REGULARITY OF QUASI-LINEAR EQUATIONS WITH HÖRMANDER VECTOR FIELDS OF STEP TWO

GIOVANNA CITTI AND SHIRSHO MUKHERJEE

ABSTRACT. If the smooth vector fields $X_1, \dots, X_m$ and their commutators span the tangent space at every point in $\Omega \subseteq \mathbb{R}^N$ for any fixed $m \leq N$, then we establish the full interior regularity theory of quasi-linear equations $\sum_{i=1}^m X_i^* A_i(X_1 u, \dots, X_m u) = 0$ with p-Laplacian type growth condition. In other words, we show that a weak solution of the equation is locally $C^{1,\alpha}$.

CONTENTS

1. Introduction 1
2. Notations and Preliminaries 4
3. Apriori Estimates 8
4. Local Lipschitz Continuity 20
5. The local $C^{1,\alpha}$ regularity 23
Acknowledgement 45
References 45

1. INTRODUCTION

The development of partial differential equations arising from families of noncommuting vector fields has been conspicuous since the fundamental work of Hörmander [29]. Such equations, nowadays referred to as *sub-elliptic equations*, are not elliptic in general. Nevertheless, they have a remarkable structural similarity with elliptic equations. This has been evident in the theory that sub-elliptic linear equations entail, where the given non-commuting vector fields $X_1, \dots, X_m$ on a domain $\Omega \subseteq \mathbb{R}^N$ satisfy the, so-called, Hörmander's rank condition

$$\dim\big(\mathtt{Lie}(X_1, \dots, X_m)\big) = N, \tag{1.1}$$

at every point in Ω. It was first used in [29] to show that the linear operator $\mathcal{L}u = \sum_{i=1}^m X_i^* X_i u$, called sub-elliptic Laplacian or *sub-Laplacian*, is hypoelliptic. Later, it was found that it also satisfies estimates of the form

$$\|u\|_{\mathcal{S}^{k+2,p}} \leq c(N,p)\big(\|\mathcal{L}u\|_{\mathcal{S}^{k,p}} + \|u\|_{\mathcal{S}^{k,p}}\big),$$

where $\mathcal{S}^{k,p}$ are Sobolev spaces defined by the vector fields (see Section 2 for details), which clearly shows the resemblance to the corresponding elliptic estimates. This has been unveiled at increasing levels of generality by Folland-Stein [21], Folland [20] and Rothschild-Stein [43]. Their methods also reflect the importance of the special case when the vector fields are left-invariant with respect to a Nilpotent Lie group. The Hörmander's condition (1.1) endows a stratification on the Lie algebra of such groups and they are called Carnot Groups (see [1]). Later, it has been shown by Nagel-Stein-Wainger [41] (see also Sánchez-Calle [44]) that in the general case,

Date: July 8, 2022.
2020 *Mathematics Subject Classification.* Primary 35H20, 35J92, 35B65.
Key words and phrases. Sub-elliptic equations, quasilinear, p-Laplacian, regularity.

the Carnot-Carathèodory metric is equivalent to the natural metrics intrinsic to the given vector fields, that are generalizations of the homogeneous metrics of Carnot groups.

These subsequently lead to the development of an extensive literature and nowadays, the theory of linear sub-elliptic equations is largely established. It naturally leads to the inquiry for a regularity theory of quasi-linear equations in the sub-elliptic setting, similar to that of quasi-linear elliptic equations. This area has remained relatively open over the years and the present paper is aimed towards this pursuit.

Let $\Omega \subset \mathbb{R}^N$ be an open, simply connected bounded domain and $X_1, \dots, X_m$ be smooth vector fields with $m \leq N$ such that they, along with their commutators, span the tangent space at every point in Ω. This is a special case of (1.1), referred to as the "step two hypothesis" by Nagel-Stein [40], where commutators up to second-order are considered in place of the full Lie algebra generated by the vector fields (see (2.1)). We consider the equation

$$\operatorname{div}_H(A(\mathfrak{X}u)) = \sum_{i=1}^{m} X_i^* A_i(\mathfrak{X}u) = 0 \qquad \text{in } \Omega, \tag{1.2}$$

where $\mathfrak{X}u = (X_1 u, \dots, X_m u)$ is the sub-elliptic gradient, X_i^* is the formal adjoint of X_i and $A : \mathbb{R}^m \to \mathbb{R}^m$ is a given C^1-function that satisfies the following structure condition,

$$\begin{aligned} (\delta + |z|^2)^{\frac{p-2}{2}} |\xi|^2 \leq \langle DA(z)\,\xi, \xi \rangle \leq L(\delta + |z|^2)^{\frac{p-2}{2}} |\xi|^2; \\ |A(z)| \leq L\,|z|(\delta + |z|^2)^{\frac{p-2}{2}}, \end{aligned} \tag{1.3}$$

for every $z, \xi \in \mathbb{R}^m$, with $1 < p < \infty$, $L \geq 1$ and $\delta \geq 0$. The equation (1.2) with (1.3) is modelled on the degenerated sub-elliptic equation

$$\operatorname{div}_H \left((\delta + |\mathfrak{X}u|^2)^{\frac{p-2}{2}} \mathfrak{X}u \right) = 0, \tag{1.4}$$

which, for $\delta = 0$, is called the sub-elliptic p-Laplacian equation $\operatorname{div}_H(|\mathfrak{X}u|^{p-2}\mathfrak{X}u) = 0$.

In the Euclidean setting, it is known that weak solutions of the classical p-Laplacian equation $\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0$ are locally $C^{1,\alpha}$, as shown independently by DiBenedetto [10], Lewis [32] and Tolksdorf [45] using techniques descending from the seminal work of De Giorgi [9]. Hence, for the sub-elliptic case, $C^{1,\alpha}$-regularity of weak solutions with respect to the metrics introduced in [41] is likewise expected. The sub-elliptic Poincaré inequality was first proved by Jerison [30] which leads to the Sobolev embedding theorem which has been later shown in different levels of generality by Garofalo-Nhieu [24], Franchi-Lu-Wheeden [23], Hajlasz-Koskela [28], etc. However, obtaining higher regularity of the gradient $\mathfrak{X}u$ using the classical techniques is quite difficult even in the case of step two, due to the non-commutativity of the vector fields. For previous results and similar theories in this direction, we refer to [2, 6, 7, 19, 13, 14, 35, 34, 12, 36, 42, 4] and references therein. Recently, the problem has been resolved for the case of the Heisenberg Group (see [47, 39]), which is the most trivial example of a step two Carnot group where the vector space of commutators is one dimensional. The case of non-Nilpotent groups like $\mathbb{SU}(3)$ and compact semi-simple Lie groups has been considered in [16]. In this paper, we prove the $C^{1,\alpha}$-regularity for the most general step two case.

First, we show that the sub-elliptic gradient is locally bounded as our first main result.

Theorem 1.1. *If $u \in \mathcal{S}^{1,p}(\Omega)$ is a weak solution of equation (1.2) equipped with the structure condition (1.3) for any $1 < p < \infty$ and $\delta \geq 0$, then $\mathfrak{X}u \in L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$. Moreover, for any $x_0 \in \Omega$ and metric ball $B_r = B_r(x_0) \subset \Omega$, we have*

$$\sup_{B_{\sigma r}} |\mathfrak{X}u| \leq \frac{c}{(1-\sigma)^{Q/p}} \left(\fint_{B_r} \left(\delta + |\mathfrak{X}u|^2\right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \tag{1.5}$$

for any $0 < \sigma < 1$, where $c = c(N, p, L) > 0$ and $Q = Q(x_0) = \liminf_{s \to 0^+} \frac{\log(|B_s(x_0)|)}{\log(s)}$.

The following Corollary is an easy consequence of Theorem 1.1.

Corollary 1.2. *If $u \in \mathcal{S}^{1,p}(\Omega)$ is a weak solution of equation (1.2) equipped with the structure condition (1.3) for any $1 < p < \infty$ and $\delta \geq 0$, then $u \in C^{0,1}_{\rm loc}(\Omega)$. Moreover, for any metric ball $B_r \subset \Omega$, we have*

$$|u(x) - u(y)| \leq c\bigg(\fint_{B_r} \big(\delta + |\mathfrak{X}u|^2\big)^{\frac{p}{2}}\, dx \bigg)^{\frac{1}{p}} d_c(x,y) \tag{1.6}$$

for any $x, y \in B_r$, where $c = c(N,p,L) > 0$ and d_c is the Carnot-Carathèodory metric.

Second, we show that the sub-elliptic gradient is locally Hölder continuous with respect to the Carnot-Carathèodory metric. The following theorem is the second main result of this paper.

Theorem 1.3. *If $u \in \mathcal{S}^{1,p}(\Omega)$ is a weak solution of equation (1.2) equipped with the structure condition (1.3) for any $1 < p < \infty$ and $\delta \geq 0$, then there exists a positive $\alpha = \alpha(N,p,L) \leq 1$ such that $\mathfrak{X}u \in C^{0,\alpha}_{\rm loc}(\Omega, \mathbb{R}^m)$. Moreover, for any concentric metric balls $B_r \subset B_{r_0} \subset \Omega$, we have*

$$\max_{1\leq i\leq m} \operatorname{osc}_{B_r} X_i u \leq c\Big(\frac{r}{r_0}\Big)^{\alpha} \bigg(\fint_{B_{r_0}} \big(\delta + |\mathfrak{X}u|^2\big)^{\frac{p}{2}}\, dx \bigg)^{\frac{1}{p}} \tag{1.7}$$

for any $0 < r < r_0$, where $c = c(N,p,L) > 0$ is a constant.

The following Corollary is an easy consequence of Theorem 1.3 and Schauder estimates [46].

Corollary 1.4. *If $u \in \mathcal{S}^{1,p}(\Omega)$ is a weak solution of equation (1.2) equipped with the structure condition (1.3) for any $1 < p < \infty$ and $\delta > 0$, moreover if A is smooth, then $u \in C^{\infty}(\Omega)$.*

The significance of this paper lies mainly in the fact that it shows the left invariance of the vector fields with respect to a Lie group is superfluous for the purpose of regularity. As reflected by the results in this paper, all necessary techniques and estimates are independent of any Lie group structure on $\mathbb{R}^N$. Therefore, the commutation relations between the vector fields are quite arbitrary within the restriction of the step two hypothesis. This is much more general than that for the case of the Heisenberg Group or any other Carnot groups of step two.

In order to get the full regularity, the general structure of the vector fields makes it necessary to obtain Caccioppoli type estimates of every possible type by testing the equations satisfied by $X_i u$ and $[X_i, X_j]u$ for all i, j with every combination of the vector fields and their commutators. The interplay between the estimates leading to the proofs of the theorems is similar to the earlier works. The guiding principle remains that the behavior of homogeneous sub-elliptic equations (1.2) is similar to that of inhomogeneous elliptic equations with an integrable right-hand side (this role is played by the extra terms containing the commutators for the sub-elliptic case). Hence, the high integrability of the commutators (Corollary 3.9 and similar others) allows the use of De Giorgi type techniques. The double truncation $v = \min\big(\mu(r)/8, \max(\mu(r)/4 - X_l u, 0)\big)$ with $\mu(r) = \max_{1\leq i\leq m} \sup_{B_r} |X_i u|$, has been used within the estimates particularly to deal with the singular case $1 < p < 2$, similarly as in [39]. This argument originates in the earlier works [45, 33], etc. The Riemannian approximation is used for the local Lipschitz continuity.

The paper is organized as follows. All necessary notations and preliminaries are provided in Section 2 and the apriori Caccioppoli type estimates are proved in Section 3. Then, the proof of Theorem 1.1 is provided in Section 4. In Section 5, further requisite estimates are obtained that lead to the proof of Theorem 1.3 in the end.

Finally, we remark that the present regularity results seem to be the best possible ones within the framework of the existing techniques and methods. The estimates of this paper use properties that are specific to step two in subtle ways which, apparently do not have any immediate generalization for higher-order commutators. We believe the regularity theory involving vector fields satisfying the general Hörmander's condition (1.1) would require new fundamental ideas.

2. Notations and Preliminaries

In this section, we develop relevant notations and provide basic properties and previously known results that shall be used throughout this paper.

Here onwards we assume the set of smooth vector fields $\{X_1, \dots, X_m\}$ defined on $\Omega \subseteq \mathbb{R}^N$ for some integer $m \leq N$, satisfy the step two hypothesis of Nagel-Stein [40]; precisely,

$$X_1, \dots, X_m \quad \text{and} \quad [X_j, X_k]\ \forall\, j,k \in \{1.\dots, m\} \quad \text{span the tangent space at each } x_0 \in \Omega. \tag{2.1}$$

The nontrivial case occurs if $m < N$ because the case of equality is identical to the familiar Euclidean setting. We choose an enumeration of the commutators as $\{T_1, \dots, T_n\}$ such that $\{X_1, \dots, X_m, T_1, \dots, T_n\}$ span the whole tangent space, for some integer $0 < n \leq m(m-1)/2$.

For a function $u : \Omega \to \mathbb{R}$, let us denote the gradients as

$$\mathfrak{X}u = (X_1u, \dots, X_mu) \quad \text{and} \quad Tu = (T_1u, \dots, T_nu).$$

The Hessian is denoted as $\mathfrak{X}\mathfrak{X}u = (X_jX_iu)_{i,j}$, similarly $\mathfrak{X}Tu = (X_jT_ku)_{k,j}$ and $T\mathfrak{X}u = (T_kX_ju)_{j,k}$ which are respectively $m \times n$ and $n \times m$ matrices. The commutator matrices are denoted as $[\mathfrak{X}, \mathfrak{X}]u = ([X_j, X_i]u)_{i,j}$ and similarly $[\mathfrak{X}, T]$ and $[T, \mathfrak{X}]$. The *sub-elliptic divergence* is defined as

$$\operatorname{div}_H(F) = \sum_{j=1}^{m} X_j^* f_j, \qquad \forall\ F = (f_1, \dots, f_m).$$

The Euclidean vector fields on $\mathbb{R}^m$ are denoted as D_i and $D = (D_1, \dots, D_m)$ is the Euclidean gradient; sometimes D_i is also denoted as ∂_{x_i} or something analogous in terms of coordinates.

We note that the step two hypothesis (2.1) implies

$$[X_i, X_j] = \sum_{k=1}^{n} \alpha_{i,j}^{(k)} T_k + \sum_{l=1}^{m} \beta_{i,j}^{(l)} X_l \quad \text{and} \quad [X_i, T_j] = \sum_{l=1}^{m} \gamma_{i,j}^{(l)} X_l + \sum_{k=1}^{n} \delta_{i,j}^{(k)} T_k, \tag{2.2}$$

and with appropriate choice of scaling in the norm of vector fields, we can assume without loss of generality that

$$\max_{i,j} \sum_{k,l} \big(\|\alpha_{i,j}^{(k)}\|_{L^\infty} + \|\beta_{i,j}^{(l)}\|_{L^\infty} + \|\gamma_{i,j}^{(l)}\|_{L^\infty} + \|\delta_{i,j}^{(k)}\|_{L^\infty} \big) \leq c(N), \quad \text{for some} \quad c(N) > 0, \tag{2.3}$$

and from smoothness, we can also conclude boundedness $\|\mathfrak{X}\alpha_{i,j}^{(k)}\|_{L^\infty} + \|T\alpha_{i,j}^{(k)}\|_{L^\infty} \leq c(N)$ and similar bounds for the others. Therefore, some basic inequalities are in order. Using (2.2) twice successively along with (2.3) and the bounds, we have the following inequalities :

$$(i) \qquad |[X_i, T_k]X_ju| \leq c\big(|\mathfrak{X}\mathfrak{X}u| + |T\mathfrak{X}u|\big) \leq c\big(|\mathfrak{X}\mathfrak{X}u| + |\mathfrak{X}Tu| + |\mathfrak{X}u| + |Tu|\big); \tag{2.4}$$

$$(ii) \qquad \big|[[X_i, X_l], X_j]u\big| \leq c\big(|[T, \mathfrak{X}]u| + |[\mathfrak{X}, \mathfrak{X}]u| + |\mathfrak{X}u| + |Tu|\big) \leq c\big(|\mathfrak{X}u| + |Tu|\big), \tag{2.5}$$

for every $i, j, l \in \{1, \dots, m\}$ and $k \in \{1, \dots, n\}$, where $c = c(N) > 0$ is a constant. Furthermore, we have the inequality

$$|Tu| \leq c|\mathfrak{X}\mathfrak{X}u|, \tag{2.6}$$

for some constant $c = c(N) > 0$, since $T_k = [X_l, X_j]$ for some $l, j \in \{1, \dots, m\}$.

2.1. Sub-elliptic metrics.

Let us recall the *Carnot-Carathéodory metric* (CC-metric) which is defined in terms of the length $\ell(\gamma)$ of curves $\gamma : [0,1] \to \Omega$, as

$$d_c(x,y) = \inf \{\ell(\gamma) : \gamma \in \Gamma(x,y)\}, \tag{2.7}$$

where the class Γ of absolutely continuous curves is given by

$$\Gamma(x,y) = \big\{\gamma \in AC\big([0,1]; \Omega\big) : \gamma(0) = x, \gamma(1) = y,\ \gamma'(t) \in \operatorname{span}\{X_1|_{\gamma(t)}, \dots, X_m|_{\gamma(t)}\}\big\}.$$

Chow's connectivity theorem (see [8]) gurantees $\Gamma(x,y) \neq \emptyset$ so that d_c is finite for any $x, y \in \mathbb{R}^N$. However, to determine the volume of the metric balls, another metric is required that is intrinsic

to the vector fields, as shown in [41]. Let $\{Y_1, \ldots, Y_{m+n}\} = \{X_1, \ldots, X_m, T_1, \ldots, T_n\}$. We observe that we can associate an integer, called formal degree $d_j = \deg(Y_j) \in \{1, 2\}$; indeed, we define that $\deg(Y_j) = 1$ if $Y_j \in \text{span}\{X_1, \ldots, X_m\}$ and $\deg(Y_j) = 2$ otherwise. Then, it is not difficult to see that (2.2) and (2.1) respectively imply:

(a) For every j, k we have $[Y_j, Y_k] \in \text{span}\{Y_l : d_l \leq d_j + d_k\}$;
(b) $Y_1, \ldots, Y_{m+n}$ span the whole tangent space.

Then, we can define following metric as in Nagel-Stein-Wainger [41]. For any $r > 0$, let $\Gamma(r)$ be the class of absolutely continuous curves $\gamma : [0, 1] \to \Omega$ satisfying the differential equation

$$\gamma'(t) = \sum_j a_j(t) Y_j(\gamma(t)) \quad \text{with} \quad |a_j(t)| < r^{d_j},$$

then the metric is defined as

$$\rho(x, y) = \inf \{r > 0 : \exists \gamma \in \Gamma(r), \gamma(0) = x, \gamma(1) = y\}. \tag{2.8}$$

Proposition 1.1 in [41], imply that within compact subsets of Ω, we have

$$c_1|x - y| \leq \rho(x, y) \leq c_2|x - y|^{1/2} \qquad \forall\, x, y \in \Omega \subseteq \mathbb{R}^N \tag{2.9}$$

for some positive constants c_1, c_2. Let us denote the metric balls $B_r(x) = \left\{y \in \mathbb{R}^N : \rho(x, y) < r\right\}$ and it has been shown in [41] that they follow all standard properties. In general $m + n \geq N$ and it is possible to choose a local basis $Y_{i_1}, \ldots, Y_{i_N}$ for the tangent space. For every such $I = (i_1, \ldots, i_N)$, letting $d(I) = d_{i_1} + \ldots + d_{i_N}$ and

$$\Lambda(x, r) = \sum_I |\lambda_I(x)| r^{d(I)},$$

where $\lambda_I(x) = \det(a_{j,k}(x))$ whenever $Y_{i_j} = \sum_k a_{j,k} \partial_{x_k}$, one can estimate the volume of the balls. Thus, $\Lambda \neq 0$ whenever $Y_{i_1}, \ldots, Y_{i_N}$ is a basis for the tangent space. Let $|E|$ be the Lebesgue measure of E for any $E \subset \mathbb{R}^N$. Then, from Theorem 1 in [41], there exists $c_2 > c_1 > 0$ such that we have

$$c_1 \Lambda(x, r) \leq |B_r(x)| \leq c_2 \Lambda(x, r)$$

in compact subsets of Ω. Therefore, let us denote

$$Q(x) = \liminf_{r \to 0^+} \frac{\log(|B_r(x)|)}{\log(r)}, \tag{2.10}$$

so that we have $c_1' r^Q \leq |B_r(x)| \leq c_2' r^Q$ in compact subsets of Ω for some $c_2' > c_1' > 0$ and $Q = Q(x)$. Since, Λ is a polynomial in r of fixed degree, hence Q is lower semi-continuous. It is a deep result of Nagel-Stein-Wainger [41, Theorem 4], that the metric ρ is equivalent to the CC-metric d_c. Henceforth, in the rest of the paper, we shall denote CC-metric balls by B_r without loss of generality and Q as the local dimension.

Remark 2.1 (Step 2 Carnot Groups). For a special case, when the vector fields are left invariant with respect to a Nilpotent Lie Group on $\mathbb{R}^N$ along with the hypothesis (2.1), then it is called a Carnot Group of step 2. In that case, if $X_1, \ldots, X_m$ are linearly independent, then we have $m + n = N$, the metric ρ is equivalent to the metric induced by the homogeneous norm of the group, and $Q = m + 2n$. Also, the commutation relations are more special than (2.2), we have $[X_i, T_k] = 0$ and $[X_i, X_j] = \sum_{k=1}^n \alpha_{i,j}^{(k)} T_k$ where $\alpha_{i,j}^{(k)}$'s are precisely prescribed by the composition law of the group. Furthermore, if it is a free group then we have $n = m(m-1)/2$.

The Hölder classes $C^{k,\alpha}$ (also denoted as $\Gamma^{k,\alpha}$) has been introduced by Folland and Stein, see [20, 22]. The functions in these classes are Hölder continuous with respect to the CC-metric.

Definition 2.2 (Folland-Stein classes)**.** Given an open set $\Omega \subset \mathbb{R}^N$, a function $u : \Omega \to \mathbb{R}$ and $0 < \alpha \leq 1$, we say $u \in C^{0,\alpha}(\Omega)$ if there exists a constant $M > 0$, such that

$$|u(x) - u(y)| \leq M\, d_c(x,y)^\alpha, \qquad \forall\, x, y \in \Omega.$$

It is a Banach space with the norm $\|u\|_{C^{0,\alpha}(\Omega)} = \|u\|_{L^\infty(\Omega)} + [u]_{C^{0,\alpha}(\Omega)}$, where the Hölder semi-norm is defined as

$$[u]_{C^{0,\alpha}(\Omega)} = \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{d_c(x,y)^\alpha}. \tag{2.11}$$

For any $k \in \mathbb{N}$, the spaces $C^{k,\alpha}(\Omega)$ are defined inductively as follows: we say that $u \in C^{k,\alpha}(\Omega)$ if $X_i u \in C^{k-1,\alpha}(\Omega)$ for every $i \in \{1, \dots, m\}$.

Definition 2.3 (Sub-elliptic Sobolev spaces)**.** Given an open set $\Omega \subset \mathbb{R}^N$, a function $u : \Omega \to \mathbb{R}$ and $1 \leq p < \infty$, the sub-elliptic Sobolev spaces are defined as

$$\mathcal{S}^{1,p}(\Omega) = \{u \in L^p(\Omega) : X_j u \in L^p(\Omega),\ \forall\, j \in \{1, \dots, m\}\},$$

which is a Banach space with the norm $\|u\|_{\mathcal{S}^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\mathfrak{X} u\|_{L^p(\Omega,\mathbb{R}^m)}$. The subspaces $\mathcal{S}^{1,p}_{\mathrm{loc}}(\Omega)$ and $\mathcal{S}^{1,p}_0(\Omega)$ are defined as usual. Moreoever, for any $k \in \mathbb{N}$, the space $\mathcal{S}^{k,p}(\Omega)$ is also defined by standard extension.

For any metric ball $B_r = B_r(x_0)$, the following Poincaré inequality for $u \in \mathcal{S}^{1,p}(B_r)$

$$\Big(\fint_{B_r} |u - (u)_{B_r}|^p \, dx \Big)^{1/p} \leq c\, r \Big(\fint_{B_r} |\mathfrak{X} u|^p \, dx \Big)^{1/p} \tag{2.12}$$

holds, see [30, 28]. This leads to the following Sobolev inequality,

$$\left(\fint_{B_r} |v|^{\frac{Qq}{Q-q}} \, dx \right)^{\frac{Q-q}{Qq}} \leq c\, r \left(\fint_{B_r} |\mathfrak{X} v|^q \, dx \right)^{\frac{1}{q}}, \tag{2.13}$$

for all $v \in \mathcal{S}^{1,q}_0(B_r)$ and $1 < q < Q$ with $Q = Q(x_0)$.

2.2. **Sub–elliptic equations.** If $u : \Omega \to \mathbb{R}$ is a weak solution of (1.2), then for every test function $\phi \in C^\infty_0(\Omega)$, we have

$$\int_\Omega \big\langle A(\mathfrak{X} u), \mathfrak{X}\phi \big\rangle \, dx = \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X} u) X_i \phi \, dx = 0. \tag{2.14}$$

Furthermore, for all non-negative $\phi \in C^\infty_0(\Omega)$, if the integral above is positive (resp. negative) then u is called a weak supersolution (resp. subsolution) of the equation (1.2).

It is easy to see that (1.3) implies the following ellipticity condition

$$\big\langle A(z), z \big\rangle \geq c(p)\, |z|^2 (\delta + |z|^2)^{\frac{p-2}{2}} \tag{2.15}$$

and also the following strong monotonicity inequality

$$\big\langle A(z) - A(w), z - w \big\rangle \geq \frac{1}{L'} (\delta + |z|^2 + |w|^2)^{\frac{p-2}{2}} |z - w|^2 \tag{2.16}$$

for all $z, w \in \mathbb{R}^m$, where $c(p) > 0$ and $L' = L'(p, L) > 0$.

The inequalities (2.15) and (2.16) can be used to prove the existence of a weak solution of equation (1.2), by standard variational methods. In fact, given any $u_0 \in \mathcal{S}^{1,p}(\Omega)$, there exists a unique weak solution $u \in \mathcal{S}^{1,p}(\Omega)$ of the following Dirichlet problem

$$\begin{cases} \mathrm{div}_H(A(\mathfrak{X} u)) = 0 \ \text{ in } \Omega; \\ u - u_0 \in \mathcal{S}^{1,p}_0(\Omega). \end{cases} \tag{2.17}$$

The uniqueness follows from the comparison principle, i.e. if u and v respectively are weak super and subsolution of the equation (1.2) and $u \geq v$ on $\partial\Omega$ in the trace sense, then we have $u \geq v$ a.e. in Ω. This can be shown easily by appropriate test fuctions on the equation (2.17).

Remark 2.4 (Divergence-free vector fields). Given a vector field X, its formal adjoint with respect to the Euclidean volume form (Lebesgue measure) is defined as X^*, so that for any $v \in C_0^\infty(\Omega)$, we have

$$\int_\Omega v X^* u \, dx = \int_\Omega u X v \, dx, \tag{2.18}$$

where $X^* u = -Xu - \mathrm{div}(X)u$. For simplicity, we assume that the given vector fields $X_1, \dots, X_m$ are divergence-free, i.e. $\mathrm{div}(X_i) = 0$. Hence, we have that $T_1, \dots, T_n$ are also divergence-free, from the formula $\mathrm{div}([X_i, X_j]) = X_i(\mathrm{div}(X_j)) - X_j(\mathrm{div}(X_i))$. The general case is a perturbation by a lower order term and can be obtained with minor modification. Moreover, in particular, if the X_i's are left-invariant with respect to a Lie group then they are always divergence-free with respect to the Haar measure of the group which is the Lebesgue measure up to multiples. In more general cases if the vector fields are defined on a manifold and the adjoint defined with respect to a volume form, then with appropriate geometric conditions (e.g. metrics with certain curvature conditions or manifolds with Killing forms, contact forms, etc.) divergence-free frames can exist and it is possible to write the equation (1.2) with respect to such frames having structure conditions similar to (1.3). We refer to [4] for details.

Henceforth, we assume $X_1, \dots, X_m$ are divergence-free, so that $X_i^* = -X_i$ and $T_k^* = -T_k$.

Remark 2.5 (The case of $\delta = 0$). If apriori estimates are available for the case of $\delta > 0$ and if the constants are independent of δ, then the case $\delta = 0$ follows as well by a standard regularization technique. A brief outline is as follows.

Let A satisfy the structure condition (1.3) for the case $\delta = 0$, i.e. for all $z, \xi \in \mathbb{R}^m$,

$$\begin{gathered} |z|^{p-2}|\xi|^2 \leq \langle DA(z)\,\xi, \xi\rangle \leq L|z|^{p-2}|\xi|^2; \\ |A(z)| \leq L\,|z|^{p-1}. \end{gathered} \tag{2.19}$$

There are many ways to define a regularization. In particular, for any $0 < \delta < 1$, one can define

$$A_\delta(z) = \big(1 - \eta_\delta(|z|)\big)A(z) + \eta_\delta(|z|)(\delta + |z|^2)^{\frac{p-2}{2}} z \tag{2.20}$$

where $\eta_\delta \in C^{0,1}([0,\infty))$ can be chosen (see [38]) such that $A_\delta \to A$ uniformly on compact subsets of $\mathbb{R}^m$ as $\delta \to 0^+$ and the structure condition

$$\begin{gathered} \frac{1}{L'}(\delta + |z|^2)^{\frac{p-2}{2}}|\xi|^2 \leq \langle DA_\delta(z)\,\xi, \xi\rangle \leq L'(\delta + |z|^2)^{\frac{p-2}{2}}|\xi|^2; \\ |A_\delta(z)| \leq L'\,|z|(\delta + |z|^2)^{\frac{p-2}{2}}, \end{gathered} \tag{2.21}$$

holds for some $L' = L'(p, L) > 1$. Given a weak solution $u \in \mathcal{S}^{1,p}(\Omega)$ of (1.2), we consider u_δ as the weak solution of the following regularized equation

$$\begin{cases} \mathrm{div}_H(A_\delta(\mathfrak{X}u_\delta)) = 0 \ \text{ in } \Omega'; \\ u_\delta - u \in \mathcal{S}_0^{1,p}(\Omega'), \end{cases} \tag{2.22}$$

for any $\Omega' \subset\subset \Omega$. Now, for the case of $\delta > 0$ whenever we obtain uniform estimates for $\mathfrak{X}u_\delta$ (with constants independent of δ), taking limit $\delta \to 0^+$, one can show that (1.5) and (1.7) also hold for the case of $\delta = 0$.

Therefore, we shall assume $\delta > 0$ without loss of generality, hereafter.

3. Apriori Estimates

In this section, we consider the equation

$$\operatorname{div}_H(A(\mathfrak{X}u)) = 0 \qquad \text{in } \Omega, \tag{3.1}$$

where $A : \mathbb{R}^m \to \mathbb{R}^m$ is a given C^1-function that satisfies the structure condition

$$\begin{aligned} \mathrm{F}(|z|)|\xi|^2 \le \langle DA(z)\,\xi, \xi\rangle \le L\,\mathrm{F}(|z|)|\xi|^2; \\ |A(z)| \le L\,|z|\mathrm{F}(|z|), \end{aligned} \tag{3.2}$$

for every $z, \xi \in \mathbb{R}^m$, given $\mathrm{F} : [0,\infty) \to [0,\infty)$ and $L > 1$. We assume F to be continuous in $(0,\infty)$ and there exists $\delta_0 > 0$ such that $\mathrm{F}(t) \ge \delta_0$ for all $t > 0$.

Here onwards, we denote $u \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega)$ as a weak solution of the equation (3.1). Furthermore, in this section we assume apriori that

$$\mathfrak{X}u \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m), \qquad Tu \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^n). \tag{3.3}$$

Thus, (3.3) implies $YX_ju, [X_i, Y]u \in L^2_{\mathrm{loc}}(\Omega)$ for any i, j, where $Y = X_j$ or $Y = T_k$. We only require the regularity (3.3) for the apriori estimates and the assumption shall be removed later by an approximation argument.

Lemma 3.1. *Let $Y = X_j$ or $Y = T_k$. For every $\varphi \in C^\infty_0(\Omega)$, we have*

$$\sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u) Y X_j u X_i \varphi \, dx = \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u)[X_i, Y]\varphi \, dx. \tag{3.4}$$

Proof. Given any $\varphi \in C^\infty_0(\Omega)$, we choose $\phi = Y\varphi$ on (2.14) and use integral by parts, to obtain

$$\begin{aligned} 0 &= \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u) X_i Y \varphi \, dx \\ &= \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u) Y X_i \varphi \, dx + \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u)[X_i, Y]\varphi \, dx \\ &= -\sum_{i=1}^m \int_\Omega Y(A_i(\mathfrak{X}u)) X_i \varphi \, dx + \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u)[X_i, Y]\varphi \, dx. \end{aligned} \tag{3.5}$$

The proof is finished by using chain rule $Y(A_i(\mathfrak{X}u)) = \sum_{j=1}^m D_j A_i(\mathfrak{X}u) Y X_j u$ on (3.5). □

Using $YX_j = X_jY - [X_j, Y]$ on (3.4), we have the following variant;

$$\begin{aligned} \sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u) X_j Y u X_i \varphi \, dx = \sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u)[X_j, Y] u X_i \varphi \, dx \\ + \sum_{i=1}^m \int_\Omega A_i(\mathfrak{X}u)[X_i, Y]\varphi \, dx \end{aligned} \tag{3.6}$$

for every $\varphi \in C^\infty_0(\Omega)$. Thus for $Y = X_l$, we have

$$\begin{aligned} \sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u) X_j X_l u X_i \varphi \, dx = \sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u)[X_j, X_l] u X_i \varphi \, dx \\ + \sum_i \int_\Omega A_i(\mathfrak{X}u)[X_i, X_l]\varphi \, dx, \end{aligned} \tag{3.7}$$

and for $Y = T_k$, using integral by parts on the last term, we have

$$\begin{aligned}\sum_{i,j}\int_\Omega D_jA_i(\mathfrak{X}u)X_jT_kuX_i\varphi\,dx &= \sum_{i,j}\int_\Omega D_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i\varphi\,dx\\ &\quad-\sum_{i,j}\int_\Omega \varphi D_jA_i(\mathfrak{X}u)[X_i,T_k]X_ju\,dx.\end{aligned}\tag{3.8}$$

3.1. **Caccioppoli type inequalities.** Here we provide Caccioppoli type estimates satisfied by the derivatives of a weak solution u of equation (3.1).

Lemma 3.2. *For any $\beta \geq 0$ and $\eta \in C_0^\infty(\Omega)$, there exists $c = c(N, L) > 0$ such that*

$$\begin{aligned}\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}Tu|^2\,dx \;\leq\; &\frac{c}{\tau}(\beta+1)^2\int_\Omega(\eta^2+|\mathfrak{X}\eta|^2)\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta\big(|\mathfrak{X}u|^2+|Tu|^2\big)\,dx\\ &+\tau\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx,\end{aligned}\tag{3.9}$$

for any arbitrary $0 < \tau < 1$.

Proof. In (3.8), we choose $\varphi = \varphi_k = \eta^2|Tu|^\beta T_ku$ on (3.8) and sum over $k \in \{1, \dots, n\}$, to obtain the following,

$$\begin{aligned}&\sum_{i,j,k}\int_\Omega \eta^2|Tu|^\beta D_jA_i(\mathfrak{X}u)X_jT_kuX_iT_ku\,dx\\ &\quad+\beta\sum_{i,j,k}\int_\Omega\eta^2|Tu|^{\beta-1}T_kuD_jA_i(\mathfrak{X}u)X_jT_kuX_i(|Tu|)\,dx\\ &\quad=-2\sum_{i,j,k}\int_\Omega\eta\,|Tu|^\beta T_kuD_jA_i(\mathfrak{X}u)X_jT_kuX_i\eta\,dx\\ &\qquad+\sum_{i,j,k}\int_\Omega D_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i\varphi_k\,dx\\ &\qquad+\sum_{i,j,k}\int_\Omega\varphi_kD_jA_i(\mathfrak{X}u)[X_i,T_k]X_ju\,dx\\ &\quad=R_1+R_2+R_3.\end{aligned}\tag{3.10}$$

Then, we use the structure condition (3.2) to estimate both sides of (3.10). The left hand side of the above is estimated as

$$\begin{aligned}\text{LHS of (3.10)} &\geq \int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}Tu|^2\,dx+\beta\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}(|Tu|)|^2\,dx\\ &\geq\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}Tu|^2\,dx.\end{aligned}\tag{3.11}$$

We estimate the right hand side of (3.10), one by one.

Using (3.2) and Young's inequality, we have

$$\begin{aligned}|R_1| &\leq c\int_\Omega|\eta|\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+1}|\mathfrak{X}Tu||\mathfrak{X}\eta|\,dx\\ &\leq\varepsilon\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}Tu|^2\,dx+\frac{c}{\varepsilon}\int_\Omega|\mathfrak{X}\eta|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx.\end{aligned}\tag{3.12}$$

Notice that

$$
\begin{aligned}
R_2 &= \sum_{i,j,k} \int_\Omega \eta^2 D_j A_i(\mathfrak{X}u)[X_j, T_k] u X_i(|Tu|^\beta T_k u)\, dx \\
&\quad + 2\sum_{i,j,k} \int_\Omega \eta D_j A_i(\mathfrak{X}u)[X_j, T_k] u |Tu|^\beta T_k u X_i \eta \, dx = R_{2,1} + R_{2,2}.
\end{aligned} \tag{3.13}
$$

Using (3.2) and (2.2) followed by Young's inequality, we have

$$
\begin{aligned}
|R_{2,1}| &\le c(\beta+1) \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u| + |Tu|\big)|Tu|^\beta |\mathfrak{X}Tu|\, dx \\
&\le \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2\, dx \\
&\quad + \frac{c}{\varepsilon}(\beta+1)^2 \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2\big)\, dx
\end{aligned} \tag{3.14}
$$

Similarly, we have

$$
|R_{2,2}| \le c \int_\Omega |\eta| \mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+1}\big(|\mathfrak{X}u| + |Tu|\big)|\mathfrak{X}\eta|\, dx \tag{3.15}
$$

which is clearly majorised by the right hand side of (3.9) from Young's inequality.

Then, R_3 is estimated using (3.2),(2.4) and Young's inequality on each term as

$$
\begin{aligned}
|R_3| &\le c \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+1}\big(|\mathfrak{X}\mathfrak{X}u| + |\mathfrak{X}Tu| + |\mathfrak{X}u| + |Tu|\big)\, dx \\
&\le \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2\, dx + c\Big(\frac{1}{\varepsilon} + \frac{1}{\tau}\Big) \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2\big)\, dx \\
&\quad + \tau \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}\mathfrak{X}u|^2\, dx,
\end{aligned} \tag{3.16}
$$

for any $0 < \tau < 1$. Combining (3.12),(3.14),(3.15),(3.16) with $\varepsilon = 1/6$, it is easy to obtain (3.9) and the proof is finished. □

Remark 3.3. For the special case when $X_1, \dots, X_m$ are left invariant with respect to a step 2 Carnot Group, we have $[X_i, T_k] = 0$ and hence, it corresponds to the case of $\tau = 0$ in (3.9), i.e.

$$
\int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2\, dx \le c(\beta+1)^2 \int_\Omega |\mathfrak{X}\eta|^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\, dx.
$$

Nevertheless, it does not make any difference in the subsequent results.

The following is the Caccioppoli type estimate of the vector fields.

Lemma 3.4. *For any $\beta \ge 0$ and $\eta \in C_0^\infty(\Omega)$, there exists $c = c(N, L) > 0$ such that*

$$
\begin{aligned}
\int_\Omega \eta^2\, \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2\, dx &\le c \int_\Omega \big(|\mathfrak{X}\eta|^2 + |\eta T\eta|\big)\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\, dx \\
&\quad + c(\beta+1)^4 \int_\Omega \eta^2\, \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2\big)\, dx.
\end{aligned} \tag{3.17}
$$

Proof. Let $\varphi_l = \eta^2 |\mathfrak{X}u|^\beta X_l u$ with $\eta \in C_0^\infty(\Omega)$ and $\beta \ge 0$. Notice that

$$
X_i \varphi_l = \eta^2 |\mathfrak{X}u|^\beta X_i X_l u + \beta \eta^2 |\mathfrak{X}u|^{\beta-1} X_l u X_i(|\mathfrak{X}u|) + 2\eta |\mathfrak{X}u|^\beta X_l u X_i \eta.
$$

We take $Y = X_l$ and $\varphi = \varphi_l$ in (3.6) and take summation over $l \in \{1, \dots, m\}$ to obtain

$$\begin{aligned}
&\sum_{i,j,l} \int_\Omega \eta^2 |\mathfrak{X}u|^\beta D_j A_i(\mathfrak{X}u) X_j X_l u X_i X_l u \, dx \\
&\quad + \beta \sum_{i,j,l} \int_\Omega \eta^2 |\mathfrak{X}u|^{\beta-1} X_l u D_j A_i(\mathfrak{X}u) X_j X_l u X_i(|\mathfrak{X}u|) \, dx \\
&\quad = \sum_{i,j,l} \int_\Omega \eta^2 |\mathfrak{X}u|^\beta D_j A_i(\mathfrak{X}u) [X_j, X_l] u X_i X_l u \, dx \\
&\quad\quad + \beta \sum_{i,j,l} \int_\Omega \eta^2 |\mathfrak{X}u|^{\beta-1} X_l u D_j A_i(\mathfrak{X}u) [X_j, X_l] u X_i(|\mathfrak{X}u|) \, dx \\
&\quad\quad - 2 \sum_{i,j,l} \int_\Omega \eta |\mathfrak{X}u|^\beta X_l u D_j A_i(\mathfrak{X}u) X_l X_j u X_i \eta \, dx \\
&\quad\quad + \sum_{i,l} \int_\Omega A_i(\mathfrak{X}u) [X_i, X_l] \varphi_l \, dx = J_1 + J_2 + J_3 + J_4,
\end{aligned} \tag{3.18}$$

where J_i are the consecutive items of the right hand side of the above.

First, we estimate the left hand side of (3.18) using (3.2) to have

$$\begin{aligned}
\text{LHS of (3.18)} &\geq \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \beta \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}(|\mathfrak{X}u|)|^2 \, dx \\
&\geq \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx.
\end{aligned} \tag{3.19}$$

For the right hand side of (3.18), we claim that the following holds,

$$\begin{aligned}
|J_i| &\leq \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \frac{c}{\varepsilon} \int_\Omega \big(|\mathfrak{X}\eta|^2 + |\eta T \eta| \big) \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+2} \, dx \\
&\quad + \frac{c}{\varepsilon} (\beta+1)^4 \int_\Omega \eta^2 \, \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2 \big) \, dx,
\end{aligned} \tag{3.20}$$

for any $0 < \varepsilon < 1$ and $i \in \{1, 2, 3, 4\}$, where $c = c(N, L) > 0$. Then (3.20) with a choice of small enough $\varepsilon = \varepsilon(N, L) > 0$ together with (3.19) would imply (3.17), thereby finishing the proof. We shall prove (3.20) by estimating each J_i of (3.18), one by one.

Using (3.2), (2.2) and Young's inequality, note that

$$\begin{aligned}
|J_1| + |J_2| &\leq c(\beta+1) \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u| \big(|\mathfrak{X}u| + |Tu| \big) \, dx \\
&\leq \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx \\
&\quad + \frac{c}{\varepsilon} (\beta+1)^2 \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2 \big) \, dx,
\end{aligned} \tag{3.21}$$

and the claim (3.20) follows for $i = 1, 2$. Similarly, for J_3 we have

$$\begin{aligned}
|J_3| &\leq c \int_\Omega |\eta| \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+1} |\mathfrak{X}\mathfrak{X}u| |\mathfrak{X}\eta| \, dx \\
&\leq \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \frac{c}{\varepsilon} \int_\Omega |\mathfrak{X}\eta|^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+2} \, dx.
\end{aligned} \tag{3.22}$$

To estimate J_4, first note that

$$[X_i, X_l] \varphi_l = \eta^2 |\mathfrak{X}u|^\beta [X_i, X_l] X_l u + \beta \eta^2 |\mathfrak{X}u|^{\beta-1} X_l u [X_i, X_l](|\mathfrak{X}u|) + 2\eta |\mathfrak{X}u|^\beta X_l u [X_i, X_l] \eta,$$

and we rewrite J_4 as

$$
\begin{aligned}
J_4 &= \sum_{i,l} \int_\Omega \eta^2 |\mathfrak{X}u|^\beta A_i(\mathfrak{X}u)[X_i, X_l] X_l u \, dx \\
&\quad + \beta \sum_{i,l} \int_\Omega \eta^2 |\mathfrak{X}u|^{\beta-1} X_l u A_i(\mathfrak{X}u)[X_i, X_l](|\mathfrak{X}u|) \, dx \\
&\quad + 2 \sum_{i,l} \int_\Omega \eta |\mathfrak{X}u|^\beta X_l u A_i(\mathfrak{X}u)[X_i, X_l]\eta \, dx = J_{4,1} + J_{4,2} + J_{4,3}.
\end{aligned} \tag{3.23}
$$

The estimate of the last term is straightforward. Using (3.2) and (2.2), we have

$$
|J_{4,3}| \le c \int_\Omega (|\eta T\eta| + |\eta \mathfrak{X}\eta|) \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+2} \, dx, \tag{3.24}
$$

which is majorised by the right hand side of (3.20). To estimate the rest of (3.23), we shall use $[X_i, X_l]X_l u = X_l[X_i, X_l]u + [[X_i, X_l], X_l]u$ and

$$
[X_i, X_l](|\mathfrak{X}u|) = |\mathfrak{X}u|^{-1} \sum_{k=1}^m X_k u [X_i, X_l] X_k u = |\mathfrak{X}u|^{-1} \sum_{k=1}^m X_k u \big(X_k [X_i, X_l] u + [[X_i, X_l], X_k] u \big),
$$

to split the first two terms as $J_{4,1} = J_{4,1,1} + J_{4,1,2}$ and $J_{4,2} = J_{4,2,1} + J_{4,2,2}$ so that the second terms i.e. $J_{4,1,2}$ and $J_{4,2,2}$ contain the terms with the commutator of commutators.

We rewrite the first terms by integral by parts as

$$
\begin{aligned}
J_{4,1,1} + J_{4,2,1} &= \sum_{i,l} \int_\Omega \eta^2 |\mathfrak{X}u|^\beta A_i(\mathfrak{X}u) X_l [X_i, X_l] u \, dx \\
&\quad + \beta \sum_{i,l,k} \int_\Omega \eta^2 |\mathfrak{X}u|^{\beta-2} X_l u A_i(\mathfrak{X}u) X_k u X_k [X_i, X_l] u \, dx \\
&= - \sum_{i,l} \int_\Omega X_l \Big(\eta^2 |\mathfrak{X}u|^\beta A_i(\mathfrak{X}u) \Big) [X_i, X_l] u \, dx \\
&\quad - \beta \sum_{i,l,k} \int_\Omega X_k \Big(\eta^2 |\mathfrak{X}u|^{\beta-2} X_l u X_k u A_i(\mathfrak{X}u) \Big) [X_i, X_l] u \, dx.
\end{aligned} \tag{3.25}
$$

Then, further computation of the terms in (3.25) followed by the use of (3.2) and (2.2) yields

$$
\begin{aligned}
|J_{4,1,1}| + |J_{4,2,1}| &\le c(\beta+1)^2 \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u| \big(|\mathfrak{X}u| + |Tu| \big) \, dx \\
&\quad + c(\beta+1) \int_\Omega |\eta| \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+1} |\mathfrak{X}\eta| \big(|\mathfrak{X}u| + |Tu| \big) \, dx \\
&\le \varepsilon \int_\Omega \eta^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx \\
&\quad + \frac{c}{\varepsilon} (\beta+1)^4 \int_\Omega \eta^2 \, \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^\beta \big(|\mathfrak{X}u|^2 + |Tu|^2 \big) \, dx \\
&\quad + c \int_\Omega |\mathfrak{X}\eta|^2 \mathrm{F}(|\mathfrak{X}u|) |\mathfrak{X}u|^{\beta+2} \, dx,
\end{aligned} \tag{3.26}
$$

where the latter inequality of the above follows by using Young's inequality on both items. Now we are only left with

$$\begin{aligned} J_{4,1,2}+J_{4,2,2} &= \sum_{i,l}\int_\Omega \eta^2|\mathfrak{X}u|^\beta A_i(\mathfrak{X}u)[[X_i,X_l],X_l]u\,dx \\ &\quad +\beta\sum_{i,l,k}\int_\Omega \eta^2|\mathfrak{X}u|^{\beta-2}X_luA_i(\mathfrak{X}u)X_ku[[X_i,X_l],X_k]u\,dx \end{aligned} \tag{3.27}$$

To estimate (3.27), we use (2.5) to have

$$|J_{4,1,2}|+|J_{4,2,2}| \le c(\beta+1)\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+1}\big(|\mathfrak{X}u|+|Tu|\big)\,dx, \tag{3.28}$$

which is estimated easily just as in (3.26). Thus, (3.26), (3.27) and (3.24) together imply that the claim (3.20) holds for $i=4$ as well. This completes the proof of the the claim (3.20) and hence the proof of the lemma. □

The following is a reverse type inequality similar to earlier papers, see [36, 47, 38], etc.

Lemma 3.5. *For any $\beta\ge 2$ and all non-negative $\eta\in C_0^\infty(\Omega)$, we have*

$$\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\beta+1)^4K_\eta\int_\Omega \eta^\beta\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^{\beta-2}|\mathfrak{X}\mathfrak{X}u|^2\,dx \tag{3.29}$$

for some constant $c=c(N,L)>0$, where $K_\eta=\|\eta\|_{L^\infty}^2+\|\mathfrak{X}\eta\|_{L^\infty}^2$.

Proof. Recalling (3.4) with $Y=X_l$ and $\varphi=\varphi_l$, we have

$$\sum_{i,j}\int_\Omega D_jA_i(\mathfrak{X}u)X_lX_juX_i\varphi_l\,dx = \sum_{i=1}^m\int_\Omega A_i(\mathfrak{X}u)[X_i,X_l]\varphi_l\,dx, \tag{3.30}$$

where we use $\varphi_l=\eta^{\beta+2}|Tu|^\beta X_lu$, where $\eta\in C_0^\infty(\Omega)$ is non-negative and $\beta\ge 2$. Note that

$$X_i\varphi_l=\eta^{\beta+2}|Tu|^\beta X_iX_lu+\beta\eta^{\beta+2}|Tu|^{\beta-1}X_luX_i(|Tu|)+(\beta+2)\eta^{\beta+1}X_i\eta|Tu|^\beta X_lu.$$

Using this on the above, we obtain

$$\begin{aligned} &\sum_{i,j}\int_\Omega \eta^{\beta+2}|Tu|^\beta D_jA_i(\mathfrak{X}u)X_lX_juX_iX_lu\,dx \\ &\qquad = -\beta\sum_{i,j}\int_\Omega \eta^{\beta+2}|Tu|^{\beta-1}X_luD_jA_i(\mathfrak{X}u)X_lX_juX_i(|Tu|)\,dx \\ &\qquad\quad -(\beta+2)\sum_{i,j}\int_\Omega \eta^{\beta+1}|Tu|^\beta X_luD_jA_i(\mathfrak{X}u)X_lX_juX_i\eta\,dx \\ &\qquad\quad -\sum_{i,j}\int_\Omega \varphi_lD_jA_i(\mathfrak{X}u)[X_i,X_l]X_ju\,dx, \end{aligned} \tag{3.31}$$

where we have used integral by parts on the term in the right hand side of (3.30). Now we use $X_iX_l=X_lX_i+[X_i,X_l]$ on the left hand side term and then sum over $l\in\{1,\dots,m\}$ both sides

of (3.31) to obtain

$$
\begin{aligned}
&\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta D_jA_i(\mathfrak{X}u)X_lX_juX_lX_iu\,dx\\
&\qquad=-\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta D_jA_i(\mathfrak{X}u)X_lX_ju[X_i,X_l]u\,dx\\
&\qquad\quad-\beta\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^{\beta-1}X_luD_jA_i(\mathfrak{X}u)X_lX_juX_i(|Tu|)\,dx\\
&\qquad\quad-(\beta+2)\sum_{i,j,l}\int_\Omega \eta^{\beta+1}|Tu|^\beta X_luD_jA_i(\mathfrak{X}u)X_lX_juX_i\eta\,dx\\
&\qquad\quad-\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta X_luD_jA_i(\mathfrak{X}u)[X_i,X_l]X_ju\,dx\\
&\qquad=I_1+I_2+I_3+I_4.
\end{aligned}
\tag{3.32}
$$

We will estimate both sides of (3.32) as follows.

For the left hand side, the structure condition (3.2) implies that

$$
\text{LHS of (3.32)}\ge\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx.
$$

For the right hand side of (3.32), we will show that for each item, the following estimate holds,

$$
\begin{aligned}
|I_k|\ &\le c\varepsilon\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\\
&\quad+\frac{c}{\varepsilon}(\beta+1)^4K_\eta\int_\Omega \eta^{\beta}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^{\beta-2}|\mathfrak{X}\mathfrak{X}u|^2\,dx,
\end{aligned}
\tag{3.33}
$$

for $k=1,2,3,4$. Then, a choice of a small enough $\varepsilon=\varepsilon(N,L)>0$, completes the proof.

To prove the claim (3.33), first we note that using (2.6) on (3.9) of Lemma 3.2 along with minor modifications, we can obtain

$$
\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}Tu|^2\,dx\le c(\beta+1)^2K_\eta\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta\big(|\mathfrak{X}u|^2+|\mathfrak{X}\mathfrak{X}u|^2\big)\,dx
\tag{3.34}
$$

where $K_\eta=\|\eta\|_{L^\infty}^2+\|\mathfrak{X}\eta\|_{L^\infty}^2$. This shall be used multiple times below.

First, to prove that (3.33) holds for I_1, we note that

$$
\begin{aligned}
I_1&=-\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta D_jA_i(\mathfrak{X}u)X_lX_ju[X_i,X_l]u\,dx\\
&=-\sum_{i,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta X_l(A_i(\mathfrak{X}u))[X_i,X_l]u\,dx,
\end{aligned}
$$

from chain rule. Then integration by parts yields

$$
\begin{aligned}
I_1&=\sum_{i,l}\int_\Omega A_i(\mathfrak{X}u)X_l\big(\eta^{\beta+2}|Tu|^\beta[X_i,X_l]u\big)\,dx\\
&=\sum_{i,l}\int_\Omega \eta^{\beta+2}A_i(\mathfrak{X}u)X_l\big(|Tu|^\beta[X_i,X_l]u\big)\,dx\\
&\quad-(\beta+2)\sum_{i,l}\int_\Omega \eta^{\beta+1}A_i(\mathfrak{X}u)X_l\eta\,|Tu|^\beta[X_i,X_l]u\,dx=I_{1,1}+I_{1,2}.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
I_{1,1} &= \sum_{i,l} \int_\Omega \eta^{\beta+2} A_i(\mathfrak{X}u)|Tu|^\beta X_l[X_i, X_l]u \, dx \\
&\quad + \beta \sum_{i,l} \int_\Omega \eta^{\beta+2} A_i(\mathfrak{X}u)|Tu|^{\beta-1}[X_i, X_l]u \, X_l(|Tu|) \, dx \\
&= I_{1,1,1} + I_{1,1,2}.
\end{aligned} \tag{3.35}
$$

For $I_{1,1,1}$, using (2.2) and (3.2), we have

$$
\begin{aligned}
|I_{1,1,1}| &\le c \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta |\mathfrak{X}Tu| \, dx + c \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta |\mathfrak{X}\mathfrak{X}u| \, dx \\
&\le \frac{\varepsilon}{(\beta+1)^2 K_\eta} \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2 \, dx + \frac{c}{\varepsilon}(\beta+1)^2 K_\eta \int_\Omega \eta^\beta \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2 |Tu|^\beta \, dx \\
&\quad + \varepsilon \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \frac{c}{\varepsilon} \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}u|^2 \, dx,
\end{aligned}
$$

where the latter inequality is obtained by using Young's inequality on both terms. Then, using (3.34) to estimate the first term of the right hand side of the above, it is not hard to see that

$$
|I_{1,1,1}| \le c\varepsilon \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \frac{c}{\varepsilon}(\beta+1)^2 K_\eta \int_\Omega \eta^\beta \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}u|^2 \, dx \tag{3.36}
$$

holds for $0 < \varepsilon < 1$, $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2$ and some large enough $c = c(N, L) > 0$. Thus, we have that $I_{1,1,1}$ satisfies the estimate of the claim (3.33).

For $I_{1,1,2}$, we just use $|[X_i, X_l]u| \le 2|\mathfrak{X}\mathfrak{X}u|$ along with (3.2) and Young's inequality to get

$$
\begin{aligned}
|I_{1,1,2}| &\le c\beta \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^{\beta-1} |\mathfrak{X}\mathfrak{X}u||\mathfrak{X}Tu| \, dx \\
&\le \frac{\varepsilon}{(\beta+1)^2 K_\eta} \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2 \, dx \\
&\quad + \frac{c}{\varepsilon}\beta^2(\beta+1)^2 K_\eta \int_\Omega \eta^\beta \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2 |Tu|^{\beta-2} |\mathfrak{X}\mathfrak{X}u|^2 \, dx.
\end{aligned} \tag{3.37}
$$

Then, using (3.34) to estimate the first term similarly as before, it is easy to see that $I_{1,1,2}$ also satisfies the estimate of the claim (3.33).

For $I_{1,2}$, we similarly use (3.2), $|[X_i, X_l]u| \le 2|\mathfrak{X}\mathfrak{X}u|$ and Young's inequality to obtain

$$
\begin{aligned}
|I_{1,2}| &\le c(\beta+1) \int_\Omega \eta^{\beta+1} |\mathfrak{X}\eta| \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta |\mathfrak{X}\mathfrak{X}u| \, dx \\
&\le c\varepsilon \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}\mathfrak{X}u|^2 \, dx + \frac{c}{\varepsilon}(\beta+1)^2 \int_\Omega \eta^\beta |\mathfrak{X}\eta|^2 \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}u|^2 \, dx.
\end{aligned} \tag{3.38}
$$

Combining (3.36),(3.37) and (3.38), we conclude that the claim (3.33) holds for I_1.

Similarly, we estimate I_2 by structure condition (3.2) and Young's inequality to get

$$
|I_2| \le c\beta \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^{\beta-1} |\mathfrak{X}\mathfrak{X}u||\mathfrak{X}Tu| \, dx,
$$

which is estimated exactly as $I_{1,1,2}$ and thus the claim (3.33) holds for I_2.

For I_3, we have by structure condition (3.2), that

$$
|I_3| \le c(\beta+1) \int_\Omega \eta^{\beta+1} |\mathfrak{X}\eta| \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta |\mathfrak{X}\mathfrak{X}u| \, dx,
$$

which is estimated exactly as $I_{1,2}$ and thus the claim (3.33) holds for I_3 as well.

Finally, we note that

$$\begin{aligned} I_4 &= -\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta X_l u D_j A_i(\mathfrak{X}u)X_j[X_i,X_l]u\,dx \\ &\quad +\sum_{i,j,l}\int_\Omega \eta^{\beta+2}|Tu|^\beta X_l u D_j A_i(\mathfrak{X}u)[[X_i,X_l],X_j]u\,dx \\ &= I_{4,1}+I_{4,2}. \end{aligned}$$

From (2.2) and (3.2), we find that

$$|I_{4,1}| \le c\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta|\mathfrak{X}Tu|\,dx + c\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta|\mathfrak{X}\mathfrak{X}u|\,dx \tag{3.39}$$

which is estimated exactly as $I_{1,1,1}$. For $I_{4,2}$, we use (3.2) followed by (2.5) and (2.6) used successively to obtain

$$|I_{4,2}| \le c\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||Tu|^\beta\big(|\mathfrak{X}u|+|\mathfrak{X}\mathfrak{X}u|\big)\,dx, \tag{3.40}$$

the first term of the above is estimated from (2.6) and the second term is exactly equal to the second term of (3.39). Thus, from (3.39) and (3.40) we conclude that the claim (3.33) holds for I_4 and the proof is finished. □

Remark 3.6. It is not difficult to see that from the above proof, the case $\beta = 0$ leads to

$$\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c\big(\|\eta\|^2_{L^\infty}+\|\mathfrak{X}\eta\|^2_{L^\infty}\big)\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2\,dx, \tag{3.41}$$

since all the terms containing $\beta|Tu|^{\beta-1}$ in the proof vanish in the first steps of the estimates.

The following corollary is proved by using (2.6) and Hölder's inequality on Lemma 3.5.

Corollary 3.7. *Given $\beta \ge 2$ and non-negative $\eta \in C_0^\infty(\Omega)$, we have $c = c(N,L) > 0$ such that*

$$\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c^{\frac{\beta}{2}}(\beta+1)^{2\beta}K_\eta^{\beta/2}\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \tag{3.42}$$

where $K_\eta = \|\eta\|^2_{L^\infty}+\|\mathfrak{X}\eta\|^2_{L^\infty}$ and $c = c(N,L) > 0$.

Proof. The right hand side of (3.29) is estimated by Hölder's inequality as

$$\begin{aligned} &\int_\Omega \eta^\beta \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^{\beta-2}|\mathfrak{X}\mathfrak{X}u|^2\,dx \\ &\qquad \le \Big(\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{\beta-2}{\beta}}\Big(\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{2}{\beta}}. \end{aligned} \tag{3.43}$$

Then, using Young's inequality on (3.43) and combining with (3.29), the proof is finished. □

Now, we can remove the error term in (3.17) and obtain the following apriori estimate. This shall be used to prove local Lipschitz regularity of weak solutions.

Lemma 3.8. *For any $\beta \ge 2$ and all non-negative $\eta \in C_0^\infty(\Omega)$, we have that*

$$\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\beta+1)^{12}K_\eta\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx, \tag{3.44}$$

where $K_\eta = \|\eta\|^2_{L^\infty}+\|\mathfrak{X}\eta\|^2_{L^\infty}+\|\eta T\eta\|_{L^\infty}$ and $c = c(N,L) > 0$.

Proof. Let us recall (3.17) and rewrite it as

$$\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\beta+1)^4\int_\Omega \big(\eta^2+|\mathfrak{X}\eta|^2+|\eta T\eta|\big)\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx + c(\beta+1)^4\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|Tu|^2\,dx. \tag{3.45}$$

It is clear that we need to estimate only the last term of the right hand side of (3.45). We estimate it by Hölder's inequality and Young's inequality as

$$\begin{aligned} &c(\beta+1)^4\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|Tu|^2\,dx\\ &\le c(\beta+1)^4\Big(\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx\Big)^{\frac{2}{\beta+2}}\Big(\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx\Big)^{\frac{\beta}{\beta+2}}\\ &\le c\,h^{\frac{\beta+2}{2}}\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx+\frac{c}{h^{\frac{\beta+2}{\beta}}}(\beta+1)^{\frac{4(\beta+2)}{\beta}}\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx \end{aligned} \tag{3.46}$$

for some $h>0$ to be chosen later. The first term of the right hand side of (3.46) is estimated using (2.6) and Corollary 3.7 as

$$\begin{aligned} c\,h^{\frac{\beta+2}{2}}\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx &\le c\,h^{\frac{\beta+2}{2}}\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\\ &\le (ch)^{\frac{\beta+2}{2}}(\beta+1)^{2\beta}K_\eta^{\beta/2}\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx. \end{aligned} \tag{3.47}$$

Now, for a large enough $c=c(N,L)>0$ and small enough $\varepsilon>0$, we take

$$h=\frac{\varepsilon}{c(\beta+1)^{4\beta/(\beta+2)}K_\eta^{\beta/(\beta+2)}}, \tag{3.48}$$

so that the above implies

$$c\,h^{\frac{\beta+2}{2}}\int_\Omega \eta^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx \le \varepsilon^{\frac{\beta+2}{2}}\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \tag{3.49}$$

Then, it is not difficult to observe that (3.49) used in (3.46) followed by (3.46) used in (3.45) with the choice of h in (3.48) for a small enough $\varepsilon=\varepsilon(N,L)>0$, leads to

$$\int_\Omega \eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\beta+1)^{\frac{4(\beta+2)}{\beta}+4}K_\eta\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx,$$

for some $c=c(N,L)>0$. This completes the proof. □

The following corollaries would be essential for proving the $C^{1,\alpha}$-regularity of weak solutions.

Corollary 3.9. *For any $\beta\ge 2$ and all non-negative $\eta\in C_0^\infty(\Omega)$, we have that*

$$\int_\Omega \eta^{\beta+2}\,\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx \le c(\beta)K_\eta^{\frac{\beta+2}{2}}\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2}\,dx, \tag{3.50}$$

where $K_\eta=\|\eta\|_{L^\infty}^2+\|\mathfrak{X}\eta\|_{L^\infty}^2+\|\eta T\eta\|_{L^\infty}$ and $c(\beta)=c(N,L,\beta)>0$.

Proof. It is easy to see that from (2.6), we have

$$\int_\Omega \eta^{\beta+2}\,\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2}\,dx \le c\int_\Omega \eta^{\beta+2}\,\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx$$

and then (3.50) follows easily from Corollary 3.7 and (3.44). This completes the proof. □

Corollary 3.10. *For any $\beta \geq 2$ and all non-negative $\eta \in C_0^\infty(\Omega)$, we have that*

$$\int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2 \, dx \leq c(\beta) K_\eta^{\frac{\beta+4}{2}} \int_{\mathrm{supp}(\eta)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2} \, dx, \tag{3.51}$$

where $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty}$ and $c(\beta) = c(N, L, \beta) > 0$.

Proof. Recalling the variant (3.34) of (3.9) of Lemma 3.2, we have

$$\int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}Tu|^2 \, dx \leq c(\beta+1)^2 K_\eta \int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta \big(|\mathfrak{X}u|^2 + |\mathfrak{X}\mathfrak{X}u|^2\big) \, dx.$$

The second term of the right hand side of the above is estimated from Corrollary 3.7 and (3.44) and the first term of the right hand side of the above is estimated by Hölder's inequality as

$$\int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta |\mathfrak{X}u|^2 \, dx \leq \Big(\int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|Tu|^{\beta+2} \, dx \Big)^{\frac{\beta}{\beta+2}} \Big(\int_\Omega \eta^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2} \, dx \Big)^{\frac{2}{\beta+2}}$$

and it is further estimated from (3.50). Combining all the estimates, the proof is finished. □

Finally, we need the following Corrollary, which also relies on (3.42), (3.44) and (3.50).

Corollary 3.11. *For any $\beta \geq 2$ and all non-negative $\eta \in C_0^\infty(\Omega)$, we have that*

$$\int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}Tu|^2 \, dx \leq c(\beta) K_\eta^{\frac{\beta+4}{2}} \int_{\mathrm{supp}(\eta)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2} \, dx, \tag{3.52}$$

where $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty}$ and $c(\beta) = c(N, L, \beta) > 0$.

Proof. By testing the equation (3.8) with $\varphi_k = \eta^{\beta+4}|\mathfrak{X}u|^\beta T_k u$ and summing over $k \in \{1, \dots, n\}$, we get

$$\begin{aligned}
&\sum_{i,j,k} \int_\Omega \eta^{\beta+4} |\mathfrak{X}u|^\beta D_j A_i(\mathfrak{X}u) X_j T_k u X_i T_k u \, dx \\
&\qquad = -(\beta+4) \sum_{i,j,k} \int_\Omega \eta^{\beta+3} |\mathfrak{X}u|^\beta T_k u D_j A_i(\mathfrak{X}u) X_j T_k u X_i \eta \, dx \\
&\qquad\quad - \beta \sum_{i,j,k} \int_\Omega \eta^{\beta+4} |\mathfrak{X}u|^{\beta-1} T_k u D_j A_i(\mathfrak{X}u) X_j T_k u X_i(|\mathfrak{X}u|) \, dx \\
&\qquad\quad + \sum_{i,j,k} \int_\Omega D_j A_i(\mathfrak{X}u)[X_j, T_k] u X_i \varphi_k \, dx \\
&\qquad\quad + \sum_{i,j,k} \int_\Omega \varphi_k D_j A_i(\mathfrak{X}u)[X_i, T_k] X_j u \, dx \\
&\qquad = L_1 + L_2 + L_3 + L_4.
\end{aligned} \tag{3.53}$$

From the structure condition (3.2),

$$\begin{aligned}
\text{LHS of (3.53)} &\geq \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}Tu|^2 \, dx + \beta \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}(|Tu|)|^2 \, dx \\
&\geq \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}Tu|^2 \, dx.
\end{aligned} \tag{3.54}$$

We estimate the right hand side of (3.53), one by one. We claim that for $i \in \{1, 2, 3, 4\}$,

$$|L_i| \leq \varepsilon \int_\Omega \eta^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |\mathfrak{X}Tu|^2 \, dx + \frac{c(\beta)}{\varepsilon} K_\eta^{\frac{\beta+4}{2}} \int_{\mathrm{supp}(\eta)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta+2} \, dx, \tag{3.55}$$

which, with a small enough $\varepsilon = \varepsilon(N, L) > 0$, is enough to conclude the proof.

First, using (3.2) and Young's inequality, we have

$$
\begin{aligned}
|L_1| &\le c(\beta+1)\int_\Omega |\eta|^{\beta+3}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta |Tu||\mathfrak{X}Tu||\mathfrak{X}\eta|\,dx \\
&\le \varepsilon\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}Tu|^2\,dx + \frac{c}{\varepsilon}(\beta+1)^2\int_\Omega \eta^{\beta+2}|\mathfrak{X}\eta|^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|Tu|^2\,dx,
\end{aligned}
\tag{3.56}
$$

and the claim (3.55) follows from Hölder's inequality and (3.50). Similarly, we have

$$
\begin{aligned}
|L_2| &\le c\beta\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta-1}|Tu||\mathfrak{X}Tu||\mathfrak{X}\mathfrak{X}u|\,dx \\
&\le \varepsilon\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}Tu|^2\,dx + \frac{c}{\varepsilon}\beta^2\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta-2}|Tu|^2|\mathfrak{X}\mathfrak{X}u|^2\,dx,
\end{aligned}
\tag{3.57}
$$

where the last term in the right hand side of (3.57) is estimated by Hölder's inequality as

$$
\begin{aligned}
&\int_\Omega \eta^4\mathrm{F}(|\mathfrak{X}u|)|Tu|^2|\mathfrak{X}u|^{\beta-2}|\mathfrak{X}\mathfrak{X}u|^2\,dx \\
&\qquad\le \Big(\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{\beta-2}{\beta}}\Big(\int_\Omega \eta^{\beta+2}\,\mathrm{F}(|\mathfrak{X}u|)|Tu|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{2}{\beta}}.
\end{aligned}
$$

and the claim (3.55) follows from (3.42) and (3.44).

To continue the estimates, note that

$$
\begin{aligned}
L_3 &= \sum_{i,j,k}\int_\Omega \eta^{\beta+4}|\mathfrak{X}u|^\beta D_jA_i(\mathfrak{X}u)[X_j,T_k]uX_iT_ku\,dx \\
&\quad+(\beta+4)\sum_{i,j,k}\int_\Omega \eta^{\beta+3}\,|\mathfrak{X}u|^\beta T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i\eta\,dx \\
&\quad-\beta\sum_{i,j,k}\int_\Omega \eta^{\beta+4}|\mathfrak{X}u|^{\beta-1}T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i(|\mathfrak{X}u|)\,dx.
\end{aligned}
\tag{3.58}
$$

Using (3.2) and (2.2), we obtain

$$
\begin{aligned}
|L_3| &\le c\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}Tu|\,dx \\
&\quad+c(\beta+1)\int_\Omega |\eta|^{\beta+3}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|Tu|\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}\eta|\,dx \\
&\quad+c\beta\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta-1}|Tu|\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}\mathfrak{X}u|\,dx,
\end{aligned}
$$

and then, by Young's inequality on all the above terms, we have

$$
\begin{aligned}
|L_3| &\le \varepsilon\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|\mathfrak{X}Tu|^2\,dx \\
&\quad+\frac{c}{\varepsilon}(\beta+1)^2\int_\Omega \eta^{\beta+2}(\eta^2+|\mathfrak{X}\eta|^2)\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta\big(|\mathfrak{X}u|^2+|Tu|^2\big)\,dx \\
&\quad+c\beta^2\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\beta-2}|Tu|^2|\mathfrak{X}\mathfrak{X}u|^2\,dx,
\end{aligned}
\tag{3.59}
$$

which is estimated similarly as L_1 and L_2 in (3.56) and (3.57) above.

Finally, we note that

$$
L_4 = \sum_{i,j,k}\int_\Omega \eta^{\beta+4}|\mathfrak{X}u|^\beta T_kuD_jA_i(\mathfrak{X}u)[X_i,T_k]X_ju\,dx
$$

which, by (3.2) and (2.4), leads to

$$|L_4| \le c\int_\Omega \eta^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\beta|Tu|\big(|\mathfrak{X}\mathfrak{X}u| + |\mathfrak{X}Tu| + |\mathfrak{X}u| + |Tu|\big)\,dx, \tag{3.60}$$

which is a linear combination of the right hand sides of the estimates L_1, L_2 and L_3 of the above, and hence is estimated accordingly. Thus (3.55) holds for all L_i's and the proof is finished. □

Remark 3.12. Similarly as in the case of (3.41), it is not hard to see that from the above proof, the case $\beta = 0$ leads to

$$\int_\Omega \eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}Tu|^2\,dx \le cK_\eta\int_{\mathrm{supp}(\eta)}\mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u|^2 + |Tu|^2\big)\,dx, \tag{3.61}$$

with $K_\eta = \|\eta\|^2_{L^\infty} + \|\mathfrak{X}\eta\|^2_{L^\infty} + \|\eta T\eta\|_{L^\infty}$, since all the terms containing $\beta|\mathfrak{X}u|^{\beta-1}$ vanish.

4. Local Lipschitz Continuity

In this section, we shall prove Theorem 1.1 using the Caccioppoli type estimate (3.44) obtained in Section 3 and an approximation argument. We shall fix $x_0 \in \Omega$ and denote all CC-balls as $B_r = B_r(x_0)$ with dimension $Q = Q(x_0)$. Notice that whenever $\delta > 0$, the function

$$\mathrm{F}(t) = (\delta + t^2)^{\frac{p-2}{2}} \tag{4.1}$$

satisfies all conditions in the begining of Section 3 for any $1 < p < \infty$ and (3.2) is identical to the structure condition (1.3). The substitution (4.1) on the estimate (3.44) followed by the use of Sobolev inequality and Moser's iteration, gives us the estimate of Theorem 1.1 subject to the apriori assumption (3.3). The proof is standard, we provide a brief outline below.

Theorem 4.1. *If $u \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega)$ is a weak solution of equation (1.2) and (3.3) holds, then*

$$\sup_{B_{\sigma r}}|\mathfrak{X}u| \le \frac{c}{(1-\sigma)^{Q/p}}\left(⨍_{B_r}\big(\delta + |\mathfrak{X}u|^2\big)^{\frac{p}{2}}\,dx\right)^{\frac{1}{p}} \tag{4.2}$$

holds for any $0 < \sigma < 1$, where $1 < p < \infty$, $\delta > 0$ and $c = c(N, p, L) > 0$.

Proof. We rewrite the Caccioppoli type estimate (3.44) with the substitution (4.1) to have

$$\int_\Omega \eta^2(\delta + |\mathfrak{X}u|^2)^{\frac{p-2}{2}}|\mathfrak{X}u|^\beta|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\beta+1)^{12}K_\eta\int_{\mathrm{supp}(\eta)}(\delta + |\mathfrak{X}u|^2)^{\frac{p-2}{2}}|\mathfrak{X}u|^{\beta+2}\,dx \tag{4.3}$$

for any $\beta \ge 2$, where $K_\eta = \|\eta\|^2_{L^\infty} + \|\mathfrak{X}\eta\|^2_{L^\infty} + \|\eta T\eta\|_{L^\infty}$. Letting $w = (\delta + |\mathfrak{X}u|^2)^{p/2}$ and using a standard choice of test function $\eta \in C^\infty_0(B_r)$ such that $0 \le \eta \le 1$ and $\eta \equiv 1$ in $B_{r'}$ for $0 < r' < r$, $|\mathfrak{X}\eta| \le 4/(r-r')$ and $|\mathfrak{X}\mathfrak{X}\eta| \le 16N/(r-r')^2$, (4.3) implies

$$\int_{B_{r'}} w^\gamma|\mathfrak{X}w|^2\,dx \le \frac{c(\gamma+1)^{12}}{(r-r')^2}\int_{B_r} w^{\gamma+2}\,dx \tag{4.4}$$

for large enough γ. Using Sobolev's inequality (2.13) for $q = 2$ on (4.4), we get that

$$\left(\int_{B_{r'}} w^{(\gamma+2)\kappa}\,dx\right)^{\frac{1}{\kappa}} \le \frac{c(\gamma+1)^{14}}{(r-r')^2}\int_{B_r} w^{\gamma+2}\,dx,$$

where $\kappa = Q/(Q-2)$. Standard Moser's iteration is carried out by taking an appropriate sequence γ_i and $r_i = \sigma r + (1-\sigma)r/2^i$ on the above, then the standard interpolation argument [11, p. 299–300] leads to (4.2) and thereby completes the proof. □

4.1. **Riemannian Approximation.** The sub-Riemannian metric generated by the vector fields $\{X_1,\dots,X_m\}$ are approximated in the Gromov-Hausdorff sense via Riemannian approximation generated by $\{X_1^\varepsilon,\dots,X_{m+n}^\varepsilon\}$, which is a relabelling of $\{X_1,\dots,X_m,\varepsilon T_1,\dots,\varepsilon T_n\}$, as $\varepsilon\to 0^+$ and generates the norm $|\cdot|_\varepsilon$ on $\mathbb{R}^{m+n}$. The second order vector fields remain $\{T_1^\varepsilon,\dots,T_n^\varepsilon\}=\{T_1,\dots,T_n\}$. In other words, the approximate vector fields in consideration are

$$\{X_1^\varepsilon,\dots,X_{m+n}^\varepsilon,T_1^\varepsilon,\dots,T_n^\varepsilon\}=\{X_1,\dots,X_m,\varepsilon T_1,\dots,\varepsilon T_n,T_1,\dots,T_n\} \tag{4.5}$$

which is exactly the step 2 case of the approximations provided in [5, 3]. The idea is to be able to carry out everything above with X_j^ε and T_k^ε in place of X_j and T_k. The reason for the relabelling (4.5) is that the equation with respect to X_j^ε's become automatically elliptic since $\{X_1,\dots,X_m,T_1,\dots,T_n\}$ spans the tangent space and hence, we can assume as much apriori regularity as needed. Then the conclusion shall follow with the limit $\varepsilon\to 0^+$ provided that all constants do not blow up. We briefly illustrate in the following that this indeed is the case when the set up of the begining of Section 2 is re-written in terms of the approximation (4.5).

We denote the approximation gradients as

$$\mathfrak{X}^\varepsilon u=(X_1^\varepsilon u,\dots,X_{m+n}^\varepsilon u)=(X_1u,\dots,X_mu,\varepsilon T_1u,\dots,\varepsilon T_nu),\quad\text{and}\quad T^\varepsilon u=Tu$$

and $\operatorname{div}_{_H}^\varepsilon$ is similarly defined. Hence, $\mathfrak{X}^\varepsilon u\to(\mathfrak{X}u,0)$ and $|\mathfrak{X}^\varepsilon u|_\varepsilon^2=|\mathfrak{X}u|^2+\varepsilon^2|Tu|^2\to|\mathfrak{X}u|^2$ as $\varepsilon\to 0^+$. Given $A:\mathbb{R}^m\to\mathbb{R}^m$ that satisfies (1.3), one can define $A_\varepsilon:\mathbb{R}^{m+n}\to\mathbb{R}^{m+n}$ as

$$A_\varepsilon(\bar z)=\big(1-\eta_\varepsilon(|\bar z|_\varepsilon)\big)\tilde A(\bar z)+\eta_\varepsilon(|\bar z|_\varepsilon)(\delta+|\bar z|_\varepsilon^2)^{\frac{p-2}{2}}\bar z \tag{4.6}$$

similarly as (2.20) for any $0<\varepsilon<1$, where $\tilde A(\bar z)=(A(z),0)$ for any $\bar z=(z,z')\in\mathbb{R}^{m+n}$ and $\eta_\varepsilon\in C^{0,1}([0,\infty))$ can be chosen (see [38]) such that $A_\varepsilon\to A$ uniformly on compact subsets as $\varepsilon\to 0^+$ and the structure condition

$$\begin{aligned}\frac{1}{L'}(\delta+|\bar z|_\varepsilon^2)^{\frac{p-2}{2}}|\xi|^2\le\big\langle DA_\varepsilon(\bar z)\,\xi,\xi\big\rangle\le L'(\delta+|\bar z|_\varepsilon^2)^{\frac{p-2}{2}}|\xi|^2;\\ |A_\varepsilon(\bar z)|\le L'\,|\bar z|_\varepsilon(\delta+|\bar z|_\varepsilon^2)^{\frac{p-2}{2}},\end{aligned} \tag{4.7}$$

holds for any $\bar z,\xi\in\mathbb{R}^{m+n}$, for some $L'=L'(p,L)>1$. Since (4.7) ensures that the equation $\operatorname{div}_{_H}^\varepsilon(A_\varepsilon(\mathfrak{X}^\varepsilon u))=0$ is uniformly elliptic on Ω, hence it has all necessary regularities including (3.3). As an example we see that

$$\operatorname{div}_{_H}^\varepsilon\big((\delta+|\mathfrak{X}^\varepsilon u|_\varepsilon^2)^{\frac{p-2}{2}}\mathfrak{X}^\varepsilon u\big)=0\quad\text{in }\Omega, \tag{4.8}$$

is a natural regularization of the sub-elliptic p-Laplacian (1.4).

From the step 2 hypothesis (2.1), note that $\{X_1,\dots,X_m,T_1,\dots,T_n\}$ spans the tangent space automatically implies $\{X_1,\dots,X_m,(1+\varepsilon)T_1,\dots,(1+\varepsilon)T_n\}$ spans the tangent space for any $0\le\varepsilon\le 1$ and therefore, any vector field can be written as linear combinations of $\{X_1^\varepsilon,\dots,X_m^\varepsilon,T_1^\varepsilon,\dots,T_n^\varepsilon\}$ with bounded coefficients; in particular the commutators $[X_i^\varepsilon,X_j^\varepsilon]$ and $[X_i^\varepsilon,T_j^\varepsilon]$ satisfy commutation relations similar to (2.2) with uniformly bounded coefficients as in (2.3) independent of ε. Hence, all constants in the following inequalities are independent of ε as well. But establishing sub-elliptic metrics similarly as in Section 2 is non-trivial. Indeed, we have to similarly define

$$\{Y_1^\varepsilon,\dots,Y_{m+2n}^\varepsilon\}=\{X_1^\varepsilon,\dots,X_{m+n}^\varepsilon,T_1^\varepsilon,\dots,T_n^\varepsilon\}$$

and define an approximate degree $d_j^\varepsilon=\deg^\varepsilon(Y_j^\varepsilon)\in\{1,2\}$ so that $\deg^\varepsilon(X_j^\varepsilon)=1$ and $\deg^\varepsilon(T_j^\varepsilon)=2$ (as $\varepsilon\to 0^+$ it is easy to see that $\deg^\varepsilon\to\deg$ with deg as defined earlier in Section 2). As before, $Y_1^\varepsilon,\dots,Y_{m+2n}^\varepsilon$ span the whole tangent space and $[Y_j^\varepsilon,Y_k^\varepsilon]\in\operatorname{span}\{Y_l^\varepsilon:d_l^\varepsilon\le d_j^\varepsilon+d_k^\varepsilon\}$ and we are in the set up of Nagel-Stein-Wainger [41]. However, when we have

$$[Y_j^\varepsilon,Y_k^\varepsilon]\;=\sum_{d_l^\varepsilon\le d_j^\varepsilon+d_k^\varepsilon}c_{j,k}^l(\varepsilon)Y_l^\varepsilon,$$

the coefficients $c^l_{j,k}(\varepsilon)$ will be unbounded as $\varepsilon \to 0^+$. In principle this could be a problem as the doubling constant for the metric balls in [41] depends on these coefficients. In other words, letting $d^\varepsilon(I)$ defined with respect to d^ε_j's and $\Lambda^\varepsilon(x,r) = \sum_I |\lambda^\varepsilon_I(x)| r^{d^\varepsilon(I)}$ where $\lambda^\varepsilon_I(x) = \det(a^\varepsilon_{j,k}(x))$ whenever $Y^\varepsilon_{i_j} = \sum_k a^\varepsilon_{j,k}\partial_{x_k}$ as earlier, there exists $c_2(\varepsilon) > c_1(\varepsilon) > 0$ such that we have

$$c_1(\varepsilon)\Lambda^\varepsilon(x,r) \le |B^\varepsilon_r(x)| \le c_2(\varepsilon)\Lambda^\varepsilon(x,r)$$

due to [41], where $B^\varepsilon_r \subset \Omega$ are the metric balls defined similarly as in Section 2. Nevertheless, it has been shown by Capogna-Citti-Rea [5, Proposition 4.4] (see also [3]) that there exists $0 < \bar\varepsilon < 1$ such that for every $0 < \varepsilon < \bar\varepsilon$, the constants $c_1(\varepsilon)$ and $c_2(\varepsilon)$ may be chosen to be independent of ε in compact subsets of Ω. The argument involves a very delicate use of exponentiation of sub-families of the vector fields depending on $\varepsilon < r$ or $r < \varepsilon < \bar\varepsilon$.

Combining all the above statements, we can conclude that the approximate vector fields (4.5) has all the required properties to carry out the estimates of Section 3 for the equation $\mathrm{div}^\varepsilon_H(A_\varepsilon(\mathfrak{X}^\varepsilon u)) = 0$ and the constants are independent of ε. The constants of Poincaré and Sobolev inequalities are also stable under $\varepsilon \to 0$, we refer to [3]. Hence, we conclude from Theorem 4.1 that

$$\sup_{B^\varepsilon_{\sigma r}} |\mathfrak{X}^\varepsilon u|_\varepsilon \le \frac{c(N,p,L)}{(1-\sigma)^{Q/p}} \left(\fint_{B^\varepsilon_r} \left(\delta + |\mathfrak{X}^\varepsilon u|^2_\varepsilon\right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \tag{4.9}$$

holds for any $0 < \sigma < 1$. It is known that $B^\varepsilon_r \to B_r$ in terms of Hausdorff distance as $\varepsilon \to 0^+$, where $B_r \subseteq \Omega$ is the concentric CC-ball (see [25]). This shall be used in the following.

4.2. **Proof of Theorem 1.1.** We are ready to prove Theorem 1.1 by a standard approximation argument and similar as in the Euclidean setting, that incorporates the approximation in order to remove the apriori regularity (3.3).

Proof of Theorem 1.1. Let $B_r = B_r(x_0) \subset \Omega$ be fixed and B^ε_r be a sequence of metric balls with limit B_r as the Hausdorfff distance limit. We shall denote similarly also for concentric balls of smaller radii. Given $u \in \mathcal{S}^{1,p}(\Omega)$ as a weak solution of (1.2), there exists a smooth approximation $\phi_k \in C^\infty(B_r)$ such that $\phi_k \to u$ in $\mathcal{S}^{1,p}(B_r)$ as $k \to \infty$. Now, let u^ε_k be the weak solution of the following Dirichlet problem,

$$\begin{cases} \mathrm{div}^\varepsilon_H(A_\varepsilon(\mathfrak{X}^\varepsilon u^\varepsilon_k)) = 0 & \text{in } B^\varepsilon_{\theta r} \\ u^\varepsilon_k - \phi_k \in \mathcal{S}^{1,p}_0(B^\varepsilon_{\theta r}). \end{cases} \tag{4.10}$$

for some $\theta = \theta(N) < 1$. The choice of test function $u^\varepsilon_k - \phi_k$ on (4.10), yields

$$\int_{B^\varepsilon_{\theta r}} \left\langle A_\varepsilon(\mathfrak{X}^\varepsilon u^\varepsilon_k), \mathfrak{X}^\varepsilon u^\varepsilon_k \right\rangle dx = \int_{B^\varepsilon_{\theta r}} \left\langle A_\varepsilon(\mathfrak{X}^\varepsilon u^\varepsilon_k), \mathfrak{X}^\varepsilon \phi_k \right\rangle dx. \tag{4.11}$$

In this case, we have an ellipticity condition similar to (2.15), which together with the structure condition (4.7) leads to

$$\begin{aligned} \int_{B^\varepsilon_{\theta r}} \left(\delta + |\mathfrak{X}^\varepsilon u^\varepsilon_k|^2_\varepsilon\right)^{\frac{p}{2}} dx &\le c \int_{B^\varepsilon_{\theta r}} \left(\delta + |\mathfrak{X}^\varepsilon \phi_k|^2_\varepsilon\right)^{\frac{p}{2}} dx \le c \int_{B_{\theta r}} \left(\delta + |\mathfrak{X}\phi_k|^2\right)^{\frac{p}{2}} dx + o(\varepsilon) \\ &\le c \int_{B_{\theta r}} \left(\delta + |\mathfrak{X}u|^2\right)^{\frac{p}{2}} dx + o(\varepsilon) + o(1/k) \end{aligned} \tag{4.12}$$

for $c = c(N,p,L) > 0$ and $o(\varepsilon), o(1/k) \to 0$ as $\varepsilon \to 0^+$ and $k \to \infty$. Now, from (4.9), we have

$$\sup_{B^\varepsilon_{\sigma\theta r}} |\mathfrak{X}^\varepsilon u^\varepsilon_k|_\varepsilon \le \frac{c}{(1-\sigma)^{Q/p}} \left(\fint_{B^\varepsilon_{\theta r}} \left(\delta + |\mathfrak{X}^\varepsilon u^\varepsilon_k|^2_\varepsilon\right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \tag{4.13}$$

for some $c = c(N,p,L) > 0$, $\sigma \in (0,1)$. Then a standard argument follows, since (4.12) ensures that there exists $u' \in \mathcal{S}^{1,p}(B_{\theta r})$ such that up to a subsequence $u^\varepsilon_k \rightharpoonup u'$. As $u^\varepsilon_k - \phi_k \in \mathcal{S}^{1,p}_0(B^\varepsilon_{\theta r})$,

hence we have $u' - u \in \mathcal{S}_0^{1,p}(B_{\theta r})$. We have a monotonicity inequality similar to (2.16), using that one can show that u' is a weak solution of (1.2). From uniqueness, $u' = u$. Taking $\varepsilon \to 0^+$ and $k \to \infty$ along the convergent subsequence in (4.13) and (4.12), we conclude

$$\sup_{B_{\sigma\theta r}} |\mathfrak{X}u| \le \frac{c}{(1-\sigma)^{Q/p}} \left(\fint_{B_r} \left(\delta + |\mathfrak{X}u|^2\right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}$$

for some $c = c(N, p, L) > 0$ and (1.5) follows from the above by a simple covering argument. This, together with Remark 2.5, concludes the proof. □

5. The local $C^{1,\alpha}$ regularity

Throughout this section, we denote $u \in \mathcal{S}^{1,p}(\Omega)$ as a weak solution of equation (1.2) satisfying the structure condition (1.3) with $\delta > 0$. Equipped with Theorem 1.1, we have $\mathfrak{X}u \in L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, which together with $\delta > 0$, implies that

$$\begin{aligned} \nu^{-1}|\xi|^2 \le \left\langle DA(\mathfrak{X}u)\,\xi, \xi \right\rangle \le \nu\,|\xi|^2; \\ |A(\mathfrak{X}u)| \le \nu\,|\mathfrak{X}u|, \end{aligned} \tag{5.1}$$

holds locally in Ω, for some $\nu = \nu(N, p, L, \|\mathfrak{X}u\|_{L^\infty}, \delta) > 0$. The equation $\mathrm{div}_H(A(\mathfrak{X}u)) = 0$ with condition (5.1), has been considered by Domokos-Manfredi [15] (and previously by Capogna [2] for the case of the vector fields $X_1, \dots, X_m$ being left invariant with respect to the Heisenberg Group). From [15, Theorem 5.2], we have that

$$\mathfrak{X}u \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m) \cap C^{0,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^m), \qquad Tu \in \mathcal{S}^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap C^{0,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^n), \tag{5.2}$$

which is, in fact, stronger than (3.3). Also, we resume the notation (4.1), i.e. $\mathrm{F}(t) = (\delta + t^2)^{\frac{p-2}{2}}$ throughout this section and therefore all lemmas and apriori estimates of Section 3 are also valid in the settings of this section. The local boundedness of $\mathfrak{X}u$ allows us to define

$$\mu_i(r) = \sup_{B_r} |X_i u|, \quad \mu(r) = \max_{1 \le i \le m} \mu_i(r), \tag{5.3}$$

for any ball $B_r \subset \Omega$ and $i \in \{1, \dots, m\}$.

Note that the function $h(t) = (\delta + t^2)^{\frac{p-2}{2}} t^q = \mathrm{F}(t) t^q$ is non-decreasing on $[0, \infty)$ whenever $q \ge 0$ and $p + q - 2 \ge 0$ and hence, we have the inequality

$$\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^q = \left(\delta + |\mathfrak{X}u|^2\right)^{\frac{p-2}{2}} |\mathfrak{X}u|^q \le c(m,p,q)\left(\delta + \mu(r)^2\right)^{\frac{p-2}{2}} \mu(r)^q \quad \text{in } B_r, \tag{5.4}$$

where $c(m, p, q) = m^{(q+p-2)/2}$ if $p \ge 2$ and $c(m, p, q) = m^{q/2}$ if $1 < p < 2$. Inequalities of the type (5.4) shall be used multiple times in the local estimates.

Similarly as [39], we fix any $l \in \{1, \dots, m\}$ and consider the following double truncation

$$v := \min\left(\mu(r)/8\,,\, \max\left(\mu(r)/4 - X_l u, 0\right)\right). \tag{5.5}$$

In the setting of Euclidean spaces, similar truncations have been used previously in [45, 33], etc. The main property of v that is used in the integral estimates is the fact that it avoids any possible singularities within the set

$$E = \{x \in \Omega : \mu(r)/8 < X_l u < \mu(r)/4\}, \tag{5.6}$$

as easily seen from the following inequality

$$\mu(r)/8 \le |\mathfrak{X}u| \le m^{\frac{1}{2}} \mu(r) \quad \text{in } E \cap B_r, \tag{5.7}$$

which shall be used several times throughout this section. Furthermore, note that

$$\mathfrak{X}v = \begin{cases} -\mathfrak{X}X_l u & \text{a.e. in } E; \\ 0 & \text{a.e. in } \Omega \setminus E, \end{cases} \quad \text{and} \quad Tv = \begin{cases} -TX_l u & \text{a.e. in } E; \\ 0 & \text{a.e. in } \Omega \setminus E, \end{cases} \tag{5.8}$$

which shall be exploited in the following integral estimates. The technique of involving the truncation in the Caccioppoli type estimates is useful particularly for the singular case $1 < p < 2$, whereas the required estimates can be obtained in easier ways for the case $p \geq 2$.

Finally, we remark that properties similar to the above also holds corresponding to

$$\mathsf{v}' = \min\big(\mu(r)/8, \max(\mu(r)/4 + X_l u, 0)\big), \tag{5.9}$$

and all following estimates invovling v shall also hold for v'. Henceforth, we shall only be using the truncation v without loss of generality.

5.1. **Caccioppoli type inequalities with truncation.** We prove certain integral estimates of the gradient by testing (3.4) with the truncation v being incorporated in the test function. Due to (5.7), these will eventually lead to an estimate similar to that of a uniformly elliptic equation without the singular weight.

The following lemma is a truncated analogue of Lemma 3.4. The proof follows along the direction of [39, Lemma 3.1] but it is lengthier due to the commutation relations (2.2).

Lemma 5.1. *For any $\beta, \kappa \geq 0$ and all non-negative $\eta \in C_0^\infty(\Omega)$, we have that*

$$\begin{aligned}
&\int_\Omega \eta^{\beta+2}\mathsf{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa|\mathfrak{X}\mathfrak{X}u|^2\,dx \\
&\qquad\leq c(\beta+1)^2\int_\Omega \eta^\beta\big(|\mathfrak{X}\eta|^2 + \eta|T\eta|\big)\mathsf{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\,dx \\
&\qquad + c(\beta+1)^2(\kappa+1)^2\int_\Omega \eta^{\beta+2}\mathsf{v}^\beta\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}|\mathfrak{X}\mathsf{v}|^2\,dx \\
&\qquad + c(\kappa+1)^4\int_\Omega \eta^{\beta+2}\mathsf{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa\big(|Tu|^2 + |\mathfrak{X}u|^2\big)\,dx,
\end{aligned} \tag{5.10}$$

where v is as in (5.5) and $c = c(N, p, L) > 0$.

Proof. For any $k \in \{1, \dots, m\}$, we take $Y = X_k$ in (3.6) to have

$$\begin{aligned}
\sum_{i,j}\int_\Omega D_j A_i(\mathfrak{X}u)X_jX_kuX_i\varphi\,dx &= \sum_{i,j}\int_\Omega D_jA_i(\mathfrak{X}u)[X_j, X_k]uX_i\varphi\,dx \\
&\quad + \sum_i\int_\Omega A_i(\mathfrak{X}u)[X_i, X_k]\varphi\,dx,
\end{aligned} \tag{5.11}$$

and then, we choose $\varphi = \eta^{\beta+2}\mathsf{v}^{\beta+2}|\mathfrak{X}u|^\kappa X_k u$ with a non-negative test function $\eta \in C_0^\infty(\Omega)$ and $\beta, \kappa \geq 0$ on (5.11) to obtain

$$\begin{aligned}
&\int_\Omega \sum_{i,j}\eta^{\beta+2}\mathsf{v}^{\beta+2}D_jA_i(\mathfrak{X}u)X_jX_kuX_i\big(|\mathfrak{X}u|^\kappa X_ku\big)\,dx \\
&\quad= -(\beta+2)\int_\Omega\sum_{i,j}\eta^{\beta+1}\mathsf{v}^{\beta+2}|\mathfrak{X}u|^\kappa X_kuD_jA_i(\mathfrak{X}u)X_jX_kuX_i\eta\,dx \\
&\qquad -(\beta+2)\int_\Omega\sum_{i,j}\eta^{\beta+2}\mathsf{v}^{\beta+1}|\mathfrak{X}u|^\kappa X_kuD_jA_i(\mathfrak{X}u)X_jX_kuX_i\mathsf{v}\,dx \\
&\qquad + \sum_{i,j}\int_\Omega D_jA_i(\mathfrak{X}u)[X_j, X_k]uX_i\big(\eta^{\beta+2}\mathsf{v}^{\beta+2}|\mathfrak{X}u|^\kappa X_ku\big)\,dx \\
&\qquad - \sum_i\int_\Omega [X_i, X_k]\big(A_i(\mathfrak{X}u)\big)\eta^{\beta+2}\mathsf{v}^{\beta+2}|\mathfrak{X}u|^\kappa X_ku\,dx \\
&\quad= I_1^k + I_2^k + I_3^k + I_4^k,
\end{aligned} \tag{5.12}$$

where integral by parts is applied to the last item in (5.11). Then, by taking summation over $k \in \{1, \dots, m\}$, we end up with

$$\int_\Omega \sum_{i,j,k} \eta^{\beta+2} \mathrm{v}^{\beta+2} D_j A_i(\mathfrak{X}u) X_j X_k u X_i\big(|\mathfrak{X}u|^\kappa X_k u\big)\, dx = \sum_k \sum_{j=1}^4 I_j^k, \tag{5.13}$$

where I_j^k's are the respective terms of the right hand side of (5.12).

In the following, we estimate both sides of (5.13). For the left hand of (5.13), first note that $X_i\big(|\mathfrak{X}u|^\kappa X_k u\big) = |\mathfrak{X}u|^\kappa X_i X_k u + X_i(|\mathfrak{X}u|^\kappa) X_k u$, hence by the structure condition (1.3), we have

$$\sum_{i,j,k} D_j A_i(\mathfrak{X}u) X_j X_l u X_i\big(|\mathfrak{X}u|^\kappa X_k u\big) \ge \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2,$$

which gives us the following estimate for the left hand side of (5.13) as

$$\text{left of (5.13)} \ge \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2\, dx. \tag{5.14}$$

To estimate the right hand side of (5.13), we shall show that I_j^k satisfies

$$\begin{aligned} |I_j^k| \le{}& c\varepsilon \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2\, dx \\ &+ \frac{c}{\varepsilon}(\beta+1)^2 \int_\Omega \eta^\beta \big(|\mathfrak{X}\eta|^2 + \eta|T\eta|\big) \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\, dx \\ &+ \frac{c}{\varepsilon}(\beta+1)^2(\kappa+1)^2 \int_\Omega \eta^{\beta+2} \mathrm{v}^\beta \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2} |\mathfrak{X}\mathrm{v}|^2\, dx \\ &+ \frac{c}{\varepsilon}(\kappa+1)^4 \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa \big(|Tu|^2 + |\mathfrak{X}u|^2\big)\, dx, \end{aligned} \tag{5.15}$$

for each $k \in \{1, \dots, m\}$ and each $j = 1, 2, 3, 4$, where $c = c(N, p, L) > 0$ and $0 < \varepsilon < 1$ is small enough. Then (5.10) follows from the above estimates (5.14) and (5.15) for both sides of (5.13). Thus the proof of the lemma is finished, modulo the proof of (5.15). In the rest, we prove (5.15) in the order of $j = 1, 2, 3, 4$.

First, for $j = 1$, by the structure condition (1.3) and Young's inequality that

$$\begin{aligned} |I_1^k| &\le c(\beta+1) \int_\Omega \eta^{\beta+1} |\mathfrak{X}\eta| \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1} |\mathfrak{X}\mathfrak{X}u|\, dx, \\ &\le c\varepsilon \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2\, dx \\ &\quad + \frac{c}{\varepsilon}(\beta+1)^2 \int_\Omega \eta^\beta |\mathfrak{X}\eta|^2 \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\, dx. \end{aligned} \tag{5.16}$$

Thus (5.15) holds for I_1^k.

Second, when $j = 2$, by the structure condition (1.3) and Young's inequality,

$$\begin{aligned} |I_2^k| &\le c(\beta+1) \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+1} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1} |\mathfrak{X}\mathfrak{X}u| |\mathfrak{X}\mathrm{v}|\, dx, \\ &\le c\varepsilon \int_\Omega \eta^{\beta+2} \mathrm{v}^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2\, dx \\ &\quad + \frac{c}{\varepsilon}(\beta+1)^2 \int_\Omega \eta^{\beta+2} \mathrm{v}^\beta \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2} |\mathfrak{X}\mathrm{v}|^2\, dx, \end{aligned} \tag{5.17}$$

which proves (5.15) for I_2^k.

Third, when $j=3$, we note that

$$\begin{aligned}\big|X_i\big(\eta^{\beta+2}\mathrm{v}^{\beta+2}|\mathfrak{X}u|^{\kappa}X_ku\big)\big| &\le (\kappa+1)\eta^{\beta+2}\mathrm{v}^{\beta+2}|\mathfrak{X}u|^{\kappa}|\mathfrak{X}\mathfrak{X}u| + (\beta+2)\eta^{\beta+1}\mathrm{v}^{\beta+2}|\mathfrak{X}u|^{\kappa+1}|\mathfrak{X}\eta|\\ &\quad + (\beta+2)\eta^{\beta+2}\mathrm{v}^{\beta+1}|\mathfrak{X}u|^{\kappa+1}|\mathfrak{X}\mathrm{v}|.\end{aligned}$$

Thus by the structure condition (1.3) and (2.2), we have

$$\begin{aligned}|I_3^k| \le{}& c(\kappa+1)\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa}|\mathfrak{X}\mathfrak{X}u|\big(|Tu|+|\mathfrak{X}u|\big)\,dx\\ &+ c(\beta+1)\int_\Omega \eta^{\beta+1}|\mathfrak{X}\eta|\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1}\big(|Tu|+|\mathfrak{X}u|\big)\,dx\\ &+ c(\beta+1)\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+1}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1}|\mathfrak{X}\mathrm{v}|\big(|Tu|+|\mathfrak{X}u|\big)\,dx,\end{aligned}$$

from which, by Young's inequality on each term, we get

$$\begin{aligned}|I_3^k| \le{}& c\varepsilon\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa}|\mathfrak{X}\mathfrak{X}u|^2\,dx\\ &+ \frac{c}{\varepsilon}(\kappa+1)^2\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa}\big(|Tu|^2+|\mathfrak{X}u|^2\big)\,dx\\ &+ c(\beta+1)^2\int_\Omega \eta^{\beta}|\mathfrak{X}\eta|^2\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\,dx\\ &+ c(\beta+1)^2\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}|\mathfrak{X}\mathrm{v}|^2\,dx.\end{aligned}\tag{5.18}$$

This proves (5.15) for I_3^k.

Finally, we are left with the case $j=4$ and the estimate is more involved. Recalling our choice of test function $\varphi=\eta^{\beta+2}\mathrm{v}^{\beta+2}|\mathfrak{X}u|^{\kappa}X_ku$, we note that

$$\begin{aligned}I_4^k &= -\sum_i\int_\Omega [X_i,X_k]\big(A_i(\mathfrak{X}u)\big)\varphi\,dx\\ &= -\sum_i\int_\Omega X_iX_k\big(A_i(\mathfrak{X}u)\big)\varphi\,dx + \sum_i\int_\Omega X_kX_i\big(A_i(\mathfrak{X}u)\big)\varphi\,dx\\ &= \sum_i\int_\Omega \Big(X_k\big(A_i(\mathfrak{X}u)\big)X_i\varphi - X_i\big(A_i(\mathfrak{X}u)\big)X_k\varphi\Big)\,dx,\end{aligned}\tag{5.19}$$

where integration by parts is used to get the last term of the above. Let us denote

$$w=\eta^{\beta+2}|\mathfrak{X}u|^{\kappa}X_ku.\tag{5.20}$$

so that we can rewrite $\varphi=\mathrm{v}^{\beta+2}w$ and hence, we have $\mathfrak{X}\varphi=(\beta+2)\mathrm{v}^{\beta+1}w\mathfrak{X}\mathrm{v}+\mathrm{v}^{\beta+2}\mathfrak{X}w$. Using this, note that (5.19) becomes

$$\begin{aligned}I_4^k &= (\beta+2)\int_\Omega \mathrm{v}^{\beta+1}w\sum_i\Big(X_k\big(A_i(\mathfrak{X}u)\big)X_i\mathrm{v} - X_i\big(A_i(\mathfrak{X}u)\big)X_k\mathrm{v}\Big)\,dx\\ &\quad + \int_\Omega \mathrm{v}^{\beta+2}\sum_i\Big(X_k\big(A_i(\mathfrak{X}u)\big)X_iw - X_i\big(A_i(\mathfrak{X}u)\big)X_kw\Big)\,dx\\ &= R^k+P^k,\end{aligned}\tag{5.21}$$

where we denote the first and the second integral in the right hand side of the above by R^k and P^k, respectively. We shall estimate both of them one by one. First, by the structure condition

(1.3), the definition of w as in (5.20), and Young's inequality, we have

$$
\begin{aligned}
|R^k| &\le c(\beta+1)\int_\Omega \eta^{\beta+2}v^{\beta+1}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1}|\mathfrak{X}\mathfrak{X}u||\mathfrak{X}v|\,dx \\
&\le c\varepsilon\int_\Omega \eta^{\beta+2}v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa}|\mathfrak{X}\mathfrak{X}u|^2\,dx \\
&\quad+\frac{c}{\varepsilon}(\beta+1)^2\int_\Omega \eta^{\beta+2}v^{\beta}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}|\mathfrak{X}v|^2\,dx.
\end{aligned}
\tag{5.22}
$$

To estimate P^k, we use integration by parts again to get

$$
\begin{aligned}
P^k &= (\beta+2)\int_\Omega v^{\beta+1}\sum_i A_i(\mathfrak{X}u)\Big(X_ivX_kw - X_kvX_iw\Big)\,dx \\
&\quad+\int_\Omega v^{\beta+2}\sum_i A_i(\mathfrak{X}u)[X_i,X_k]w\,dx \\
&= P_1^k + P_2^k.
\end{aligned}
\tag{5.23}
$$

For P_1^k, we have by the structure condition (1.3) that

$$
\begin{aligned}
|P_1^k| &\le c(\beta+1)(\kappa+1)\int_\Omega \eta^{\beta+2}v^{\beta+1}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1}|\mathfrak{X}\mathfrak{X}u||\mathfrak{X}v|\,dx \\
&\quad+c(\beta+1)^2\int_\Omega \eta^{\beta+1}v^{\beta+1}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}|\mathfrak{X}v||\mathfrak{X}\eta|\,dx
\end{aligned}
$$

from which it follows by Young's inequality that

$$
\begin{aligned}
|P_1^k| &\le c\varepsilon\int_\Omega \eta^{\beta+2}v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa}|\mathfrak{X}\mathfrak{X}u|^2\,dx \\
&\quad+\frac{c}{\varepsilon}(\beta+1)^2(\kappa+1)^2\int_\Omega \eta^{\beta+2}v^{\beta}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}|\mathfrak{X}v|^2\,dx \\
&\quad+c(\beta+1)^2\int_\Omega \eta^{\beta}|\mathfrak{X}\eta|^2v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\,dx.
\end{aligned}
\tag{5.24}
$$

To estimate P_2^k, we note that

$$
\begin{aligned}
[X_i,X_k]w &= \eta^{\beta+2}|\mathfrak{X}u|^{\kappa}[X_i,X_k]X_ku+\kappa\eta^{\beta+2}|\mathfrak{X}u|^{\kappa-1}X_ku[X_i,X_k](|\mathfrak{X}u|) \\
&\quad+(\beta+2)\eta^{\beta+1}|\mathfrak{X}u|^{\kappa}X_ku[X_i,X_k]\eta,
\end{aligned}
$$

and hence, we rewrite

$$
\begin{aligned}
P_2^k &= \int_\Omega \eta^{\beta+2}v^{\beta+2}\sum_i A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa}[X_i,X_k]X_ku\,dx \\
&\quad+\kappa\int_\Omega \eta^{\beta+2}v^{\beta+2}\sum_i A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa-1}X_ku[X_i,X_k](|\mathfrak{X}u|)\,dx \\
&\quad+(\beta+2)\int_\Omega \eta^{\beta+1}v^{\beta+2}\sum_i A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa}X_ku[X_i,X_k]\eta\,dx \\
&= P_{2,1}^k+P_{2,2}^k+P_{2,3}^k,
\end{aligned}
\tag{5.25}
$$

which is estimated similarly as in the last part of the proof of Lemma 3.4. The estimate of $P_{2,3}^k$ is straightforward. Using (3.2) and (2.2), we have

$$
|P_{2,3}^k| \le c(\beta+1)\int_\Omega \eta^{\beta+1}(|T\eta|+|\mathfrak{X}\eta|)v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\,dx,
\tag{5.26}
$$

which is clearly majorised by the right hand side of (5.15).

To estimate the rest of (5.25), we shall use $[X_i, X_k]X_k u = X_k[X_i, X_k]u + [[X_i, X_k], X_k]u$ and

$$[X_i, X_k](|\mathfrak{X}u|) = |\mathfrak{X}u|^{-1}\sum_{j=1}^{m} X_j u[X_i, X_k]X_j u = |\mathfrak{X}u|^{-1}\sum_{j=1}^{m} X_j u\big(X_j[X_i, X_k]u + [[X_i, X_k], X_j]u\big),$$

to split the first two terms as $P^k_{2,1} = P^k_{2,1,1} + P^k_{2,1,2}$ and $P^k_{2,2} = P^k_{2,2,1} + P^k_{2,2,2}$ so that the second terms i.e. $P^k_{2,1,2}$ and $P^k_{2,2,2}$ contain the terms with the commutator of commutators.

Now we rewrite the first terms of $P^k_{2,1} + P^k_{2,2}$ using integral by parts as

$$\begin{aligned} P^k_{2,1,1} + P^k_{2,2,1} &= \sum_i \int_\Omega \eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^\kappa X_k[X_i, X_k]u\, dx \\ &\quad + \kappa \sum_{i,j} \int_\Omega \eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa-2} X_k u X_j u X_j[X_i, X_k]u\, dx \\ &= -\sum_i \int_\Omega X_k\Big(\eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^\kappa\Big)[X_i, X_k]u\, dx \\ &\quad - \kappa \sum_{i,j} \int_\Omega X_j\Big(\eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa-2} X_k u X_j u\Big)[X_i, X_k]u\, dx. \end{aligned} \tag{5.27}$$

Then, further computation of the terms in (5.27) followed by the use of (3.2) and (2.2) yields

$$\begin{aligned} |P^k_{2,1,1}| + |P^k_{2,2,1}| &\le c(\kappa+1)^2 \int_\Omega \eta^{\beta+2} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|\big(|\mathfrak{X}u| + |Tu|\big)\, dx \\ &\quad + c(\beta+1)(\kappa+1) \int_\Omega \eta^{\beta+2} v^{\beta+1} |\mathfrak{X}u|^{\kappa+1} |\mathfrak{X}v|\big(|\mathfrak{X}u| + |Tu|\big)\, dx \\ &\quad + c(\beta+1)(\kappa+1) \int_\Omega \eta^{\beta+1} v^{\beta+2} |\mathfrak{X}u|^{\kappa+1} |\mathfrak{X}\eta|\big(|\mathfrak{X}u| + |Tu|\big)\, dx, \end{aligned}$$

and then, by Young's inequality on each term, we get

$$\begin{aligned} |P^k_{2,1,1}| + |P^k_{2,2,1}| &\le c\varepsilon \int_\Omega \eta^{\beta+2} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa |\mathfrak{X}\mathfrak{X}u|^2\, dx \\ &\quad + \frac{c}{\varepsilon}(\kappa+1)^4 \int_\Omega \eta^{\beta+2} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^\kappa \big(|\mathfrak{X}u|^2 + |Tu|^2\big)\, dx \\ &\quad + c(\beta+1)^2 \int_\Omega \eta^{\beta+2} v^{\beta} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2} |\mathfrak{X}v|^2\, dx \\ &\quad + c(\beta+1)^2 \int_\Omega \eta^{\beta} |\mathfrak{X}\eta|^2 v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+2}\, dx \end{aligned} \tag{5.28}$$

Now we are only left with

$$\begin{aligned} P^k_{2,1,2} + P^k_{2,2,2} &= \sum_i \int_\Omega \eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^\kappa [[X_i, X_k], X_k]u\, dx \\ &\quad + \kappa \sum_{i,j} \int_\Omega \eta^{\beta+2} v^{\beta+2} A_i(\mathfrak{X}u)|\mathfrak{X}u|^{\kappa-2} X_k u X_j u [[X_i, X_k], X_j]u\, dx \end{aligned} \tag{5.29}$$

To estimate (5.29), we use (3.2) and (2.5) to get

$$|P^k_{2,1,2}| + |P^k_{2,2,2}| \le c(\kappa+1) \int_\Omega \eta^{\beta+2} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa+1}\big(|\mathfrak{X}u| + |Tu|\big)\, dx, \tag{5.30}$$

which is estimated easily to get the right hand side of (5.15). Thus, combining the estimates above we can see that the claimed estimate (5.15) holds also for I^k_4 and the proof is finished. □

The following lemma is a truncated analogue of Corollary 3.11.

Lemma 5.2. *Let $B_r \subset \Omega$ and $\eta \in C_0^\infty(B_r)$ be fixed. Then, for any $\tau \in (\frac{1}{2}, 1)$ and $\beta \geq 0$,*

$$\begin{aligned}
&\int_\Omega \eta^{\tau(\beta+2)+4} v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4 |\mathfrak{X}Tu|^2 \, dx \\
&\qquad \leq c_0 \int_\Omega \eta^{\tau(\beta+2)+4} v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4 \big(|Tu|^2 + |\mathfrak{X}u|^2\big) \, dx \\
&\qquad\quad + c\mu(r)^{4(1-\tau)} K_\eta^{2-\tau} \Big(\int_{B_r} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\frac{2}{1-\tau}} \, dx \Big)^{1-\tau} \mathcal{I}^\tau
\end{aligned} \tag{5.31}$$

holds for any small enough $c_0 = c_0(N, p, L, \tau) > 0$, where $c = c(N, p, L, \tau) > 0$ is large enough, $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty}$, and

$$\begin{aligned}
\mathcal{I} &= c(\beta+1)^2 \int_\Omega \eta^{\beta+2} \big(\eta^2 + |\mathfrak{X}\eta|^2 + \eta|T\eta|\big) v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4 \, dx \\
&\quad + c(\beta+1)^2 \int_\Omega \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4 |\mathfrak{X}v|^2 \, dx \\
&\quad + c \int_\Omega \eta^{\beta+4} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2 |Tu|^2 \, dx.
\end{aligned} \tag{5.32}$$

Proof. Let us denote the left hand side of (5.31) by M, i.e.

$$M = \int_\Omega \eta^{\tau(\beta+2)+4} \, v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4 |\mathfrak{X}Tu|^2 \, dx, \tag{5.33}$$

for any fixed $1/2 < \tau < 1$ and $\beta \geq 0$. Now, we recall (3.8), i.e.

$$\begin{aligned}
\sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u) X_j T_k u X_i \varphi \, dx &= \sum_{i,j} \int_\Omega D_j A_i(\mathfrak{X}u)[X_j, T_k] u X_i \varphi \, dx \\
&\quad - \sum_{i,j} \int_\Omega \varphi D_j A_i(\mathfrak{X}u)[X_i, T_k] X_j u \, dx.
\end{aligned} \tag{5.34}$$

We use $\varphi = \varphi_k = \eta^{\tau(\beta+2)+4} \, v^{\tau(\beta+4)} |\mathfrak{X}u|^4 \, T_k u$ on (5.34) and take summation over $k \in \{1, \dots, n\}$, to obtain

$$\begin{aligned}
&\sum_{i,j,k} \int_\Omega \eta^{\tau(\beta+2)+4} \, v^{\tau(\beta+4)} |\mathfrak{X}u|^4 D_j A_i(\mathfrak{X}u) X_j T_k u \, X_i T_k u \, dx \\
&\quad = -(\tau(\beta+2)+4) \sum_{i,j,k} \int_\Omega \eta^{\tau(\beta+2)+3} \, v^{\tau(\beta+4)} |\mathfrak{X}u|^4 T_k u D_j A_i(\mathfrak{X}u) X_j T_k u \, X_i \eta \, dx \\
&\qquad - \tau(\beta+4) \sum_{i,j,k} \int_\Omega \eta^{\tau(\beta+2)+4} \, v^{\tau(\beta+4)-1} |\mathfrak{X}u|^4 T_k u D_j A_i(\mathfrak{X}u) X_j T_k u \, X_i v \, dx \\
&\qquad - 4 \sum_{i,j,k,l} \int_\Omega \eta^{\tau(\beta+2)+4} \, v^{\tau(\beta+4)} |\mathfrak{X}u|^2 X_l u T_k u D_j A_i(\mathfrak{X}u) X_j T_k u \, X_i X_l u \, dx \\
&\qquad + \sum_{i,j,k} \int_\Omega D_j A_i(\mathfrak{X}u)[X_j, T_k] u X_i \varphi_k \, dx \\
&\qquad - \sum_{i,j,k} \int_\Omega \varphi_k D_j A_i(\mathfrak{X}u)[X_i, T_k] X_j u \, dx \\
&\quad = K_1 + K_2 + K_3 + K_4 + K_5,
\end{aligned} \tag{5.35}$$

where the integrals in the right hand side of (5.35) are denoted by K_1, K_2, K_3, K_4, K_5 in order. To prove the lemma, we estimate both sides of (5.35) as follows.

For the left hand side, we have by the structure condition (1.3) that

$$\text{left of (5.35)} \geq \int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}Tu|^2\,dx = M, \tag{5.36}$$

and for the right hand side of (5.35), we estimate each item K_i, one by one. Let us denote

$$\tilde{K} = \int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2|\mathfrak{X}Tu|^2\,dx, \tag{5.37}$$

which shall appear frequently in the following.

First, we estimate K_1 by the structure condition (1.3) and Hölder's inequality, to get

$$\begin{aligned} |K_1| &\leq c(\beta+1)\int_\Omega \eta^{\tau(\beta+2)+3}\, v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu||\mathfrak{X}Tu||\mathfrak{X}\eta|\,dx \\ &\leq c(\beta+1)\Big(\int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1}{2}} \\ &\qquad \times \Big(\int_\Omega \eta^{\beta+2} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}\eta|^2\,dx\Big)^{\frac{1}{2}} \\ &= c(\beta+1)\tilde{K}^{\frac{1}{2}}\Big(\int_\Omega \eta^{\beta+2} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}\eta|^2\,dx\Big)^{\frac{1}{2}}. \end{aligned} \tag{5.38}$$

Second, we estimate K_2 also by the structure condition (1.3) and Hölder's inequality,

$$\begin{aligned} |K_2| &\leq c(\beta+1)\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)-1} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu||\mathfrak{X}Tu||\mathfrak{X}v|dx \\ &\leq c(\beta+1)\Big(\int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1}{2}} \\ &\qquad \times \Big(\int_\Omega \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}v|^2\,dx\Big)^{\frac{1}{2}} \\ &= c(\beta+1)\tilde{K}^{\frac{1}{2}}\Big(\int_\Omega \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}v|^2\,dx\Big)^{\frac{1}{2}}. \end{aligned} \tag{5.39}$$

Similarly, we estimate K_3 by structure condition (1.3) and Hölder's inequality. We have

$$\begin{aligned} |K_3| &\leq c\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^3|Tu||\mathfrak{X}Tu||\mathfrak{X}\mathfrak{X}u|\,dx \\ &\leq c\Big(\int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1}{2}} \\ &\qquad \times \Big(\int_\Omega \eta^{\beta+4} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{1}{2}} \\ &= c\tilde{K}^{\frac{1}{2}}\Big(\int_\Omega \eta^{\beta+4} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|\mathfrak{X}\mathfrak{X}u|^2\,dx\Big)^{\frac{1}{2}} \leq c\tilde{K}^{\frac{1}{2}}\mathcal{I}^{\frac{1}{2}}, \end{aligned} \tag{5.40}$$

where the last inequality of the above follows from Lemma 5.1 with

$$\begin{aligned} \mathcal{I} =& c(\beta+1)^2\int_\Omega \eta^{\beta+2}\big(\eta^2+|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\, v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\,dx \\ &+ c(\beta+1)^2\int_\Omega \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}v|^2\,dx \\ &+ c\int_\Omega \eta^{\beta+4} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^2\,dx, \end{aligned}$$

which is a rearrangement of the right hand side of (5.10) (with $\kappa = 2$ and $\beta \mapsto \beta+2$).

In fact, we can combine (5.38), (5.39) and (5.40) to obtain

$$|K_1| + |K_2| + |K_3| \le c\tilde{K}^{\frac{1}{2}}\mathcal{I}^{\frac{1}{2}}. \tag{5.41}$$

Then $\tilde{K}$ is estimated by Hölder's inequality as

$$\begin{aligned}\tilde{K} &= \int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2|\mathfrak{X}Tu|^2\,dx\\ &\le \Big(\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{2\tau-1}{\tau}}\\ &\quad\times\Big(\int_\Omega \eta^{\frac{2\tau}{1-\tau}+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^{\frac{2\tau}{1-\tau}}|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1-\tau}{\tau}}\\ &= M^{\frac{2\tau-1}{\tau}}H^{\frac{1-\tau}{\tau}},\end{aligned} \tag{5.42}$$

where M is as in (5.33) and H is the second integral on the right hand side of (5.42), i.e.

$$H = \int_\Omega \eta^{\frac{2\tau}{1-\tau}+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^{\frac{2\tau}{1-\tau}}|\mathfrak{X}Tu|^2\,dx, \tag{5.43}$$

which is estimated using Corollary 3.10 and (5.4) with $q = 2/(1-\tau)$, as

$$H \le c\mu(r)^4\int_\Omega \eta^{q+2}\mathrm{F}(|\mathfrak{X}u|)|Tu|^{q-2}|\mathfrak{X}Tu|^2\,dx \le cK_\eta^{\frac{q+2}{2}}\mu(r)^4\int_{B_r}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^q\,dx \tag{5.44}$$

where $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty}$ and $c = c(N,p,L,\tau) > 0$.

To estimate the rest, we note that

$$\begin{aligned}K_4 &= \sum_{i,j,k}\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)}|\mathfrak{X}u|^4 D_jA_i(\mathfrak{X}u)[X_j,T_k]u\, X_iT_ku\,dx\\ &\quad+(\tau(\beta+2)+4)\sum_{i,j,k}\int_\Omega \eta^{\tau(\beta+2)+3}\, v^{\tau(\beta+4)}|\mathfrak{X}u|^4T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]u\, X_i\eta\,dx\\ &\quad+\tau(\beta+4)\sum_{i,j,k}\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)-1}|\mathfrak{X}u|^4T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]u\, X_i\mathrm{v}\,dx\\ &\quad+4\sum_{i,j,k,l}\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)}|\mathfrak{X}u|^2X_luT_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]u\, X_iX_lu\,dx\\ &= K_{4,0}+K_{4,1}+K_{4,2}+K_{4,3}.\end{aligned} \tag{5.45}$$

Then $K_{4,0}$ is estimated by (1.3), (2.2) and Hölder's inequality as

$$\begin{aligned}|K_{4,0}| &\le c(\beta+1)\int_\Omega \eta^{\tau(\beta+2)+4}\, v^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|+|\mathfrak{X}u|\big)|\mathfrak{X}Tu|\,dx\\ &\le c(\beta+1)\Big(\int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|^2+|\mathfrak{X}u|^2\big)|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1}{2}}\\ &\qquad\times\Big(\int_\Omega \eta^{\beta+4}v^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\,dx\Big)^{\frac{1}{2}}\\ &= c(\beta+1)\bar{K}^{\frac{1}{2}}\Big(\int_\Omega \eta^{\beta+4}v^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\,dx\Big)^{\frac{1}{2}} \le c\bar{K}^{\frac{1}{2}}\mathcal{I}^{\frac{1}{2}},\end{aligned}$$

where $\bar{K}$ is the term similar to but larger than $\tilde{K}$ and is estimated similarly as

$$\begin{aligned}\bar{K} &= \int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, \mathrm{v}^{(2\tau-1)(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|^2+|\mathfrak{X}u|^2\big)|\mathfrak{X}Tu|^2\,dx\\ &\le c\Big(\int_\Omega \eta^{\tau(\beta+2)+4}\mathrm{v}^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{2\tau-1}{\tau}}\\ &\quad\times\Big(\int_\Omega \eta^{\frac{2\tau}{1-\tau}+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|^{\frac{2\tau}{1-\tau}}+|\mathfrak{X}u|^{\frac{2\tau}{1-\tau}}\big)|\mathfrak{X}Tu|^2\,dx\Big)^{\frac{1-\tau}{\tau}}\\ &= cM^{\frac{2\tau-1}{\tau}}(H+H')^{\frac{1-\tau}{\tau}},\end{aligned}\tag{5.46}$$

for some $c=c(N,p,L,\tau)>0$, H is in (5.43) and

$$H'=\int_\Omega \eta^{\frac{2\tau}{1-\tau}+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{4+\frac{2\tau}{1-\tau}}|\mathfrak{X}Tu|^2\,dx,\tag{5.47}$$

which is estimated using Corollary 3.11 and (5.4) with $q=2/(1-\tau)$, as

$$H'\le c\mu(r)^4\int_\Omega \eta^{q+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{q-2}|\mathfrak{X}Tu|^2\,dx\le cK_\eta^{\frac{q+2}{2}}\mu(r)^4\int_{B_r}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^q\,dx\tag{5.48}$$

where $K_\eta=\|\eta\|_{L^\infty}^2+\|\mathfrak{X}\eta\|_{L^\infty}^2+\|\eta T\eta\|_{L^\infty}$ and $c=c(N,p,L,\tau)>0$.

Since we have $|[X_j,T_k]u|^2\le c\big(|Tu|^2+|\mathfrak{X}u|^2\big)$ from (2.2), it is not hard to see that the estimates of $K_{4,1},K_{4,2},K_{4,3}$ can be performed by Hölder's inequality in exactly the same way as that of K_1,K_2,K_3, i.e. (5.38),(5.39),(5.40) respectively, where $|\mathfrak{X}Tu|^2$ being replaced by $\big(|Tu|^2+|\mathfrak{X}u|^2\big)$. In other words, we have

$$|K_{4,1}|+|K_{4,2}|+|K_{4,3}|\le c\mathcal{K}^{\frac12}\mathcal{I}^{\frac12},\tag{5.49}$$

where $\mathcal{I}$ is as in (5.32) and

$$\mathcal{K}=\int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\,\mathrm{v}^{(2\tau-1)(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2\big(|Tu|^2+|\mathfrak{X}u|^2\big)\,dx,\tag{5.50}$$

which can be estimated similarly as above to obtain

$$\begin{aligned}\mathcal{K} &\le c\Big(\int_\Omega \eta^{\tau(\beta+2)+4}\mathrm{v}^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2\,dx\Big)^{\frac{2\tau-1}{\tau}}\\ &\quad\times\Big(\int_\Omega \eta^{\frac{2\tau}{1-\tau}+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|^{\frac{2\tau}{1-\tau}}+|\mathfrak{X}u|^{\frac{2\tau}{1-\tau}}\big)|Tu|^2\,dx\Big)^{\frac{1-\tau}{\tau}}\\ &= c\bar{M}^{\frac{2\tau-1}{\tau}}\bar{H}^{\frac{1-\tau}{\tau}},\end{aligned}\tag{5.51}$$

where $\bar{H}$ is the term analogous to H and H' as in (5.43) and (5.47) which can be estimated similarly by (3.50) and

$$\bar{M}=\int_\Omega \eta^{\tau(\beta+2)+4}\,\mathrm{v}^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|^2\,dx.\tag{5.52}$$

Finally, we recall that

$$K_5=-\sum_{i,j,k}\int_\Omega \eta^{\tau(\beta+2)+4}\,\mathrm{v}^{\tau(\beta+4)}|\mathfrak{X}u|^4\,T_kuD_jA_i(\mathfrak{X}u)[X_i,T_k]X_ju\,dx,$$

which is estimated by (1.3) and (2.4) as

$$\begin{aligned}|K_5| &\le c\int_\Omega \eta^{\tau(\beta+2)+4}\,\mathrm{v}^{\tau(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|Tu|\big(|\mathfrak{X}\mathfrak{X}u|+|\mathfrak{X}Tu|+|\mathfrak{X}u|+|Tu|\big)\,dx\\ &\le c(\mathcal{K}_0^{\frac12}+\tilde{K}^{\frac12}+\bar{K}^{\frac12}+\mathcal{K}^{\frac12})\mathcal{I}^{\frac12},\end{aligned}\tag{5.53}$$

where the latter inequality is obtained from Hölder's inequality and the previous estimates, where $\tilde{K}, \bar{K}$ and $\mathcal{K}$ are as in (5.46) and (5.50), $\mathcal{I}$ is as in (5.32) and the first term of (5.53) estimated similarly as

$$\mathcal{K}_0 = \int_\Omega \eta^{(2\tau-1)(\beta+2)+6}\, v^{(2\tau-1)(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^8\, dx \le c M_0^{\frac{2\tau-1}{\tau}} H_0^{\frac{1-\tau}{\tau}}, \tag{5.54}$$

where M_0 and H_0 are the analogous terms containing only $\mathfrak{X}u$ in place of Tu or $\mathfrak{X}Tu$. Combining all the estimates of the above, we finally have

$$\begin{aligned} M &\le c\big(\tilde{K}^{\frac12} + \bar{K}^{\frac12} + \mathcal{K}^{\frac12} + \mathcal{K}_0^{\frac12}\big)\mathcal{I}^{\frac12} \\ &\le c\Big(M^{\frac{2\tau-1}{2\tau}} + \bar{M}^{\frac{2\tau-1}{2\tau}} + M_0^{\frac{2\tau-1}{2\tau}}\Big)\big(H + H' + \bar{H} + H_0\big)^{\frac{1-\tau}{2\tau}}\mathcal{I}^{\frac12} \end{aligned} \tag{5.55}$$

which, by Young's inequality, leads to

$$M \le c_0(\bar{M} + M_0) + c\big(H + H' + \bar{H} + H_0\big)^{1-\tau}\mathcal{I}^{\tau},$$

which, with the estimates (5.44) and (5.48), implies (5.31) and the proof is finished. □

Remark 5.3. For the special case when $X_1, \dots, X_m$ are left invariant with respect to a step 2 Carnot Group, since we have $[X_i, T_k] = 0$, hence it corresponds to the case of $c_0 = 0$ for (5.31). However, it does not make any difference because of the following corollary.

Corollary 5.4. *Let $B_r \subset \Omega$ and $\eta \in C_0^\infty(B_r)$ be fixed. Then, for any $\tau \in (\frac12, 1)$ and $\beta \ge 0$,*

$$\int_\Omega \eta^{\tau(\beta+2)+4} v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}Tu|^2\, dx \le c\mu(r)^{4(1-\tau)} K_\eta^{2-\tau}\Big(\int_{B_r} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\frac{2}{1-\tau}}\, dx\Big)^{1-\tau}\mathcal{I}^{\tau}$$

where $\mathcal{I}$ is as in (5.32), $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty}$ and $c = c(N, p, L, \tau) > 0$.

Proof. The first term in the right hand side of (5.31) can be estimated by Hölder's inequality as

$$\begin{aligned} &\int_\Omega \eta^{\tau(\beta+2)+4} v^{\tau(\beta+4)} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(|Tu|^2 + |\mathfrak{X}u|^2\big)\, dx \\ &\le \Big(\int_\Omega \eta^{\beta+4} v^{\beta+4} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2\big(|Tu|^2 + |\mathfrak{X}u|^2\big)\, dx\Big)^{\tau} \\ &\quad\times \Big(\int_\Omega \eta^{2+\frac{2}{1-\tau}} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{2+\frac{2}{1-\tau}}\big(|Tu|^2 + |\mathfrak{X}u|^2\big)\, dx\Big)^{1-\tau} \le c\mathcal{I}^{\tau}(\bar{H} + H_0)^{1-\tau}, \end{aligned} \tag{5.56}$$

where $\mathcal{I}$ is as in (5.32) and $\bar{H}, H_0$ are as in the proof of Lemma 5.2, which can be estimated easily by (3.50) as earlier. This concludes the proof. □

5.2. **The main Lemma.** We are ready to prove the main lemma for the proof of the $C^{1,\alpha}$ regularity. This is a Caccioppoli type estimate of the truncation that is devoid of the weight $\mathrm{F}(|\mathfrak{X}u|) = (\delta + |\mathfrak{X}u|^2)^{\frac{p-2}{2}}$ which is similar to estimates for equations with uniform ellipticity.

Towards this, we fix $B_r \subset \Omega$ and use a standard test function $\eta \in C_0^\infty(B_r)$ that satisfies

$$\begin{aligned} &0 \le \eta \le 1 \quad \text{in } B_r, \ \eta \equiv 1 \quad \text{in } B_{r/2}, \\ &|\mathfrak{X}\eta| \le 4/r \quad \text{and} \quad |\mathfrak{X}\mathfrak{X}\eta| \le 16N/r^2 \quad \text{in } B_r, \end{aligned} \tag{5.57}$$

and hence from (2.6), $K_\eta = \|\eta\|_{L^\infty}^2 + \|\mathfrak{X}\eta\|_{L^\infty}^2 + \|\eta T\eta\|_{L^\infty} \le c/r^2$ for large enough $c = c(N) > 0$.

The following is the main lemma of this section, which has been proved previously in [39] for the special case when $X_1, \dots, X_m$ are left invariant with respect to the Heisenberg Group. In the Euclidean case, similar estimates have been proved earlier, see [45, 33], etc.

Lemma 5.5. *Let v be the truncation as in (5.5) and $\eta \in C_0^\infty(B_r)$ satisfy (5.57). Then we have the inequality*

$$\int_{B_r} \eta^{\beta+4} v^{\beta+2} |\mathfrak{X} v|^2 \, dx \le \frac{c}{r^2} (\beta+1)^2 \mu(r)^4 |B_r|^{1-1/\gamma} \Big(\int_{B_r} \eta^{\gamma\beta} v^{\gamma\beta} \, dx \Big)^{1/\gamma}, \tag{5.58}$$

for any $\beta \ge 0$ and $\gamma > 1$, where $c = c(N, p, L, \gamma) > 0$ is a constant.

Proof. We shall assume $1 < \gamma < 3/2$ without loss of generality, since (5.58) can be extended to the full range by applying Hölder's inequality to the right hand side.

Recall that $v = \min\big(\mu(r)/8\,,\, \max\left(\mu(r)/4 - X_l u, 0\right)\big)$ as in (5.5), for some fixed $l \in \{1, \dots, m\}$. For this l, we test (3.7) with the function $\varphi = \eta^{\beta+4} v^{\beta+3}$ for $\beta \ge 0$ and $\eta \in C_0^\infty(B_r)$ as given, which is allowed since it follows from (5.2) and (5.8) that $\mathfrak{X} v \in L^2_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ and $Tv \in L^2_{\mathrm{loc}}(\Omega; \mathbb{R}^n)$. Thus, we obtain

$$\begin{aligned}
&-(\beta+3) \sum_{i,j} \int_\Omega \eta^{\beta+4} v^{\beta+2} D_j A_i(\mathfrak{X} u) X_j X_l u X_i v \, dx \\
&\qquad = (\beta+4) \sum_{i,j} \int_\Omega \eta^{\beta+3} v^{\beta+3} D_j A_i(\mathfrak{X} u) X_j X_l u X_i \eta \, dx \\
&\qquad\quad - (\beta+4) \sum_{i,j} \int_\Omega \eta^{\beta+3} v^{\beta+3} D_j A_i(\mathfrak{X} u) [X_j, X_l] u X_i \eta \, dx \\
&\qquad\quad - (\beta+3) \sum_{i,j} \int_\Omega \eta^{\beta+4} v^{\beta+2} D_j A_i(\mathfrak{X} u) [X_j, X_l] u X_i v \, dx \\
&\qquad\quad - \sum_i \int_\Omega A_i(\mathfrak{X} u) [X_i, X_l] (\eta^{\beta+4} v^{\beta+3}) \, dx.
\end{aligned} \tag{5.59}$$

We can combine the first two terms of the right hand side and rewrite (5.59) as

$$\begin{aligned}
&-(\beta+3) \sum_{i,j} \int_\Omega \eta^{\beta+4} v^{\beta+2} D_j A_i(\mathfrak{X} u) X_j X_l u X_i v \, dx \\
&\qquad = (\beta+4) \sum_{i,j} \int_\Omega \eta^{\beta+3} v^{\beta+3} D_j A_i(\mathfrak{X} u) X_l X_j u X_i \eta \, dx \\
&\qquad\quad - (\beta+3) \sum_{i,j} \int_\Omega \eta^{\beta+4} v^{\beta+2} D_j A_i(\mathfrak{X} u) [X_j, X_l] u X_i v \, dx \\
&\qquad\quad - \sum_i \int_\Omega A_i(\mathfrak{X} u) [X_i, X_l] (\eta^{\beta+4} v^{\beta+3}) \, dx \\
&\qquad = I_1 + I_2 + I_3,
\end{aligned} \tag{5.60}$$

where we denote the terms in the right hand side of (5.60) by I_1, I_2, I_3, respectively and we estimate both sides of (5.60) in the following.

For the left hand side, from (5.8), the structure condition (1.3) and (5.7), we have

$$\begin{aligned}
\text{left of (5.60)} &\ge (\beta+3) \int_E \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X} u|) |\mathfrak{X} v|^2 \, dx \\
&\ge c_0 (\beta+1) \mathrm{F}(\mu(r)) \int_{B_r} \eta^{\beta+4} v^{\beta+2} |\mathfrak{X} v|^2 \, dx,
\end{aligned} \tag{5.61}$$

for some constant $c_0 = c_0(N,p,L) > 0$, where E as in (5.6). For the right hand side of (5.60), we claim that each item I_1, I_2, I_3 satisfies

$$
\begin{aligned}
|I_k| \le & \frac{c_0}{6}(\beta+1)\mathrm{F}(\mu(r)) \int_{B_r} \eta^{\beta+4}\mathrm{v}^{\beta+2}|\mathfrak{X}\mathrm{v}|^2\,dx \\
& + \frac{c}{r^2}(\beta+1)^3\mathrm{F}(\mu(r))\mu(r)^4|B_r|^{1-1/\gamma}\Big(\int_{B_r}\eta^{\gamma\beta}\mathrm{v}^{\gamma\beta}\,dx\Big)^{1/\gamma},
\end{aligned}
\tag{5.62}
$$

where $k = 1,2,3$, $1 < \gamma < 3/2$ and $c = c(N,p,L,\gamma) > 0$. Then (5.58) follows from the estimate (5.61) for the left hand side of (5.60)) and the above claim (5.62)) for each item in the right, thereby completing the proof. Thus, we are only left with proving the claim (5.62). To this end, we estimate each I_k one by one, in the rest of the proof.

First, from chain rule, notice that

$$
\begin{aligned}
I_1 &= (\beta+4)\sum_i \int_\Omega \eta^{\beta+3}\mathrm{v}^{\beta+3}X_l\big(A_i(\mathfrak{X}u)\big)X_i\eta\,dx \\
&= -(\beta+4)\sum_i \int_\Omega A_i(\mathfrak{X}u)X_l\big(\eta^{\beta+3}\mathrm{v}^{\beta+3}X_i\eta\big)\,dx,
\end{aligned}
$$

where the latter equality is due to integral by parts. Then by structure condition (1.3),

$$
\begin{aligned}
|I_1| \le & c(\beta+1)^2\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+3}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|\big(|\mathfrak{X}\eta|^2+\eta|\mathfrak{X}\mathfrak{X}\eta|\big)\,dx \\
& + c(\beta+1)^2\int_\Omega \eta^{\beta+3}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u||\mathfrak{X}\mathrm{v}||\mathfrak{X}\eta|\,dx \\
\le & \frac{c}{r^2}(\beta+1)^2\mathrm{F}(\mu(r))\mu(r)^4\int_{B_r}\eta^\beta\mathrm{v}^\beta\,dx \\
& + \frac{c}{r}(\beta+1)^2\mathrm{F}(\mu(r))\mu(r)^2\int_{B_r}\eta^{\beta+2}\mathrm{v}^{\beta+1}|\mathfrak{X}\mathrm{v}|\,dx,
\end{aligned}
\tag{5.63}
$$

where $c = c(N,p,L) > 0$. For the latter inequality of (5.63), we have used (5.4). Now we apply Young's inequality to the last term of (5.63) to end up with

$$
\begin{aligned}
|I_1| \le & \frac{c_0}{6}(\beta+1)\mathrm{F}(\mu(r))\int_{B_r}\eta^{\beta+4}\mathrm{v}^{\beta+2}|\mathfrak{X}\mathrm{v}|^2\,dx \\
& + \frac{c}{r^2}(\beta+1)^3\mathrm{F}(\mu(r))\mu(r)^4\int_{B_r}\eta^\beta\mathrm{v}^\beta\,dx,
\end{aligned}
\tag{5.64}
$$

where $c = c(N,p,L) > 0$ and c_0 is the same constant as in (5.61). The claimed estimate (5.62) for I_1, follows from the above estimate (5.64) and Hölder's inequality.

Second, we estimate I_2 by structure condition (1.3) and (2.2), to have

$$
|I_2| \le c(\beta+1)\int_\Omega \eta^{\beta+4}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}\mathrm{v}|\big(|Tu|+|\mathfrak{X}u|\big)\,dx,
$$

and it is important to note that the above integral is supported on E from (5.8). Therefore, it follows by Hölder's inequality that

$$
\begin{aligned}
|I_2| \le & c(\beta+1)\Big(\int_E \eta^{\beta+4}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}\mathrm{v}|^2\,dx\Big)^{\frac12}\Big(\int_E \eta^{\gamma(\beta+2)}\mathrm{v}^{\gamma(\beta+2)}\mathrm{F}(|\mathfrak{X}u|)\,dx\Big)^{\frac{1}{2\gamma}} \\
& \times\Big(\int_\Omega \eta^q\,\mathrm{F}(|\mathfrak{X}u|)\big(|Tu|^q+|\mathfrak{X}u|^q\big)\,dx\Big)^{\frac1q},
\end{aligned}
\tag{5.65}
$$

where $q = 2\gamma/(\gamma-1)$. The fact that the first two integrals are on E is crucial as (5.7) can be exploited to carry out the following estimates,

$$\int_E \eta^{\beta+4} v^{\beta+2} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}v|^2\,dx \le c\mathrm{F}(\mu(r)) \int_{B_r} \eta^{\beta+4} v^{\beta+2}|\mathfrak{X}v|^2\,dx, \tag{5.66}$$

and

$$\int_E \eta^{\gamma(\beta+2)} v^{\gamma(\beta+2)} \mathrm{F}(|\mathfrak{X}u|)\,dx \le c\mathrm{F}(\mu(r))\mu(r)^{2\gamma} \int_{B_r} \eta^{\gamma\beta} v^{\gamma\beta}\,dx, \tag{5.67}$$

which cannot be done otherwise (unless $t \mapsto \mathrm{F}(t)$ is increasing which corresponds to $p \ge 2$ for the p-Laplacian). The last integral of (5.65) is estimated from (3.50) and (5.4) to get

$$\int_\Omega \eta^q\, \mathrm{F}(|\mathfrak{X}u|)\big(|Tu|^q + |\mathfrak{X}u|^q\big)\,dx \le \frac{c}{r^q}\int_{B_r} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^q\,dx \le \frac{c}{r^q}|B_r|\mathrm{F}(\mu(r))\mu(r)^q, \tag{5.68}$$

for some $c = c(N,p,L,\gamma) > 0$. Combining the above three estimates (5.66), (5.67) and (5.68) for the three integrals in (5.65) respectively, we end up with

$$|I_2| \le \frac{c}{r}(\beta+1)\mathrm{F}(\mu(r))\mu(r)^2|B_r|^{\frac{\gamma-1}{2\gamma}}\Big(\int_{B_r} \eta^{\beta+4}v^{\beta+2}|\mathfrak{X}v|^2\,dx\Big)^{\frac12}\Big(\int_{B_r}\eta^{\gamma\beta}v^{\gamma\beta}\,dx\Big)^{\frac{1}{2\gamma}},$$

from which, together with Young's inequality, the claim (5.62) for I_2 follows.

Finally, for I_3, we have

$$\begin{aligned} I_3 &= -\sum_i \int_\Omega A_i(\mathfrak{X}u)[X_i,X_l](\eta^{\beta+4}v^{\beta+3})\,dx \\ &= -(\beta+4)\sum_i\int_\Omega \eta^{\beta+3}v^{\beta+3}A_i(\mathfrak{X}u)[X_i,X_l]\eta\,dx \\ &\quad -(\beta+3)\sum_i\int_\Omega \eta^{\beta+4}v^{\beta+2}A_i(\mathfrak{X}u)[X_i,X_l]v\,dx = I_3^1 + I_3^2, \end{aligned} \tag{5.69}$$

where we denote the last two integrals in the above equality by I_3^1 and I_3^2, respectively. The estimate for I_3^1 easily follows from the structure condition (1.3) and (5.4) as

$$\begin{aligned} |I_3^1| &\le c(\beta+1)\int_\Omega \eta^{\beta+3}v^{\beta+3}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|\big(|T\eta| + |\mathfrak{X}\eta|\big)\,dx \\ &\le \frac{c}{r^2}(\beta+1)\mathrm{F}(\mu(r))\mu(r)^4\int_{B_r}\eta^\beta v^\beta\,dx, \end{aligned} \tag{5.70}$$

and by Hölder's inequality, I_3^1 satisfies estimate (5.62). To estimate I_3^2, we use the structure condition (1.3),(2.2) and (5.8) to get

$$\begin{aligned} |I_3^2| &\le c(\beta+1)\int_E \eta^{\beta+4}v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|\big(|\mathfrak{X}v| + |Tv|\big)\,dx, \\ &\le c(\beta+1)\int_E \eta^{\beta+4}v^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|\big(|\mathfrak{X}\mathfrak{X}u| + |\mathfrak{X}Tu| + |\mathfrak{X}u| + |Tu|\big)\,dx. \end{aligned} \tag{5.71}$$

As before, it is crucial that the above integral is supported on E of (5.6) since we require more weight of $|\mathfrak{X}u|$ to estimate (5.71), which can be endowed by virtue of (5.7). First, we note that by Hölder's inequality, we have

$$\begin{aligned} |I_3^2| \le c(\beta+1)&\Big(\int_E \eta^{\gamma(\beta+2)}v^{\gamma\beta+4(\gamma-1)}\mathrm{F}(|\mathfrak{X}u|)\,dx\Big)^{\frac12} \\ &\times\Big(\int_E \eta^{(2-\gamma)(\beta+2)+4}v^{(2-\gamma)(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2\big(|\mathfrak{X}\mathfrak{X}u|^2 + |\mathfrak{X}Tu|^2 + |\mathfrak{X}u|^2 + |Tu|^2\big)\,dx\Big)^{\frac12}. \end{aligned}$$

Then, using (5.7), we can obtain

$$|I_3^2| \le c(\beta+1)\mathrm{F}(\mu(r))^{\frac{1}{2}}\mu(r)^{2(\gamma-1)-1}\Big(\int_{B_r}\eta^{\gamma\beta}\mathrm{v}^{\gamma\beta}\,dx\Big)^{\frac{1}{2}}(M'+M+\bar{M}+M_0)^{\frac{1}{2}} \tag{5.72}$$

where, we denote

$$M=\int_\Omega \eta^{(2-\gamma)(\beta+2)+4}\,\mathrm{v}^{(2-\gamma)(\beta+4)}|\mathfrak{X}u|^4\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}Tu|^2\,dx, \tag{5.73}$$

and $M', \bar{M}, M_0$ are the other analogous terms containing $|\mathfrak{X}\mathfrak{X}u|^2, |Tu|^2$ and $|\mathfrak{X}u|^2$ respectively, in place of $|\mathfrak{X}Tu|^2$ in (5.73). Note that $M, \bar{M}$ and M_0 are the same terms as defined in the proof of Lemma 5.2, with $\tau = 2-\gamma$ when $1<\gamma<3/2$. Then, M is estimated by Lemma 5.2 and Corollary 5.4, $M'+\bar{M}+M_0$ is estimated by Hölder's inequality exactly as in (5.56) in the proof of Corollary 5.4 (to estimate the term corresponding to M', Lemma 5.1 and (3.44) are further used) and hence, we have

$$M'+M+\bar{M}+M_0 \le \frac{c}{r^{2\gamma}}\mu(r)^{4(\gamma-1)}\Big(\int_{B_r}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\frac{2}{\gamma-1}}\,dx\Big)^{\gamma-1}\mathcal{I}^{2-\gamma} \tag{5.74}$$

for some $c=c(N,p,L,\gamma)>0$, where $\mathcal{I}$ is as in (5.32), i.e.

$$\begin{aligned}\mathcal{I} =& c(\beta+1)^2\int_\Omega \eta^{\beta+2}\big(\eta^2+|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\,\mathrm{v}^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\,dx\\ &+c(\beta+1)^2\int_\Omega \eta^{\beta+4}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}\mathrm{v}|^2\,dx\\ &+c\int_\Omega \eta^{\beta+4}\mathrm{v}^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^2\,dx.\end{aligned}$$

It is evident that each term of $\mathcal{I}$ are estimable from (5.4). Indeed, the first term of $\mathcal{I}$ is estimated by Hölder's inequality and (5.4) to get

$$\begin{aligned}&\int_\Omega \eta^{\beta+2}\mathrm{v}^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4\big(\eta^2+|\mathfrak{X}\eta|^2+\eta|T\eta|\big)\,dx\\ &\qquad\le \frac{c}{r^2}\mathrm{F}(\mu(r))\mu(r)^8|B_r|^{1-\frac{1}{\gamma}}\Big(\int_{B_r}\eta^{\gamma\beta}\mathrm{v}^{\gamma\beta}\,dx\Big)^{\frac{1}{\gamma}}.\end{aligned} \tag{5.75}$$

For the second term of $\mathcal{I}$, we have

$$\int_\Omega \eta^{\beta+4}\mathrm{v}^{\beta+2}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^4|\mathfrak{X}\mathrm{v}|^2\,dx \le c\mathrm{F}(\mu(r))\mu(r)^4\int_{B_r}\eta^{\beta+4}\mathrm{v}^{\beta+2}|\mathfrak{X}\mathrm{v}|^2\,dx. \tag{5.76}$$

For the third term of $\mathcal{I}$, we use Hölder's inequality, (3.50) and (5.4) to obtain

$$\begin{aligned}&\int_\Omega \eta^{\beta+4}\mathrm{v}^{\beta+4}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^2\,dx\\ &\qquad\le\Big(\int_\Omega \eta^{\frac{2\gamma}{\gamma-1}}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2|Tu|^{\frac{2\gamma}{\gamma-1}}\,dx\Big)^{1-\frac{1}{\gamma}}\\ &\qquad\quad\times\Big(\int_\Omega \eta^{\gamma(\beta+2)}\mathrm{v}^{\gamma(\beta+4)}\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^2\,dx\Big)^{\frac{1}{\gamma}}\\ &\qquad\le\frac{c}{r^2}\mathrm{F}(\mu(r))\mu(r)^8|B_r|^{1-\frac{1}{\gamma}}\Big(\int_{B_r}\eta^{\gamma\beta}\mathrm{v}^{\gamma\beta}\,dx\Big)^{\frac{1}{\gamma}}\end{aligned} \tag{5.77}$$

for some $c=c(N,p,L,\gamma)>0$. Combining (5.75), (5.76) and (5.77), we get

$$\mathcal{I}\le c(\beta+1)^2\mathrm{F}(\mu(r))\mu(r)^4\mathcal{J} \tag{5.78}$$

where we denote

$$\mathcal{J} = \int_{B_r} \eta^{\beta+4} v^{\beta+2} |\mathfrak{X} v|^2 \, dx + \frac{\mu(r)^4}{r^2} |B_r|^{1-\frac{1}{\gamma}} \Big(\int_{B_r} \eta^{\gamma\beta} v^{\gamma\beta} \, dx \Big)^{\frac{1}{\gamma}}. \tag{5.79}$$

Using (5.78) and (5.4) on (5.74), we obtain

$$M' + M + \bar{M} + M_0 \leq \frac{c}{r^{2\gamma}} (\beta+1)^{2(2-\gamma)} \mathrm{F}(\mu(r)) \mu(r)^6 |B_r|^{\gamma-1} \mathcal{J}^{2-\gamma},$$

which is further used in (5.72) to obtain

$$|I_3^2| \leq \frac{c}{r^\gamma} (\beta+1)^{3-\gamma} \mathrm{F}(\mu(r)) \mu(r)^{2\gamma} \, \mathcal{J}^{\frac{2-\gamma}{2}} |B_r|^{\frac{\gamma-1}{2}} \Big(\int_{B_r} \eta^{\gamma\beta} v^{\gamma\beta} \, dx \Big)^{\frac{1}{2}}.$$

Then, by Young's inequality, we end up with

$$|I_3^2| \leq \frac{c_0}{12} (\beta+1) \mathrm{F}(\mu(r)) \mathcal{J} + \frac{c}{r^2} (\beta+1)^{\frac{4}{\gamma}-1} \mathrm{F}(\mu(r)) \mu(r)^4 |B_r|^{1-\frac{1}{\gamma}} \Big(\int_{B_r} \eta^{\gamma\beta} v^{\gamma\beta} \, dx \Big)^{\frac{1}{\gamma}},$$

where $c_0 > 0$ is the same constant as in (5.62) and $\mathcal{J}$ as in (5.79). Thus, I_3^2 satisfies an estimate similar to (5.62) and hence the claim (5.62) for I_3 follows, since both I_3^1 and I_3^2 satisfy similar estimates. This concludes the proof of the claim (5.62), and hence the proof of the lemma. □

The following corollary follows from Lemma 5.5 by using Sobolev's inequality (2.13) on (5.58) and carrying out Moser's iteration. We refer to [39, Corollary 3.5] for the proof. In the Euclidean setting, similar statements have been proved earlier in [10, 45], etc.

Corollary 5.6. *There exists a constant $\theta = \theta(N, p, L) > 0$ such that the following holds for any ball $B_r \subset \Omega$. If we have*

$$|\{x \in B_r : X_l u < \mu(r)/4\}| \leq \theta |B_r| \tag{5.80}$$

for an index $l \in \{1, \dots, m\}$, then

$$\inf_{B_{r/2}} X_l u \geq 3\mu(r)/16.$$

Analogously, if we have

$$|\{x \in B_r : X_l u > -\mu(r)/4\}| \leq \theta |B_r|, \tag{5.81}$$

for an index $l \in \{1, \dots, m\}$, then

$$\sup_{B_{r/2}} X_l u \leq -3\mu(r)/16.$$

5.3. **De Giorgi's method.** It is evident that, equipped with Lemma 5.5 and Corollary 5.6, the problem is reduced to the case for that of a uniformly elliptic equation where the Hölder continuity of the gradient has been shown in the fundamental work of De Giorgi [9]. The techniques have been adopted for degenerated equations for the Euclidean case in [31, 18, 10, 45], etc. and for the case of the Heisenberg Group in [36, 47, 39], etc.

To this end, some notations are in order. For any $B_{\rho_0} \subset \mathbb{R}^N$, the De Giorgi's class $DG^+(B_{\rho_0})$ consists of functions $w \in \mathcal{S}^{1,2}(B_{\rho_0}) \cap L^\infty(B_{\rho_0})$, which satisfy the inequality

$$\int_{B_{\rho'}} |\mathfrak{X}(w-k)^+|^2 \, dx \leq \frac{\gamma}{(\rho-\rho')^2} \int_{B_\rho} |(w-k)^+|^2 \, dx + \chi^2 |A^+_{k,\rho}(w)|^{1-\frac{2}{Q}+\epsilon} \tag{5.82}$$

for some $\gamma, \chi, \epsilon > 0$, where $A^+_{k,\rho}(w) = \{x \in B_\rho : (w-k)^+ = \max(w-k, 0) > 0\}$ for any arbitrary $k \in \mathbb{R}$, the balls $B_{\rho'}, B_\rho$ and B_{ρ_0} are concentric with $0 < \rho' < \rho \leq \rho_0$. Also, $A^-_{k,\rho}(w)$ and the class $DG^-(B_{\rho_0})$ are similarly defined and $DG(B_{\rho_0}) = DG^+(B_{\rho_0}) \cap DG^-(B_{\rho_0})$. All properties of classical De Giorgi class functions, also hold for these classes.

In order to establish integral inequality of the type (5.82) for the gradient, we require the following lemma. The proof is similar to Lemma 4.1 of [39] but more involved.

Lemma 5.7. *Let $B_{r_0} \subset \Omega$ and $0 < r < r_0/2$ be fixed. Suppose that there is $\tau > 0$ such that*

$$|\mathfrak{X}u| \geq \tau\mu(r) \quad \text{in } A^+_{k,r}(X_l u), \tag{5.83}$$

for an index $l \in \{1, \dots, m\}$ and a constant $k \in \mathbb{R}$. Then for any $q \geq 4$ and any $0 < r'' < r' \leq r$, we have the following inequality,

$$\int_{B_{r''}} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u - k)^+|^2\,dx \leq \frac{c}{(r'-r'')^2}\int_{B_{r'}} \mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\,dx + cK|A^+_{k,r'}(X_l u)|^{1-\frac{2}{q}},$$

where $K = r_0^{-2}|B_{r_0}|^{2/q}\mu(r_0)^2\mathrm{F}(\mu(r_0))$ and $c = c(N, p, L, q, \tau) > 0$.

Proof. Note that we can assume $|k| \leq \mu(r_0)$ without loss of generality, since otherwise all the terms vanish. Recalling (3.6) with $Y = X_l$, we have

$$\begin{aligned}\sum_{i,j}\int_\Omega D_j A_i(\mathfrak{X}u)X_j X_l u X_i\varphi\,dx &= \sum_{i,j}\int_\Omega D_j A_i(\mathfrak{X}u)[X_j, X_l]u X_i\varphi\,dx \\ &\quad + \sum_i \int_\Omega A_i(\mathfrak{X}u)[X_i, X_l]\varphi\,dx.\end{aligned} \tag{5.84}$$

Let $\eta \in C_0^\infty(B_{r'})$ is a standard cutoff function such that $\eta = 1$ in $B_{r''}$ and $|\eta| + |\mathfrak{X}\eta| \leq 2/(r'-r'')$, we choose $\varphi = \eta^2(X_l u - k)^+$ as a test function in equation (5.84) to get

$$\begin{aligned}&\sum_{i,j}\int_{B_r} \eta^2 D_j A_i(\mathfrak{X}u)X_j X_l u X_i((X_l u - k)^+)\,dx \\ &\qquad = -2\sum_{i,j}\int_{B_r}\eta(X_l u - k)^+ D_j A_i(\mathfrak{X}u)X_j X_l u X_i\eta\,dx \\ &\qquad\quad + \sum_{i,j}\int_\Omega D_j A_i(\mathfrak{X}u)[X_j, X_l]u X_i(\eta^2(X_l u - k)^+)\,dx \\ &\qquad\quad - \sum_{i,j}\int_{B_r}\eta^2(X_l u - k)^+ D_j A_i(\mathfrak{X}u)[X_i, X_l]X_j u\,dx \\ &\qquad = J_1 + J_2 + J_3,\end{aligned} \tag{5.85}$$

where integral by parts is performed on the last term of (5.84). From (1.3), note that

$$\text{LHS of (5.85)} \geq \int_{B_r}\eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u - k)^+|^2\,dx, \tag{5.86}$$

and we now estimate each J_i of the right hand side of (5.85), one by one.

From (1.3) and Young's inequality, we have

$$\begin{aligned}|J_1| &\leq c\int_{B_r}|\eta||(X_l u - k)^+|\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}\mathfrak{X}u||\mathfrak{X}\eta|\,dx \\ &\leq \varepsilon\int_{B_r}\eta^2\,\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u - k)^+|^2\,dx + \frac{c}{\varepsilon}\int_{B_r}|\mathfrak{X}\eta|^2\mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\,dx.\end{aligned} \tag{5.87}$$

Similarly, from (1.3), (2.2) and Young's inequality, we have

$$\begin{aligned}|J_2| &\leq c\int_\Omega \mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u| + |Tu|\big)\Big(\eta^2|\mathfrak{X}(X_l u - k)^+| + |\eta||\mathfrak{X}\eta||(X_l u - k)^+)|\Big)\,dx \\ &\leq \varepsilon\int_{B_r}\eta^2\mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u - k)^+|^2\,dx + \varepsilon\int_{B_r}|\mathfrak{X}\eta|^2\mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\,dx \\ &\quad + \frac{c}{\varepsilon}\int_{A^+_{k,r}}\eta^2\mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u|^2 + |Tu|^2\big)\,dx.\end{aligned} \tag{5.88}$$

Then the last term of the above is estimated by Hölder's inequality, (3.50) and (5.83) as

$$
\begin{aligned}
&\int_{A^+_{k,r}} \eta^2 \mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u|^2+|Tu|^2\big)\,dx \\
&\qquad \le \Big(\int_{B_{r_0/2}} \mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}u|^q+|Tu|^q\big)\,dx\Big)^{\frac{2}{q}} \Big(\int_{A^+_{k,r}} \mathrm{F}(|\mathfrak{X}u|)\,dx\Big)^{1-\frac{2}{q}} \\
&\qquad \le c\, r_0^{-2}\mu(r_0)^2 \mathrm{F}(\mu(r_0))|B_{r_0}|^{\frac{2}{q}} |A^+_{k,r}(X_l u)|^{1-\frac{2}{q}}
\end{aligned}
\tag{5.89}
$$

for some $c = c(N,p,L,q,\tau) > 0$.

Finally, for J_3, by (1.3),(2.2) and Hölder's inequality,

$$
\begin{aligned}
|J_3| &\le \int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|(X_l u-k)^+|\big(|\mathfrak{X}\mathfrak{X}u|+|\mathfrak{X}Tu|+|\mathfrak{X}u|+|Tu|\big)\,dx \\
&\le \Big(\int_{A^+_{k,r}} \eta^2 |(X_l u-k)^+|^2 \mathrm{F}(|\mathfrak{X}u|)\big(|\mathfrak{X}\mathfrak{X}u|^2+|\mathfrak{X}Tu|^2+|\mathfrak{X}u|^2+|Tu|^2\big)\,dx\Big)^{\frac12} \Big(\int_{A^+_{k,r}} \mathrm{F}(|\mathfrak{X}u|)dx\Big)^{\frac12}
\end{aligned}
$$

The estimate of integrand with the terms $|\mathfrak{X}u|^2+|Tu|^2$ of the above is same as (5.89). However, estimating the first two terms of the above integrand is more delicate; the use of (3.41),(3.61) or any other estimate of Section 3 right away would not be appropriate because the support $A^+_{k,r}$ would be lost that way, thereby producing terms larger than the required right hand side.

To circumvent this, we require the following variants of the Caccioppoli type estimates that is similar to [47, 39]. For any $\kappa \ge 0$, we test (3.8) with $\varphi_k = \eta^2|(X_l u-k)^+|^2|Tu|^\kappa T_k u$. Then, taking summation over $k \in \{1,\dots,n\}$, we obtain

$$
\begin{aligned}
&\sum_{i,j,k}\int_\Omega \eta^2|(X_l u-k)^+|^2|Tu|^\kappa D_j A_i(\mathfrak{X}u)X_j T_k u X_i T_k u\,dx \\
&\quad +\kappa\sum_{i,j,k}\int_\Omega \eta^2|Tu|^{\kappa-1}T_k u|(X_l u-k)^+|^2 D_j A_i(\mathfrak{X}u)X_j T_k u X_i(|Tu|)\,dx \\
&\qquad = -2\sum_{i,j,k}\int_\Omega \eta\,|(X_l u-k)^+|^2|Tu|^\kappa T_k u D_j A_i(\mathfrak{X}u)X_j T_k u X_i\eta\,dx \\
&\qquad\quad -2\sum_{i,j,k}\int_\Omega \eta^2(X_l u-k)^+|Tu|^\kappa T_k u D_j A_i(\mathfrak{X}u)X_j T_k u X_i((X_l u-k)^+)\,dx \\
&\qquad\quad +\sum_{i,j,k}\int_\Omega D_j A_i(\mathfrak{X}u)[X_j,T_k]u X_i\varphi_k\,dx \\
&\qquad\quad +\sum_{i,j,k}\int_\Omega \varphi_k D_j A_i(\mathfrak{X}u)[X_i,T_k]X_j u\,dx \\
&\qquad = U_1+U_2+U_3+U_4
\end{aligned}
\tag{5.90}
$$

From the structure condition (1.3), we have

$$
\text{LHS of (5.90)} \ge \int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|(X_l u-k)^+|^2|Tu|^\kappa|\mathfrak{X}Tu|^2\,dx. \tag{5.91}
$$

The right hand side of (5.90) is estimated as follows. Let us denote

$$
\mathcal{M} = \int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u-k)^+|^2\,dx + \int_{B_r}(\eta^2+|\mathfrak{X}\eta|^2)\mathrm{F}(|\mathfrak{X}u|)|(X_l u-k)^+|^2\,dx. \tag{5.92}
$$

Then, from (1.3) and Hölder's inequality, we have

$$
\begin{aligned}
|U_1|+|U_2| &\le c\int_{B_r}|\eta||(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\kappa+1}|\mathfrak{X}Tu||\mathfrak{X}\eta|\,dx\\
&\quad + c\int_{B_r}\eta^2|(X_lu-k)^+|\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\kappa+1}|\mathfrak{X}Tu||\mathfrak{X}(X_lu-k)^+|\,dx\\
&\le c\,\mathcal{M}^{\frac12}\Big(\int_{B_r}\eta^2|(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{2\kappa+2}|\mathfrak{X}Tu|^2dx\Big)^{\frac12}.
\end{aligned}
\tag{5.93}
$$

To continue the estimates, note that

$$
\begin{aligned}
U_3 &= \sum_{i,j,k}\int_\Omega \eta^2|(X_lu-k)^+|^2|Tu|^\kappa D_jA_i(\mathfrak{X}u)[X_j,T_k]uX_iT_ku\,dx\\
&\quad+\kappa\sum_{i,j,k}\int_\Omega\eta^2|Tu|^{\kappa-1}T_ku|(X_lu-k)^+|^2D_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i(|Tu|)\,dx\\
&\quad+2\sum_{i,j,k}\int_\Omega\eta\,|(X_lu-k)^+|^2|Tu|^\kappa T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i\eta\,dx\\
&\quad+2\sum_{i,j,k}\int_\Omega\eta^2(X_lu-k)^+|Tu|^\kappa T_kuD_jA_i(\mathfrak{X}u)[X_j,T_k]uX_i((X_lu-k)^+)\,dx,
\end{aligned}
\tag{5.94}
$$

which, from (1.3), (2.2) and Hölder's inequality leads to

$$
\begin{aligned}
|U_3| &\le c(\kappa+1)\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|(X_lu-k)^+|^2|Tu|^\kappa\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}Tu|\,dx\\
&\quad+c\int_\Omega|\eta|\mathrm{F}(|\mathfrak{X}u|)|(X_lu-k)^+|^2|Tu|^{\kappa+1}\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}\eta|\,dx\\
&\quad+c\int_\Omega\eta^2|(X_lu-k)^+|\mathrm{F}(|\mathfrak{X}u|)|Tu|^{\kappa+1}\big(|\mathfrak{X}u|+|Tu|\big)|\mathfrak{X}(X_lu-k)^+)|\,dx\\
&\le c(\kappa+1)\mathcal{M}^{\frac12}\Big(\int_{B_r}\eta^2|(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{2\kappa}\big(|\mathfrak{X}u|^2+|Tu|^2\big)|\mathfrak{X}Tu|^2dx\Big)^{\frac12}\\
&\quad+c\,\mathcal{M}^{\frac12}\Big(\int_{B_r}\eta^2|(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{2\kappa+2}\big(|\mathfrak{X}u|^2+|Tu|^2\big)dx\Big)^{\frac12}
\end{aligned}
\tag{5.95}
$$

Finally, from (1.3), (2.4) and Hölder's inequality, we have

$$
\begin{aligned}
|U_4| &\le c\int_\Omega\eta^2\mathrm{F}(|\mathfrak{X}u|)|(X_lu-k)^+|^2|Tu|^{\kappa+1}\big(|\mathfrak{X}\mathfrak{X}u|+|\mathfrak{X}Tu|+|\mathfrak{X}u|+|Tu|\big)\,dx\\
&\le c\,\mathcal{M}^{\frac12}\Big(\int_{B_r}\eta^2|(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)|Tu|^{2\kappa+2}\big(|\mathfrak{X}\mathfrak{X}u|^2+|\mathfrak{X}Tu|^2+|\mathfrak{X}u|^2+|Tu|^2\big)dx\Big)^{\frac12},
\end{aligned}
$$

which combined with the previous estimates (5.90), (5.91), (5.93) and (5.95), leads to

$$
\int_{B_r}\eta^2\mathrm{F}(|\mathfrak{X}u|)|(X_lu-k)^+|^2|Tu|^\kappa|\mathfrak{X}Tu|^2\,dx\le c(\kappa+1)\mathcal{M}^{\frac12}a(2\kappa+2)^{\frac12}
\tag{5.96}
$$

where $c=c(N,p,L)>0$, $\mathcal{M}$ is as in (5.92) and we denote

$$
a(\kappa)=\int_{B_r}\eta^2|(X_lu-k)^+|^2\mathrm{F}(|\mathfrak{X}u|)\big(|Tu|^\kappa+|\mathfrak{X}u|^\kappa\big)\big(|\mathfrak{X}\mathfrak{X}u|^2+|\mathfrak{X}Tu|^2+|\mathfrak{X}u|^2+|Tu|^2\big)dx
\tag{5.97}
$$

for any $\kappa\ge0$. Notice that the left hand side of (5.96) contains a part of $a(\kappa)$. To get the other parts one requires similar estimates. From the symmetric structure of commutation relations

(2.2), it is not difficult to see that a similar estimate can be obtained as above by testing (3.8) with $\varphi_k = \eta^2|(X_l u - k)^+|^2|\mathfrak{X}u|^\kappa T_k u$, i.e. we can obtain

$$\int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2|\mathfrak{X}u|^\kappa|\mathfrak{X}Tu|^2\,dx \le c(\kappa+1)\mathcal{M}^{\frac{1}{2}} a(2\kappa+2)^{\frac{1}{2}}. \tag{5.98}$$

Similarly, it is also not hard to see that by taking $\varphi_l = \eta^2|(X_l u - k)^+|^2\big(|\mathfrak{X}u|^\kappa + |Tu|^\kappa\big)X_l u$ as the test function in (3.7) and following the steps similarly as above, we can obtain

$$\int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\big(|\mathfrak{X}u|^\kappa + |Tu|^\kappa\big)|\mathfrak{X}\mathfrak{X}u|^2\,dx \le c(\kappa+1)\mathcal{M}^{\frac{1}{2}} a(2\kappa+2)^{\frac{1}{2}}. \tag{5.99}$$

Finally, the following is easy to see from Young's and Hölder's inequality that

$$\int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\big(|\mathfrak{X}u|^\kappa + |Tu|^\kappa\big)\big(|\mathfrak{X}u|^2 + |Tu|^2\big)\,dx \le c\mathcal{M}^{\frac{1}{2}} a(2\kappa+2)^{\frac{1}{2}}. \tag{5.100}$$

Thus, by adding (5.96),(5.98),(5.99) and (5.100), we finally get

$$a(\kappa) \le c(\kappa+1)\mathcal{M}^{\frac{1}{2}} a(2\kappa+2)^{\frac{1}{2}}, \tag{5.101}$$

for any $\kappa \ge 0$, where $\mathcal{M}$ is as in (5.92). We iterate (5.101) with the sequence $\kappa_j = 2^j - 2$ for $j \in \mathbb{N}$ and letting $c = c(N,p,L,j)$ and $a_j = a(\kappa_j)$, so that we get

$$a_1 \le (c\,\mathcal{M})^{\frac{1}{2}} a_2^{\frac{1}{2}} \le \ldots \le (c\,\mathcal{M})^{(1-1/2^j)}\, a_{j+1}^{1/2^j} \tag{5.102}$$

for every $j \in \mathbb{N}$. Now, for a large enough j to be chosen later, we estimate using all the apriori estimates obtained finally in Section 3, i.e. (3.44), (3.50), (3.51), (3.52), to get

$$a_{j+1} \le \frac{c\mu(r_0)^2}{r_0^{\kappa_{j+1}+4}} \int_{B_{r_0}} \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}u|^{\kappa_{j+1}+2}\,dx \le c\,r_0^{-(\kappa_{j+1}+4)}\mathrm{F}(\mu(r_0))\mu(r_0)^{\kappa_{j+1}+4}\,|B_{r_0}|. \tag{5.103}$$

Now we are ready to estimate J_3. Notice that, from (5.83) we have

$$|J_3| \le c\,a_1^{1/2}\Big(\int_{A^+_{k,r}} \mathrm{F}(|\mathfrak{X}u|)dx\Big)^{\frac{1}{2}} \le c\,a_1^{1/2}\mathrm{F}(\mu(r_0))^{1/2}|A^+_{k,r}(X_l u)|^{1/2}$$

for some $c = c(N,p,L,\tau) > 0$. Then, applying the iteration (5.102) and the estimate (5.103),

$$\begin{aligned}|J_3| &\le c\,\mathcal{M}^{\frac{1}{2}(1-1/2^j)}\, a_{j+1}^{1/2^{j+1}}\mathrm{F}(\mu(r_0))^{1/2}|A^+_{k,r}(X_l u)|^{\frac{1}{2}}\\ &\le \frac{c}{r_0^{(1+1/2^j)}}\mathcal{M}^{\frac{1}{2}(1-1/2^j)}\mathrm{F}(\mu(r_0))^{\frac{1}{2}(1+1/2^j)}\mu(r_0)^{(1+1/2^j)}|B_{r_0}|^{1/2^{j+1}}|A^+_{k,r}(X_l u)|^{\frac{1}{2}}.\end{aligned}$$

Then, by Young's inequality, we finally obtain

$$|J_3| \le \mathcal{M}/2 + c\,r_0^{-2}\mathrm{F}(\mu(r_0))\mu(r_0)^2|B_{r_0}|^{1/(2^j+1)}|A^+_{k,r}(X_l u)|^{2^j/(2^j+1)} \tag{5.104}$$

for some $c = c(N,p,L,\tau,j) > 0$. This, with $j = j(q) \in \mathbb{N}$ such that $2^j/(2^j+1) \ge 1 - 2/q$, gives us the required estimate. Combining (5.87) and (5.88),(5.89) with a small enough $\varepsilon > 0$ and (5.104) with the choice of $j = j(q)$, we finally end up with

$$\begin{aligned}\int_{B_r} \eta^2 \mathrm{F}(|\mathfrak{X}u|)|\mathfrak{X}(X_l u - k)^+|^2\,dx \le c\int_{B_r} (\eta^2 + |\mathfrak{X}\eta|^2)\mathrm{F}(|\mathfrak{X}u|)|(X_l u - k)^+|^2\,dx\\ + c\,r_0^{-2}\mu(r_0)^2\mathrm{F}(\mu(r_0))|B_{r_0}|^{\frac{2}{q}}|A^+_{k,r}(X_l u)|^{1-\frac{2}{q}}\end{aligned}$$

for some $c = c(N,p,L,q,\tau) > 0$ and the proof is finished. □

Remark 5.8. Some observations on Lemma 5.7 are in order:

(1) We can obtain a similar inequality, corresponding to that of Lemma 5.7 with $(X_l u - k)^+$ replaced by $(X_l u - k)^-$ and $A^+_{k,r}(X_l u)$ replaced by $A^-_{k,r}(X_l u)$.

(2) It can be observed from the proof that Lemma 5.7 holds without the condition (5.83), for the case of $p \geq 2$ and for this range, it is possible to prove the $C^{1,\alpha}$ regularity directly following the direction of the classical result of DiBenedetto [10], without using anything related to the truncation from the previous subsections. However, for the full range $1 < p < \infty$, the application of Lemma 5.7 shall be made in the following whenever Corollary 5.6 ensures the condition (5.83).

Le us fix $B_{r_0} \subset \Omega$. For any $0 < r < r_0$, similarly as (5.3), let us denote

$$\omega_l(r) = \operatorname{osc}_{B_r} X_l u, \quad \omega(r) = \max_{1 \leq l \leq m} \omega_l(r),$$

and it is clear that we have $\omega(r) \leq 2\mu(r)$. The following oscillation Lemma is a consequence of Lemma 5.7 and Corollary 5.6. Its proof is the same as [39, Theorem 4.3] and quite similar to those in [10, 45], etc. Nevertheless, we provide a brief outline for sake of completeness.

Lemma 5.9. *There exists a constant $s = s(N, p, L) \geq 0$ such that for every $0 < r \leq r_0/16$, we have the following,*

$$\omega(r) \leq (1 - 2^{-s})\omega(8r) + 2^s\big(\delta + \mu(r_0)^2\big)^{\frac{1}{2}} \left(\frac{r}{r_0}\right)^{\alpha}, \tag{5.105}$$

where $\alpha = 1/2$ when $1 < p < 2$ and $\alpha = 1/p$ when $p \geq 2$.

Proof. Letting α be as given, note that we may assume

$$\omega(r) \geq \big(\delta + \mu(r_0)^2\big)^{\frac{1}{2}} \left(\frac{r}{r_0}\right)^{\alpha}, \tag{5.106}$$

since, otherwise, (5.105) is true with $s = 0$. In the following, we assume that (5.106) is true we divide the proof into two alternative cases, similarly as in [10, 45, 39], etc.

Case 1. For at least one index $l \in \{1, \dots, m\}$, we have either

$$|\{x \in B_{4r} : X_l u < \mu(4r)/4\}| \leq \theta |B_{4r}| \quad \text{or} \quad |\{x \in B_{4r} : X_l u > -\mu(4r)/4\}| \leq \theta |B_{4r}|, \tag{5.107}$$

where $\theta = \theta(N, p, L) > 0$ is the constant in Corollary 5.6. Then, from Corollary 5.6, we have

$$|\mathfrak{X} u| \geq 3\mu(2r)/16 \quad \text{in} \quad B_{2r}, \tag{5.108}$$

which allows us to use Lemma 5.7 with $q = 2Q$ to obtain

$$\begin{aligned} \int_{B_{r''}} |\mathfrak{X}(X_i u - k)^+|^2 \, dx \leq{}& \frac{c}{(r' - r'')^2} \int_{B_{r'}} |(X_i u - k)^+|^2 \, dx \\ &+ cK\mathrm{F}(\mu(2r))^{-1} |A^+_{k,r'}(X_i u)|^{1 - \frac{1}{Q}} \end{aligned} \tag{5.109}$$

for any $0 < r'' < r' \leq 2r$, $i \in \{1, \dots, 2n\}$ and all $k \in \mathbb{R}$, where $K = r_0^{-2} |B_{r_0}|^{1/Q} \mu(r_0)^2 \mathrm{F}(\mu(r_0))$. This implies that $X_i u$ belongs to the De Giorgi class $DG^+(B_{2r})$ for each $i \in \{1, \dots, m\}$ and hence, it satisfies an oscillation lemma, see [9, 31, 27], etc. from which, as in [39], it is not hard to obtain (5.105).

Case 2. If Case 1 does not happen, then for every $i \in \{1, \dots, m\}$, we have

$$|\{x \in B_{4r} : X_i u < \mu(4r)/4\}| > \theta |B_{4r}| \quad \text{and} \quad |\{x \in B_{4r} : X_i u > -\mu(4r)/4\}| > \theta |B_{4r}|. \tag{5.110}$$

Then, note that, on the set $\{x \in B_{8r} : X_i u > \mu(8r)/4\}$, we trivially have

$$|\mathfrak{X} u| \geq \mu(8r)/4 \quad \text{in } A^+_{k,8r}(X_i u) \tag{5.111}$$

for all $k \geq \mu(8r)/4$. Thus, we can apply Lemma 5.7 with $q = 2Q$ to conclude that

$$\begin{aligned} \int_{B_{r''}} |\mathfrak{X}(X_i u - k)^+|^2 \, dx \leq{}& \frac{c}{(r' - r'')^2} \int_{B_{r'}} |(X_i u - k)^+|^2 \, dx \\ &+ cK\,\mathrm{F}(\mu(8r))^{-1} |A^+_{k,r'}(X_i u)|^{1 - \frac{1}{Q}} \end{aligned} \tag{5.112}$$

where $K = r_0^{-2}|B_{r_0}|^{1/Q}\mu(r_0)^2\mathrm{F}(\mu(r_0))$, whenever $k \geq k_0 = \mu(8r)/4$ and $0 < r'' < r' \leq 8r$. Then, we have standard estimates for the supremum and infimum of $X_i u$, see [27, 31], etc. such that for some $s_1 = s_1(n, p, L) > 0$ he following holds,

$$\sup_{B_{2r}} X_i u \leq \sup_{B_{8r}} X_i u - 2^{-s_1}\big(\sup_{B_{8r}} X_i u - \mu(8r)/4\big) + cK^{\frac{1}{2}}\mathrm{F}(\mu(8r))^{-1/2}r^{\frac{1}{2}}. \tag{5.113}$$

and from the second part of Case 2 and Remark 5.8, we similarly have

$$\inf_{B_{2r}} X_i u \geq \inf_{B_{8r}} X_i u + 2^{-s_1}\big(-\inf_{B_{8r}} X_i u - \mu(8r)/4\big) - cK^{\frac{1}{2}}\mathrm{F}(\mu(8r))^{-1/2}r^{\frac{1}{2}}. \tag{5.114}$$

Combining (5.113) and (5.114) as in [39], one can also get (5.105) and the proof is finished. □

From (5.105) of Lemma 5.9, it is easy to prove Theorem 1.3 for $\delta > 0$ by a standard iteration argument which is classical. We refer to [27, 31, 26], etc. for details. Since all the constants in the above estimates are independent of δ, hence from the arguments of Remark 2.5, we can obtain (1.7) also for $\delta = 0$. Thus, the proof of Theorem 1.3 is complete.

5.4. **Concluding Remarks.** Here are some further remarks about the method and techniques used above throughout the paper and some discussion on further generalizations.

(1) All the apriori estimates for the equation (3.1) in Section 3, rely only on the structure conditions (3.2) and are independent of the "p" of the p-Laplacian. The other arguments in the later sections can also be appropriately generalized for the regularity theory of more general classes of equations; e.g. the equation $\mathrm{div}_{_H}\big(f(|\mathfrak{X}u|)\mathfrak{X}u\big) = 0$ with f as any doubling differentiable function on $(0,\infty)$ so that $t \mapsto tf(t)$ is monotonic, admits weak solutions in sub-elliptic Orlicz-Sobolev spaces (defined similarly as Definition (2.3)) which can be similarly shown to be locally $C^{1,\alpha}$ whenever $1 + |\mathfrak{X}u|f'(|\mathfrak{X}u|)/f(|\mathfrak{X}u|)$ is positive and finite. We refer to [38, 37] for the case of the Heisenberg group.

(2) In the special case when the vector fields are left invariant with respect to a Carnot group of step 2 (in particular, the Heisenberg group as in [47, 39]), the commutation relations are more special than (2.2) as mentioned earlier in Remark 2.1, up to the choice of exponential coordinates. This is why, in that case the estimates are simpler and some of the items in the integral estimates gets suppressed, see Remark 5.3. But, unlike the previous case for the Heisenberg group by Mukherjee-Zhong [39], here the lengths of the main crucial lemmas like Lemma 5.2 and Lemma 5.7 are drastically extended and Corollary 3.11, Corollary 5.4, etc. have to be added. Also, in [47], the approximation has been accomplished by a Hilbert-Haar theory for the case of the Heisenberg group, which depends on the group law and hence does not have an easy analogue for the case of general Hörmander vector fields (some deeper theorems and newer adaptations are needed for it). Therefore, in this paper, we have used the Riemannian approximation which is more natural and easier to use.

(3) The generalization of the above techniques from step 2 to any higher step does not seem to be immediately available. The inequalities (2.4) and (2.5) in their present form, do not lead to anything useful for the case of step 3 or any higher step. The higher order commutators have more intricate strucures. For instance, letting $X_1^{(s)}, \ldots, X_k^{(s)}$ as vector fields corresponding to s-th order commutators, the nature of the intermediate integral estimates for $\mathfrak{X}^{(s)}u$ is unclear. In case of a Carnot group of step r, i.e. $1 \leq s \leq r$, notice that $|\mathfrak{X}\mathfrak{X}^{(s)}u| \neq |\mathfrak{X}^{(s)}\mathfrak{X}u|$ except for the cases of $s = r$ (as $X_i X_j^{(r)} = X_j^{(r)} X_i$ for the last step of the Lie algebra) and $s = 1$ (from the definition of the norm even though $X_i X_j \neq X_j X_i$). This feature leads to serious difficulties in the integral estimates for the intermediary steps that do not occur in a set up where these two are the only possible

alternatives, i.e. the step 2 case with $s = 1, 2$. Similar features typical to vector fields of step 2 Carnot groups are also imbibed in more general step 2 vector fields. Lastly, we refer to [17] for some results on certain special cases of vector fields in higher steps.

Acknowledgement

The authors have been supported by European Union's Horizon 2020 research and innovation programme under the (GHAIA) Marie Skłodowska-Curie grant agreement No. 777822. The authors are thankful to the anonymous referee and to Juan Manfredi for a careful reading of the paper and providing valuable suggestions.

References

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

[2] Luca Capogna. Regularity of quasi-linear equations in the Heisenberg group. *Comm. Pure Appl. Math.*, 50(9):867–889, 1997.

[3] Luca Capogna and Giovanna Citti. Regularity for subelliptic PDE through uniform estimates in multi-scale geometries. *Bull. Math. Sci.*, 6(2):173–230, 2016.

[4] Luca Capogna, Giovanna Citti, Enrico Le Donne, and Alessandro Ottazzi. Conformality and Q-harmonicity in sub-Riemannian manifolds. *J. Math. Pures Appl. (9)*, 122:67–124, 2019.

[5] Luca Capogna, Giovanna Citti, and Garrett Rea. A subelliptic analogue of Aronson-Serrin's Harnack inequality. *Math. Ann.*, 357(3):1175–1198, 2013.

[6] Luca Capogna, Donatella Danielli, and Nicola Garofalo. An embedding theorem and the Harnack inequality for nonlinear subelliptic equations. *Comm. Partial Differential Equations*, 18(9-10):1765–1794, 1993.

[7] Luca Capogna and Nicola Garofalo. Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type. *J. Eur. Math. Soc. (JEMS)*, 5(1):1–40, 2003.

[8] Wei-Liang Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.

[9] Ennio De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.

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

[11] E. DiBenedetto and Neil S. Trudinger. Harnack inequalities for quasiminima of variational integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1(4):295–308, 1984.

[12] András Domokos. Differentiability of solutions for the non-degenerate p-Laplacian in the Heisenberg group. *J. Differential Equations*, 204(2):439–470, 2004.

[13] András Domokos and Juan J. Manfredi. $C^{1,\alpha}$-regularity for p-harmonic functions in the Heisenberg group for p near 2. 370:17–23, 2005.

[14] András Domokos and Juan J. Manfredi. Subelliptic Cordes estimates. *Proc. Amer. Math. Soc.*, 133(4):1047–1056 (electronic), 2005.

[15] András Domokos and Juan J. Manfredi. Nonlinear subelliptic equations. *Manuscripta Math.*, 130(2):251–271, 2009.

[16] András Domokos and Juan J. Manfredi. $C^{1,\alpha}$-subelliptic regularity on SU(3) and compact, semi-simple Lie groups. *Anal. Math. Phys.*, 10(1):Paper No. 4, 32, 2020.

[17] András Domokos and Juan J. Manfredi. Hilbert-Haar coordinates and Miranda's theorem in Lie groups. In *Something about nonlinear problems*, volume 11 of *Bruno Pini Math. Anal. Semin.*, pages 94–118. Univ. Bologna, Alma Mater Stud., Bologna, 2020.

[18] Lawrence C. Evans. A new proof of local $C^{1,\alpha}$ regularity for solutions of certain degenerate elliptic p.d.e. *J. Differential Equations*, 45(3):356–373, 1982.

[19] Anna Föglein. Partial regularity results for subelliptic systems in the Heisenberg group. *Calc. Var. Partial Differential Equations*, 32(1):25–51, 2008.

[20] G. B. Folland. Subelliptic estimates and function spaces on nilpotent Lie groups. *Ark. Mat.*, 13(2):161–207, 1975.

[21] G. B. Folland and E. M. Stein. Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group. *Comm. Pure Appl. Math.*, 27:429–522, 1974.

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

[23] Bruno Franchi, Guozhen Lu, and Richard L. Wheeden. A relationship between Poincaré-type inequalities and representation formulas in spaces of homogeneous type. *Internat. Math. Res. Notices*, (1):1–14, 1996.
[24] Nicola Garofalo and Duy-Minh Nhieu. Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces. *Comm. Pure Appl. Math.*, 49(10):1081–1144, 1996.
[25] Zhong Ge. Collapsing Riemannian metrics to Carnot-Carathéodory metrics and Laplacians to sub-Laplacians. *Canad. J. Math.*, 45(3):537–553, 1993.
[26] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
[27] Enrico Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
[28] Piotr Hajłasz and Pekka Koskela. Sobolev met Poincaré. *Mem. Amer. Math. Soc.*, 145(688):x+101, 2000.
[29] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
[30] David Jerison. The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J.*, 53(2):503–523, 1986.
[31] Olga A. Ladyzhenskaya and Nina N. Ural′tseva. *Linear and quasilinear elliptic equations*. Translated from the Russian by Scripta Technica, Inc. Translation editor: Leon Ehrenpreis. Academic Press, New York-London, 1968.
[32] John L. Lewis. Regularity of the derivatives of solutions to certain degenerate elliptic equations. *Indiana Univ. Math. J.*, 32(6):849–858, 1983.
[33] Gary M. Lieberman. Boundary regularity for solutions of degenerate elliptic equations. *Nonlinear Anal.*, 12(11):1203–1219, 1988.
[34] Juan J. Manfredi and Giuseppe Mingione. Regularity results for quasilinear elliptic equations in the Heisenberg group. *Math. Ann.*, 339(3):485–544, 2007.
[35] Silvana Marchi. $C^{1,\alpha}$ local regularity for the solutions of the p-Laplacian on the Heisenberg group. The case $1 + \frac{1}{\sqrt{5}} < p \leq 2$. *Comment. Math. Univ. Carolin.*, 44(1):33–56, 2003.
[36] Giuseppe Mingione, Anna Zatorska-Goldstein, and Xiao Zhong. Gradient regularity for elliptic equations in the Heisenberg group. *Adv. Math.*, 222(1):62–129, 2009.
[37] Shirsho Mukherjee. $C^{1,\alpha}$-Regularity of Quasilinear elliptic equations on the Heisenberg Group. *https://arxiv.org/abs/1805.03748*, 2018.
[38] Shirsho Mukherjee. On local Lipschitz regularity for quasilinear equations in the Heisenberg group. *Nonlinear Anal.*, 212:Paper No. 112453, 24, 2021.
[39] Shirsho Mukherjee and Xiao Zhong. $C^{1,\alpha}$-regularity for variational problems in the Heisenberg group. *Anal. PDE*, 14(2):567–594, 2021.
[40] Alexander Nagel and E. M. Stein. *Lectures on pseudodifferential operators: regularity theorems and applications to nonelliptic problems*, volume 24 of *Mathematical Notes*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1979.
[41] Alexander Nagel, Elias M. Stein, and Stephen Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985.
[42] Diego Ricciotti. On the $C^{1,\alpha}$ regularity of p-harmonic functions in the Heisenberg group. *Proc. Amer. Math. Soc.*, 146(7):2937–2952, 2018.
[43] Linda Preiss Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.
[44] Antonio Sánchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.*, 78(1):143–160, 1984.
[45] Peter Tolksdorf. Regularity for a more general class of quasilinear elliptic equations. *J. Differential Equations*, 51(1):126–150, 1984.
[46] Chao Jiang Xu. Regularity for quasilinear second-order subelliptic equations. *Comm. Pure Appl. Math.*, 45(1):77–96, 1992.
[47] Xiao Zhong. Regularity for variational problems in the Heisenberg Group. *https://arxiv.org/abs/1711.03284*.

(G. Citti) Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 - Bologna (BO), Italy.
Email address: giovanna.citti@unibo.it

(S. Mukherjee) Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 - Bologna (BO), Italy.
Email address: shirsho.mukherjee2@unibo.it