^*_{100R_0} \leqslant \varepsilon$ too, an elementary manipulation gives

$$\fint_{20B} |\mathfrak{X}u - \mathfrak{X}v|^p\, dx \leqslant c(n, p, L/\nu)\varepsilon\lambda. \tag{10.12}$$

Moreover estimates (10.6) and (10.11) also give

$$\fint_{10B} \left(\mu^s + |\mathfrak{X}v|^s\right) dx \leqslant \left[c(n,p,L/\nu)\right]^{s/p} \lambda^{s/p}. \tag{10.13}$$

We now start giving a few estimates for the restricted maximal operator relative to the ball $10B$, that in the following will be denoted by M^{**}, therefore $M^{**} \equiv M^*_{10B}$. First, let us observe that a standard geometric argument using that $M^*(\mu^p + |\mathfrak{X}u|^p)(z_1) < \lambda$, exactly the same as the one working in the Euclidean case, allows us to get the existence of an absolute constant c_*, depending on the doubling constant C_d in (2.8) and therefore ultimately on n, such that

$$M^*\left(\mu^p + |\mathfrak{X}u|^p\right)(x) \leqslant \max\left\{M^{**}\left(\mu^p + |\mathfrak{X}u|^p\right)(x), c_*\lambda\right\}, \tag{10.14}$$

whenever $x \in 5B \cap B_{R_0}$. Now, using (2.27) with $\gamma = s/p$, we have

$$\begin{aligned}
&\left|\left\{x \in 5B\colon\ M^{**}\left(\mu^p + |\mathfrak{X}u|^p\right)(x) > AK\lambda\right\}\right| \\
&\quad \leqslant \left|\left\{x \in 5B\colon\ M^{**}\left(\mu^p + |\mathfrak{X}v|^p\right)(x) > 2^{-p}AK\lambda\right\}\right| \\
&\qquad + \left|\left\{x \in 5B\colon\ M^{**}\left(|\mathfrak{X}u - \mathfrak{X}v|^p\right)(x) > 2^{-p}AK\lambda\right\}\right| \\
&\quad \overset{(2.27)}{\leqslant} \frac{2^{s/p+p}c(n,p)}{(AK\lambda)^{s/p}} \int_{10B} \left(\mu^s + |\mathfrak{X}v|^s\right) dx + \frac{c(n,p)}{AK\lambda} \int_{10B} |\mathfrak{X}u - \mathfrak{X}v|^p \, dx \\
&\quad \overset{(10.12)\text{–}(10.13)}{\leqslant} \frac{2^{s/p}[c(n,p,L/\nu)]^{s/p}|B|}{(AK)^{s/p}} + \frac{c(n,p,L/\nu)\varepsilon|B|}{AK} \\
&\quad \leqslant \frac{2^{s/p}[c_7(n,p,L/\nu,s)]^{s/p}|B \cap B_{R_0}|}{(AK)^{s/p}} + \frac{c_8(n,p,L/\nu)\varepsilon|B \cap B_{R_0}|}{AK}.
\end{aligned} \tag{10.15}$$

In the last inequality we used the fact that B is a ball centered in B_{R_0} whose radius does not exceed $2R_0$, and the doubling condition (2.8). Now we fix $A \equiv A(n,p,L/\nu) > 1 + c_*$ large enough in order to have $(2c_7/A)^{s/p} \leqslant 2c_7/A \leqslant 1/4$; here $c_* \equiv c_*(n)$ is the constant appearing in (10.14). Then we take $\varepsilon \equiv \varepsilon(n,p,L/\nu,K)$ in order to have $c_8\varepsilon K^{s/p-1} < 1/4$. Such choices fix the quantities A and ε with the dependence on the constants described in the statement of the lemma, and together with (10.15) they give

$$\left|\left\{x \in 5B \cap B_{R_0}\colon\ M^{**}\left(\mu^p + |\mathfrak{X}u|^p\right)(x) > AK\lambda\right\}\right| < K^{-s/p}|B|.$$

Now, since $K \geqslant 1$ and $A > c_*$, by (10.14) we also obtain

$$\left|\left\{x \in 5B \cap B_{R_0}\colon\ M^{*}\left(\mu^p + |\mathfrak{X}u|^p\right)(x) > AK\lambda\right\}\right| < K^{-s/p}|B|,$$

that finally contradicts (10.9), and the proof is complete. □

Proof of Theorem 1.4. The proof is actually split in two cases. The first is when $q \leqslant \tilde{q}$, and $\tilde{q} \equiv \tilde{q}(n,p,L/\nu) > p$ is the higher integrability exponent identified in Theorem 3.4. In this case the assertion follows directly from such a theorem. The other case is when $q > \tilde{q}$, to which we specialize henceforth. Therefore, with $\tilde{q} < q < \infty$ as in the statement, we fix a number s

such that $s > q$. Note that as a consequence of the choice of $s \equiv s(q)$, from now on all the constants depending on s will be actually depending on q, and as such they will be denoted, and in particular we determine the constant A when eventually using Lemma 10.2. Then we take $K > 1$ large enough in order to have

$$2K^{\frac{q-s}{p}} = A^{-\frac{q}{p}}. \tag{10.16}$$

Such a choice fixes $K \equiv K(n, p, L/\nu, q)$ and this is the number we are going to take when using Lemma 10.2. Therefore this determines the choice of $\varepsilon \equiv \varepsilon(n, p, L/\nu, q) > 0$ for the use in Lemma 10.2. Finally we determine the radius $R_0 \equiv (n, p, L/\nu, s, b(\cdot)) > 0$ in such a way that $[b]^*_{100R_0} \leqslant \varepsilon$. This is possible by (10.4). Now, let us set

$$\mu_1(t) := \bigl|\bigl\{x \in B_{R_0}\colon\ M^*\bigl(\mu^p + |\mathfrak{X}u|^p\bigr)(x) > t\bigr\}\bigr|, \tag{10.17}$$

$$\mu_2(t) := \bigl|\bigl\{x \in B_{R_0}\colon\ M^*_{q_0/p}\bigl(|F|^p\bigr)(x) > t\bigr\}\bigr|, \tag{10.18}$$

and keep in mind that the maximal operators $M^*_{q_0/p}$ are restricted to the ball B_{100R_0}. The proof will proceed by iterating the function $\mu_1(\cdot)$ using information on $\mu_2(\cdot)$, that is getting information on the measure of the level sets of $|\mathfrak{X}u|$, in terms of those of $|F|$. We choose the "starting level" λ_0 as follows:

$$\lambda_0 := 10C_d^{10} c_W K^{s/p} \fint_{B_{100R_0}} \bigl(\mu^p + |\mathfrak{X}u|^p\bigr)\,dx, \tag{10.19}$$

where C_d is the doubling constant appearing in (2.8), and $c_W \equiv c_W(n)$ is the constant appearing in (2.28) for $\gamma = 1$. Therefore using (2.28), and that $AK > 1$ we find, for any $m \in \mathbb{N}$

$$\mu_1\bigl((AK)^m\lambda_0\bigr) \leqslant \mu_1(\lambda_0) \leqslant \frac{1}{2K^{s/p}}|B_{R_0}|. \tag{10.20}$$

Now we want to combine Lemmas 10.2 and 2.1. More precisely, for every $m = 0, 1, 2, \ldots$ we want to apply Lemma 2.1 with the choice $\delta = K^{-s/p}$ and

$$E := \bigl\{z \in B_{R_0}\colon\ M^*\bigl(\mu^p + |\mathfrak{X}u|^p\bigr) > (AK)^{m+1}\lambda_0,\ \text{and}\ M^*_{q_0/p}\bigl(|F|^p\bigr) < \varepsilon(AK)^m\lambda_0\bigr\},$$
$$G := \bigl\{z \in B_{R_0}\colon\ M^*\bigl(\mu^p + |\mathfrak{X}u|^p\bigr) > (AK)^m\lambda_0\bigr\}.$$

In fact using Lemma 10.2 for $\lambda \equiv (AK)^m\lambda_0$ in the context of Lemma 2.1, keeping (10.20) in mind, and recalling that $|G| = \mu_1((AK)^m\lambda_0)$ and that $|E| \geqslant \mu_1((AK)^{m+1}\lambda_0) - \mu_2((AK)^m\varepsilon\lambda_0)$ we have

$$\mu_1\bigl((AK)^{m+1}\lambda_0\bigr) \leqslant K^{-s/p}\mu_1\bigl((AK)^m\lambda_0\bigr) + \mu_2\bigl((AK)^m\varepsilon\lambda_0\bigr),$$

for any $m = 0, 1, 2, \ldots$. Induction on the previous inequality easily gives

$$\mu_1\bigl((AK)^{m+1}\lambda_0\bigr) \leqslant K^{-s(m+1)/p}\mu_1(\lambda_0) + \sum_{i=0}^{m} K^{-s(m-i)/p}\mu_2\bigl((AK)^i\varepsilon\lambda_0\bigr),$$

and therefore, multiplying the previous inequalities by $(AK)^{q(m+1)/p}$ and summing up on $m = 0, 1, \dots, M \in \mathbb{N}$, we have

$$\begin{aligned}
&\sum_{m=0}^{M} (AK)^{q(m+1)/p} \mu_1\big((AK)^{m+1}\lambda_0\big) \\
&\quad \leqslant \left(\sum_{m=0}^{M} \big[K^{-s/p}(AK)^{q/p}\big]^{m+1} \right) \mu_1(\lambda_0) \\
&\qquad + \sum_{m=0}^{M} \sum_{i=0}^{m} (AK)^{q(m+1)/p} K^{-s(m-i)/p} \mu_2\big((AK)^i \varepsilon\lambda_0\big).
\end{aligned} \tag{10.21}$$

First, we notice that (10.16) implies

$$\sum_{m=0}^{\infty} \big[K^{-s/p}(AK)^{q/p}\big]^{m+1} = 1.$$

On the other hand, using Fubini's theorem for series it easily follows that

$$\begin{aligned}
&\sum_{m=0}^{M} \sum_{i=0}^{m} (AK)^{q(m+1)/p} K^{-s(m-i)/p} \mu_2\big((AK)^i \varepsilon\lambda_0\big) \\
&\quad \leqslant 2(AK)^{q/p} \sum_{m=0}^{M} (AK)^{qm/p} \mu_2\big((AK)^m \varepsilon\lambda_0\big).
\end{aligned}$$

Combining the last two inequalities with (10.21), and eventually letting $M \nearrow \infty$, we obtain

$$\begin{aligned}
\sum_{m=1}^{\infty} (AK)^{qm/p} \mu_1\big((AK)^m \lambda_0\big) &\leqslant \mu_1(\lambda_0) + 2(AK)^{q/p} \mu_2(\varepsilon\lambda_0) \\
&\quad + 2(AK)^{q/p} \sum_{m=1}^{\infty} (AK)^{qm/p} \mu_2\big((AK)^m \varepsilon\lambda_0\big).
\end{aligned} \tag{10.22}$$

From now on keep in mind that AK is a constant depending on $n, p, L/\nu, q$; without loss of generality we assume $AK \geqslant 2$. Now, making a few elementary manipulations on (10.22) such as $\mu_1(\cdot), \mu_2(\cdot) \leqslant |B_{R_0}|$, and using Fubini's theorem, we estimate

$$\begin{aligned}
\int_{B_{R_0}} \big(\mu + |\mathfrak{X}u|\big)^q \, dx &\leqslant c \int_{B_{R_0}} \big[M^*\big(\mu^p + |\mathfrak{X}u|^p\big)\big]^{q/p} \, dx \\
&= c \int_0^{\infty} \lambda^{q/p-1} \mu_1(\lambda) \, d\lambda
\end{aligned}$$

$$= c\int_0^{\lambda_0} [\ldots]\,d\lambda + c\int_{\lambda_0}^{\infty} [\ldots]\,d\lambda$$

$$\leqslant c\lambda_0^{q/p}|B_{R_0}| + c\sum_{m=0}^{\infty} \int_{(AK)^m\lambda_0}^{(AK)^{m+1}\lambda_0} [\ldots]\,d\lambda$$

$$\leqslant c\lambda_0^{q/p}|B_{R_0}| + c\lambda_0^{q/p}\sum_{m=0}^{\infty}(AK)^{qm/p}\mu_1\big((AK)^m\lambda_0\big)$$

$$\overset{(10.20)–(10.22)}{\leqslant} c\lambda_0^{q/p}|B_{R_0}| + c\lambda_0^{q/p}\sum_{m=1}^{\infty}(AK)^{qm/p}\mu_2\big((AK)^m\varepsilon\lambda_0\big), \quad (10.23)$$

with $c \equiv c(n, p, L/\nu, q)$; moreover, (10.19) yields

$$\lambda_0^{q/p}|B_{R_0}| \leqslant c\bigg(\fint_{B_{100R_0}} \big(\mu^p + |\mathfrak{X}u|^p\big)\,dx\bigg)^{q/p}|B_{R_0}|. \quad (10.24)$$

In turn, again by means of Fubini's theorem and elementary manipulations, we have

$$\lambda_0^{q/p}\sum_{m=1}^{\infty}(AK)^{qm/p}\mu_2\big((AK)^m\varepsilon\lambda_0\big)$$

$$\leqslant \frac{AK}{\varepsilon^{q/p}(AK-1)}\int_0^{\infty}\lambda^{q/p-1}\mu_2(\lambda)\,d\lambda$$

$$\leqslant c\int_{B_{R_0}} \big[M^*_{q_0/p}\big(|F|^p\big)\big]^{q/p}\,dx \overset{(2.29)–(10.2)}{\leqslant} c\int_{B_{100R_0}} |F|^q\,dx, \quad (10.25)$$

where, taking into account the peculiar dependence of ε, AK, and also (10.2), it turns out that the constant c in the last line depends only on $n, p, L/\nu, q$. Connecting (10.25)–(10.24) to (10.23), we finally gain, after further elementary manipulations

$$\bigg(\fint_{B_{R_0}} |\mathfrak{X}u|^q\,dx\bigg)^{1/q} \leqslant c\bigg(\fint_{B_{100R_0}} \big(\mu^p + |\mathfrak{X}u|^p\big)\,dx\bigg)^{1/p} + c\bigg(\fint_{B_{100R_0}} |F|^q\,dx\bigg)^{1/q}. \quad (10.26)$$

We have used again, and repeatedly, the doubling condition (2.8); the constant c depends on n, p, L/ν, q, but not yet on $b(\cdot)$; the dependence on q is such that c blows-up only when $q \nearrow \infty$. Now notice that the only point to use a ball with small radius R_0 in the above argumentation was to fulfill the requirement $[b]^*_{100R_0} \leqslant \varepsilon$; therefore estimate (10.26) continues to hold with R_0 replaced by any other smaller radius, and therefore

$$\left(\fint_{B_{R_1}} |\mathfrak{X}u|^q \, dx \right)^{1/q} \leqslant c \left(\fint_{B_{100R_1}} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \right)^{1/p} + c \left(\fint_{B_{100R_1}} |F|^q \, dx \right)^{1/q} \tag{10.27}$$

holds whenever $R_1 \leqslant R_0$ and $B_{100R_1} \Subset \Omega$. Summarizing, we have obtained a first form of estimate (1.23), that is (10.27), which is valid for suitably small radii; moreover when estimating the left hand side with the right-hand one we pass to an integral supported on a ball with radius magnified of a factor 100. In order to derive the precise form (1.23) we can proceed using a standard covering argument at the end of which we shall get the desired estimate, where the constant c will be the one from (10.27), magnified of a factor equal to $c(n, p, q)(R/R_0)^{Q(q-p)/p}$. Since the radius R_0 has been chosen in order to verify $[b]^*_{100R_0} \leqslant \varepsilon$ the final dependence of c on $b(\cdot)$ will follow. We hereby sketch the covering argument; we first treat the most relevant case $R \geqslant R_0$. Consider a CC-ball $B_R \Subset \Omega'$ with $R \geqslant R_0$, and cover $B_{R/2}$ with a finite family of CC-balls $\{B_i\}$ with radius equal to $R_0/1000$, centered in $B_{R/2}$, and such that the enlarged balls have locally finite intersection in the following sense: every ball $100B_i$ touches at most $c(n)$ of the other ones $100B_j$, $i \neq j$. It clearly follows that $100B_i \Subset B_R$. The existence of such a family follows considering the structure of the CC-balls; see Section 2.3. We then apply (10.27) on every ball B_i – this means we are taking $R_1 = R_0/1000$ in (10.27) – and manipulate as follows:

$$\begin{aligned}
\fint_{B_{R/2}} |\mathfrak{X}u|^q \, dx &\leqslant c \left(\frac{R_0}{R}\right)^Q \sum_i \fint_{B_i} |\mathfrak{X}u|^q \, dx \\
&\leqslant c \left(\frac{R_0}{R}\right)^Q \sum_i \left(\fint_{100B_i} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \right)^{q/p} + c \left(\frac{R_0}{R}\right)^Q \sum_i \fint_{100B_i} |F|^q \, dx \\
&\leqslant c \left(\frac{R_0}{R}\right)^Q R_0^{-Qq/p} \left(\int_{B_R} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \right)^{(q-p)/p} \sum_i \int_{100B_i} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \\
&\quad + c \fint_{B_R} |F|^q \, dx \\
&\leqslant c \left(\frac{R}{R_0}\right)^{Q(q-p)/p} \left(\fint_{B_R} \left(\mu^p + |\mathfrak{X}u|^p\right) dx \right)^{q/p} + c \fint_{B_R} |F|^q \, dx,
\end{aligned} \tag{10.28}$$

where $c \equiv c(n, p, L/\nu, q)$. Therefore estimate (1.23) follows in the case $R_0 \leqslant R$. The case $R < R_0$ can be treated in a similar way, and it is actually almost contained in (10.27), where $R_1 \leqslant R_0$: we only need to pass from a ball $B_{R/2}$ to B_R instead of passing from $B_{R/100}$ to B_R as in (10.27). This fact can be done via the same covering argument used for the case $R_0 \leqslant R$, by covering $B_{R/2}$ by small balls with radius $R/1000$ and then perform the same computation as in (10.28); this time since the radius of the balls B_i is comparable to that of B_R, when passing from estimate (10.27) to (1.23) the constant will magnify of a factor that depends only on n, p, L/ν, q but independent of R_0. $\square$

Remark 10.1. The argument at the end of the last proof leads to a statement which is dual to the one in Theorem 1.4. Indeed it follows that for every $q < \infty$ there exists a constant $c \equiv c(n, p, L/\nu, q)$ and a positive radius $R_0 \equiv R_0(n, p, L/\nu, q, b(\cdot))$ such that (1.23) holds provided $R \leqslant R_0$; this is actually the content of (10.27). In this way the constant c is independent of $b(\cdot)$, while the dependence on $b(\cdot)$ in the final estimate is shifted in R_0, that is "the radius after which estimate (1.23) starts to hold."

Remark 10.2. The constant appearing in the estimate (1.23) blow-up when $p \nearrow 4$. As far as the dependence on q is concerned, from the proof given we see that c blows-up when $q \nearrow \infty$, as it must be, while it remains stable when $q \searrow p$. This last fact is basically a consequence of the use of Theorem 3.4 to prove (1.23) when q is "close" to p – see the beginning of the section – and of inequality (10.2) applied in (10.25), when q is "larger" than p.

11. More equations

This section should be considered as an appendix to the previous one in that we are describing here a few generalizations of the results contained there. To begin with we observe that the result of Theorem 1.4 extends to the case of solutions to more general equations of the type

$$\operatorname{div}_H a(x, \mathfrak{X}u) = \operatorname{div}_H\big(|F|^{p-2}F\big), \tag{11.1}$$

with the vector field $a : \Omega \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that

$$z \mapsto a(x, z) \quad \text{satisfies (1.2)–(1.3), for every } x \in \Omega, \tag{11.2}$$

and with continuous dependence on the x-variable, that is

$$\big|a(x, z) - a(y, z)\big| \leqslant L\omega\big(d_{CC}(x, y)\big)\big(\mu + |z|\big)^{p-1}, \tag{11.3}$$

is satisfied for every $z \in \mathbb{R}^{2n}$ and $x, y \in \Omega$, where $\omega : [0, \infty) \to [0, \infty)$ is a continuous, non-decreasing function such that $\omega(0) = 0$. The function $\omega(\cdot)$ is usually called "modulus of continuity." The proof of such an extension is very close to the ones already given in the previous section and we shall therefore confine ourselves to explaining the main differences, which occur in the following points.

When using Lemma 10.1 we shall consider as a comparison function v the unique solution of the Dirichlet problem

$$\begin{cases} \operatorname{div} a(x_0, \mathfrak{X}v) = 0 & \text{in } B_R, \\ v = u & \text{on } \partial B_R, \end{cases} \tag{11.4}$$

where x_0 is the center of B_R. At this point the statement and the proof of Lemma 10.1 are even simpler, as for instance they do not need the use of Theorem 3.4; for the ease of exposition we shall nevertheless refer to the already given proof although it may be shortened at some points. Anyway we remark that Theorem 3.4 continues to hold for solutions to (11.1) under the considered assumptions. Estimate (10.5) continues to hold in a different form, that is (11.5) below; this is due to the fact that the comparison estimate (10.7) in Lemma 10.1 has to be replaced by

$$\begin{aligned}\fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p\,dx &\leqslant c \fint_{B_R} \big\langle a(x_0, \mathfrak{X}u) - a(x_0, \mathfrak{X}v), \mathfrak{X}u - \mathfrak{X}v \big\rangle\,dx \\ &= c \fint_{B_R} \big\langle a(x_0, \mathfrak{X}u) - a(x, \mathfrak{X}u), \mathfrak{X}u - \mathfrak{X}v \big\rangle\,dx \\ &\quad + c \fint_{B_R} \big\langle |F|^{p-2}F, \mathfrak{X}u - \mathfrak{X}v \big\rangle\,dx =: I + II,\end{aligned}$$

which holds in view of (11.4). The estimation of I will be done this time using (11.3), the one for II being exactly as in (10.8). This finally yields the estimate

$$\begin{aligned}\fint_{B_R} |\mathfrak{X}u - \mathfrak{X}v|^p\,dx &\leqslant c_5 \omega^*(2R_0) \fint_{B_{2R}} \big(\mu + |\mathfrak{X}u|\big)^p\,dx \\ &\quad + c_5\big[1 + \omega^*(2R_0)\big]\bigg(\fint_{B_{2R}} |F|^{q_0}\,dx\bigg)^{p/q_0},\end{aligned} \tag{11.5}$$

where $\omega^*(\cdot) := [\omega(\cdot)]^{p/(p-1)}$. Once the comparison estimate is gained we may proceed as in the proof of Lemma 10.2 but using the assumption that $\omega^*(200R_0) < \varepsilon$ instead of $[b]^*_{100R_0} \leqslant \varepsilon$. Then, when using the comparison function v, it will be defined as the unique solution to (11.4) with $20B \equiv B_R$ and x_0 is the center of $20B$, while the use of (11.5) will replace the use of (10.5). This will give the proof of the new version of Lemma 10.2.

Then, proceeding exactly as in the proof of Theorem 1.4 we arrive at the following:

Theorem 11.1. *Let $u \in HW^{1,p}(\Omega)$ be a weak solution to Eq.* (11.1) *under the assumptions* (11.2)–(11.3) *with* $2 \leqslant p < 4$. *Assume that $F \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$ for some $q > p$; then $\mathfrak{X}u \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$. Moreover there exists a constant c, depending only on n, p, L/ν, q and the function $\omega(\cdot)$, such that the inequality* (1.23) *holds for any CC-ball $B_R \Subset \Omega$.*

Again, the dependence on $\omega(\cdot)$ in the a priori estimates of Theorem 11.1 can be replaced as described in Remark 10.1.

Remark 11.1. Theorem 1.4 admits an obvious reformulation in the case the coefficient function $b(\cdot)$ in (1.20) is assumed to have a properly small BMO norm instead of being locally in VMO. Referring to (2.13), the function $b(\cdot)$ is said to have bounded mean oscillations provided $[b]_{R,\Omega} < \infty$ for some $R > 0$. Now it is easy too see that in Theorem 1.4 assumption (1.21) can be replaced in order to have the following statement: For every $q < \infty$ there exists $\varepsilon > 0$ depending only on n, p, L/ν and q such that $[b]_{R,\Omega} < \varepsilon$ for some $R > 0$ implies $\mathfrak{X}u \in L^q_{\mathrm{loc}}(\Omega, \mathbb{R}^{2n})$. This comes directly from Lemma 10.2, where $[b]^*_{100R_0} \leqslant \varepsilon$, which is later implied by the VMO condition in the proof of Theorem 1.4, is now immediately implied by the global smallness assumption $[b]_{R,\Omega} < \varepsilon$.

Acknowledgments

G.M. is supported by MUR via the national project "Calcolo delle Variazioni," and by GNAMPA via the project "Singularities and regularity in non-linear potential theory." Part of this work was done while G.M. was visiting the Universities of Warsaw, Helsinki and Erlangen-Nürnberg, in May, June and August 2007, respectively. He wishes to thank all the members of the institutions for the nice hospitality. A.Z.-G. would like to thank her hosts at the Helsinki University of Technology and at the University of Jyväskylä. X.Z. is supported by the Academy of Finland, project 207288. The authors wish to thank Anna Föglein for remarks on a first version of the manuscript. Finally, last but not least, the authors wish to thank the referee for his/her suggestions aimed at improving the presentation of the paper.

References

[1] E. Acerbi, G. Mingione, Gradient estimates for a class of parabolic systems, Duke Math. J. 136 (2007) 285–320.
[2] A. Bellaïche, The tangent space in sub-Riemannian geometry, in: Sub-Riemannian Geometry, in: Progr. Math., vol. 144, Birkhäuser, Basel, 1996, pp. 1–78.
[3] M. Bildhauer, M. Fuchs, X. Zhong, A regularity theory for scalar local minimizers of splitting-type variational integrals, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 6 (2007) 385–404.
[4] M. Bramanti, L. Brandolini, L^p estimates for nonvariational hypoelliptic operators with VMO coefficients, Trans. Amer. Math. Soc. 352 (2000) 781–822.
[5] M. Bramanti, L. Brandolini, Schauder estimates for parabolic nondivergence operators of Hörmander type, J. Differential Equations 234 (2007) 177–245.
[6] L. Caffarelli, I. Peral, On $W^{1,p}$ estimates for elliptic equations in divergence form, Comm. Pure Appl. Math. 51 (1998) 1–21.
[7] L. Capogna, Doctoral dissertation, Purdue University, 1996.
[8] L. Capogna, Regularity of quasi-linear equations in the Heisenberg group, Comm. Pure Appl. Math. 50 (1997) 867–889.
[9] L. Capogna, Regularity for quasilinear equation and 1-quasiconformal maps in Carnot groups, Math. Ann. 313 (1999) 263–295.
[10] L. Capogna, D. Danielli, N. Garofalo, An embedding theorem and the Harnack inequality for nonlinear subelliptic equations, Comm. Partial Differential Equations 18 (1993) 1765–1794.
[11] 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 (2003) 1–40.
[12] S. Chanillo, J.J. Manfredi, Sharp global bounds for the Hessian on pseudo-Hermitian manifolds, preprint, 2006.
[13] F. Chiarenza, M. Frasca, P. Longo, $W^{2,p}$-solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients, Trans. Amer. Math. Soc. 336 (1993) 841–853.
[14] W.L. Chow, Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung, Math. Ann. 117 (1939) 98–105.
[15] E. DiBenedetto, J.J. Manfredi, On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems, Amer. J. Math. 115 (1993) 1107–1134.
[16] A. Domokos, Differentiability of solutions for the non-degenerate p-Laplacian in the Heisenberg group, J. Differential Equations 204 (2004) 439–470.
[17] A. Domokos, J.J. Manfredi, Subelliptic Cordes estimates, Proc. Amer. Math. Soc. 133 (2005) 1047–1056.
[18] A. Domokos, J.J. Manfredi, $C^{1,\alpha}$-regularity for p-harmonic functions in the Heisenberg group for p near 2, in: Contemp. Math., vol. 370, 2005, pp. 17–23.
[19] L. Esposito, F. Leonetti, G. Mingione, Regularity for minimizers of irregular integrals with (p,q)-growth, Forum Math. 14 (2002) 245–272.
[20] L. Esposito, F. Leonetti, G. Mingione, Sharp regularity for functionals with (p,q)-growth, J. Differential Equations 204 (2004) 5–55.
[21] A. Föglein, Partial regularity results for sub-elliptic systems in the Heisenberg group, Calc. Var. Partial Differential Equations 32 (2008) 25–51.
[22] G.B. Folland, Applications of analysis on nilpotent groups to partial differential equations, Bull. Amer. Math. Soc. 83 (1977) 912–930.
[23] G.B. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups, Ark. Mat. 13 (1975) 161–207.
[24] E. Giusti, Direct Methods in the Calculus of Variations, World Scientific, 2003.
[25] P. Goldstein, A. Zatorska-Goldstein, Calderón–Zygmund type estimates for non-linear systems with quadratic growth on the Heisenberg group, Forum Math. 20 (2008) 679–710.

[26] P. Hajlasz, P. Koskela, Sobolev met Poincaré, Mem. Amer. Math. Soc. 145 (688) (2000), x+101 pp.
[27] P. Hajlasz, P. Strzelecki, Subelliptic p-harmonic maps into spheres and the ghost of Hardy spaces, Math. Ann. 312 (1998) 341–362.
[28] C. Hamburger, Regularity of differential forms minimizing degenerate elliptic functionals, J. Reine Angew. Math. (Crelle's J.) 431 (1992) 7–64.
[29] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1967) 147–171.
[30] T. Iwaniec, Projections onto gradient fields and L^p-estimates for degenerated elliptic operators, Studia Math. 75 (1983) 293–312.
[31] T. Iwaniec, C. Sbordone, Weak minima of variational integrals, J. Reine Angew. Math. (Crelle's J.) 454 (1994) 143–161.
[32] D. Jerison, The Poincaré inequality for vector fields satisfying Hörmander's condition, Duke Math. J. 53 (1986) 503–523.
[33] J. Kinnunen, N. Shanmugalingam, Regularity of quasi-minimizers on metric spaces, Manuscripta Math. 105 (2001) 401–423.
[34] J. Kinnunen, S. Zhou, A local estimate for nonlinear equations with discontinuous coefficients, Comm. Partial Differential Equations 24 (1999) 2043–2068.
[35] J.J. Kohn, Pseudo-differential operators and hypoellipticity, in: Partial Differential Equations, in: Proc. Sympos. Pure Math., vol. XXIII, Univ. California, Berkeley, CA, 1971, pp. 61–69.
[36] J. Kristensen, G. Mingione, The singular set of minima of integral functionals, Arch. Ration. Mech. Anal. 180 (2006) 331–398.
[37] O.A. Ladyzhenskaya, N.N. Ural'tseva, Linear and Quasilinear Elliptic Equations, Academic Press, New York/ London, 1968.
[38] G. Lu, Embedding theorems into Lipschitz and BMO spaces and applications to quasilinear subelliptic differential equations, Publ. Mat. 40 (1996) 301–329.
[39] G. Lu, The sharp Poincaré inequality for free vector fields: An endpoint result, Rev. Mat. Iberoamericana 10 (1994) 453–466.
[40] J.J. Manfredi, G. Mingione, Regularity results for quasilinear elliptic equations in the Heisenberg group, Math. Ann. 339 (2007) 485–544.
[41] P. Marcellini, Regularity of minimizers of integrals of the calculus of variations with non standard growth conditions, Arch. Ration. Mech. Anal. 105 (1989) 267–284.
[42] S. Marchi, $C^{1,\alpha}$ local regularity for the solutions of the p-Laplacian on the Heisenberg group for $2 \leqslant p \leqslant 1+\sqrt{5}$, Z. Anal. Anwend. 20 (2001) 617–636, Erratum: Z. Anal. Anwend. 22 (2003) 471–472.
[43] G. Mingione, The Calderón–Zygmund theory for elliptic problems with measure data, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 6 (2007) 195–261.
[44] A. Nagel, E.M. Stein, S. Wainger, Balls and metrics defined by vector fields. I. Basic properties, Acta Math. 155 (1985) 103–147.
[45] M.C. Shaw, L. Wang, Hölder and L^p estimates for $\Box_b$ on CR manifolds of arbitrary codimension, Math. Ann. 331 (2005) 297–343.
[46] E.M. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton Math. Ser., vol. 43, Princeton University Press, Princeton, NJ, 1993.
[47] A. Zatorska-Goldstein, Very weak solutions of nonlinear subelliptic equations, Ann. Acad. Sci. Fenn. Math. 30 (2005) 407–436.