Research article

Jialin Wang*, Maochun Zhu, Shujin Gao, and Dongni Liao

Regularity for sub-elliptic systems with VMO-coefficients in the Heisenberg group: the sub-quadratic structure case

https://doi.org/10.1515/anona-2020-0145
Received March 19, 2020; accepted June 29, 2020.

Abstract: We consider nonlinear sub-elliptic systems with VMO-coefficients for the case $1 < p < 2$ under controllable growth conditions, as well as natural growth conditions, respectively, in the Heisenberg group. On the basis of a generalization of the technique of $\mathcal{A}$-harmonic approximation introduced by Duzaar-Grotowski-Kronz, and an appropriate Sobolev-Poincaré type inequality established in the Heisenberg group, we prove partial Hölder continuity results for vector-valued solutions of discontinuous sub-elliptic problems. The primary model covered by our analysis is the non-degenerate sub-elliptic p-Laplacian system with VMO-coefficients, involving sub-quadratic growth terms.

Keywords: Partial Hölder continuity; Heisenberg group; sub-quadratic controllable growth; sub-quadratic natural growth; VMO-coefficient; p-Laplacian

MSC: 35H20, 35B65, 32A37

1 Introduction and statements of main results

In this paper, we consider discontinuous sub-elliptic systems with sub-quadratic growth coefficients that belong to the space of functions with vanishing mean oscillation (VMO, for short) in the Heisenberg group $\mathbb{H}^n$. We establish optimal partial Hölder continuity for vector-valued weak solutions in the sense that the solution is Hölder continuous on an open subset of its domain with full measure. More precisely, let Ω be a bounded domain, and horizontal gradient $X = \{X_1, \cdots X_{2n}\}$ with the horizontal vector fields X_i $(i = 1, \cdots, 2n)$ in $\mathbb{H}^n$, we consider sub-elliptic systems of the type

$$-\sum_{i=1}^{2n} X_i A_i^\alpha(\xi, u, Xu) = B^\alpha(\xi, u, Xu), \quad \text{in } \Omega, \quad \alpha = 1, 2, \cdots, N, \tag{1.1}$$

where the primary coefficient $A_i^\alpha \in VMO$ and satisfies some standard ellipticity and growth conditions with polynomial growth rate $p \in (1, 2)$, and the inhomogeneous term B^α conforms to either controllable growth conditions, or natural growth conditions under an additional smallness assumption on the weak solutions. For the precise statement of the assumptions, and more details about the Heisenberg group, we refer to **(H1)**-**(H4)**-**(HC)** and **(HN)** below, and Section 2, respectively.

***Corresponding Author: Jialin Wang,** School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, Jiangxi, P.R. China, E-mail: wangjialin@gnnu.edu.cn. Tel.:+86 0797 8393663; fax: +86 0797 8395933
Shujin Gao, Dongni Liao, School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, Jiangxi, P.R. China
Maochun Zhu, School of Science, Jiangsu University, Zhenjiang 212013, Jiangsu, P.R. China

The main new aspect of this paper is the fact that we are able to deal with the inhomogeneity $B^\alpha : \mathbb{R}^{2n+1} \times \mathbb{R}^N \times \mathbb{R}^{2n\times N} \to \mathbb{R}^N$ that satisfies the sub-quadratic controllable growth conditions, as well as sub-quadratic natural growth conditions, respectively, and the primary coefficient $A_i^\alpha : \mathbb{R}^{2n+1} \times \mathbb{R}^N \times \mathbb{R}^{2n\times N} \to \mathbb{R}^{2n\times N}$ that satisfies only a VMO-condition in ξ and is continuous in u. More precisely, we assume that the partial mapping $\xi \mapsto A_i^\alpha(\xi, u, P)/(1+|P|)^{p-1}$ is VMO uniformly in (u, P), in the sense of (1.5) below and, moreover, $u \mapsto A_i^\alpha(\xi, u, P)/(1+|P|)^{p-1}$ is continuous in the sense of (1.3) below. Our tool of choice is the use of an appropriate Sobolev-Poincaré inequality, and the harmonic approximation lemma; see Lemma 3.1, Lemma 3.3 below, respectively. The method of proof employed here will avoid the use of $L^q - L^p$-estimates for the horizontal gradient and reverse Hölder inequalities. Our results essentially extend those results that the coefficients are continuous with respect to variables ξ and u to the case of the coefficients being VMO in the first variable ξ. We point out that partial Hölder continuity is the best one can expect under such weak assumptions concerning regularity of the structural functions A_i^α and B^α in the (ξ, u)-variables.

We now impose the precise structure assumptions for coefficients A_i^α and B^α we are dealing with.

(H1). The primary coefficient A_i^α satisfies following ellipticity and growth conditions for a growth exponent $1 < p < 2$:

$$\begin{cases} \left\langle D_P A_i^\alpha(\xi, u, P)P_0, P_0\right\rangle \geq \nu(1+|P|)^{p-2}|P_0|^2, \\ |A_i^\alpha(\xi, u, P)| + (1+|P|)|D_P A_i^\alpha(\xi, u, P)| \leq L(1+|P|)^{p-1}, \end{cases} \tag{1.2}$$

for any choice of $\xi \in \Omega$, $u, u_0 \in \mathbb{R}^N$ and $P, P_0 \in \mathbb{R}^{2n\times N}$. Here structure constants $\nu \leq 1 \leq L < \infty$.

(H2). The vector field A_i^α is continuous with respect to the second variable u. More precisely, there exists a bounded, concave and non-decreasing moduli of continuity $\omega : [0,\infty) \to [0,1]$ with $\lim_{s\to 0}\omega(s) = 0 = \omega(0)$ such that

$$|A_i^\alpha(\xi, u, P) - A_i^\alpha(\xi, u_0, P)| \leq L\omega\left(|u-u_0|^p\right)(1+|P|)^{p-1}, \quad 1 < p < 2. \tag{1.3}$$

(H3). The vector field A_i^α is differentiable in the third variable P with continuous derivatives. This infers the bounded, concave and non-decreasing modulus $\mu : [0,\infty) \to [0,1]$ such that $\mu(t) \leq t$, $\lim_{s\to 0}\mu(s) = 0 = \mu(0)$, and we have

$$|D_P A_i^\alpha(\xi, u, P) - D_P A_i^\alpha(\xi, u, P_0)| \leq L\mu\left(\frac{|P-P_0|}{1+|P|+|P_0|}\right)(1+|P|+|P_0|)^{p-2}, \quad 1 < p < 2. \tag{1.4}$$

With respect to the dependence on the first variable ξ, we do not impose a continuity condition, but we merely assume the following VMO-condition.

(H4). The mapping $\xi \mapsto A_i^\alpha(\xi, u, P)/(1+|P|)^{p-1}$ satisfies the following VMO-condition uniformly in u and P:

$$\left|A_i^\alpha(\xi, u, P) - \left(A_i^\alpha(\cdot, u, P)\right)_{\xi_0, r}\right| \leq \mathbf{v}_{\xi_0}(\xi, r)(1+|P|)^{p-1}, \text{for all } \xi \in B_r(\xi_0), \quad 1 < p < 2. \tag{1.5}$$

where $\mathbf{v}_{\xi_0} : \mathbb{R}^{2n+1} \times [0, \rho_0] \to [0, 2L]$ are bounded functions satisfying

$$\lim_{\rho\to 0}\mathbf{V}(\rho) = 0, \text{ where } \mathbf{V}(\rho) = \sup_{\xi_0\in\Omega}\sup_{0<r\leq\rho_0}\fint_{B_r(\xi_0)\cap\Omega}\mathbf{v}_{\xi_0}(\xi, r)d\xi. \tag{1.6}$$

Here we have used the short-hand notation

$$\left(A_i^\alpha(\cdot, u, P)\right)_{\xi_0, r} := \fint_{B_r(\xi_0)\cap\Omega} A_i^\alpha(\zeta, u, P)d\zeta = \left|B_r(\xi_0)\cap\Omega\right|^{-1}\int_{B_r(\xi_0)\cap\Omega} A_i^\alpha(\zeta, u, P)d\zeta.$$

(HC) (Controllable growth condition). The inhomogeneity B^α satisfies sub-quadratic controllable growth condition

$$|B^\alpha(\xi, u, P)| \leq C\left(1 + |u|^{p^\star - 1} + |P|^{p(1-\frac{1}{p^\star})}\right), \quad p^\star = \begin{cases} \dfrac{pQ}{Q-p}, & 1 < p < Q, \\ \text{any constant } p^\star \geq p, & p \geq Q, \end{cases} \tag{1.7}$$

where C is a positive constant. We note that $Q \geq 3$ is the homogeneous dimension in non-Abelian Heisenberg groups (see (2.1) below), and the exponent $p \in (1,2)$. So those infer that $p < Q$, and then, $p^\star = \frac{pQ}{Q-p}$ in our

setting.

(HN) (Natural growth condition). For $|u| \le M = \sup_{\Omega} |u|$. The term B^α satisfies sub-quadratic natural growth condition

$$|B^\alpha(\xi, u, P)| \le a|P|^p + b, \quad 1 < p < 2, \tag{1.8}$$

where $a = a(M)$ and $b = b(M)$ are constants possibly depending on M.

Now we mention some results on elliptic systems. Duzaar and Grotowski [9] prove optimal partial Hölder continuity for nonlinear elliptic systems with quadratic growth $p = 2$, by a new method so-called $\mathcal{A}$-harmonic approximation introduced by Duzaar and Steffen [15]. Then, the method was extended to non-quadratic growth cases. Duzaar and Mingione [12, 13] consider systems of p-Laplacian type. Many partial regularity results have been established for more general nonlinear elliptic problems with Hölder, or Dini continuous coefficients; see, for example, [6, 8, 11, 28]. Furthermore, with respect to discontinuous elliptic problems, we refer to Bögelein, Duzaar, Habermann and Scheven [1], Ragusa [23], Zheng [35], Kanazawa [26], Goodrich, Ragusa and Scapellato [20], Polidoro and Ragusa [22], Scapellato [24], and Tan, Wang and Chen [27] and the references therein.

Several regularity results were focused on sub-elliptic systems in Heisenberg groups, or Hörmander vector fields; see Bramanti [2]. Xu and Zuily [34], Capogna and Garofalo [5], and Shores [25] showed partial regularity for quasi-linear sub-elliptic systems with quadratic growth $p = 2$. Their methods depend on generalization of classical freezing coefficient method. Then, by the generalization of the method of $\mathcal{A}$-harmonic approximation, Föglein [16] treated homogeneous nonlinear sub-elliptic systems with Hölder continuous coefficients, under super-quadratic growth conditions $p \ge 2$ in the Heisenberg group, and established partial Hölder continuity for the horizontal gradient Xu. Later Wang and Liao [30] considered the case of $1 < p < 2$ for inhomogeneous systems in Carnot groups. Furthermore, Wang, Liao and Gao [31] weakened assumptions on coefficients A_i^α with Hölder continuity in the variables (ξ, u) to the assumptions of Dini continuity, and proved partial regularity result with optimal estimates for the modulus of continuity for the horizontal derivative Xu.

Regularity results for discontinuous sub-elliptic systems with VMO coefficients instead of continuous coefficients have been established in the work [7] by Di Fazio and Fanciullo, and [19] by Gao, Niu and Wang for the case of quadratic growth; [32] by Wang and Manfredi, [14] by Dong and Niu, [36] by Zheng and Feng, and [33] by Wang, Zhang and Yang for non-quadratic growth conditions. We note that the regularity results in [14] and [36] have a limitation of p near 2, and the result in [19] holds only under a strong smallness condition for the dimension. In contrast, our partial Hölder continuity result stated below, is valid for the full range $1 < p < 2$ in any dimension.

The typical strategy in partial regularity depends on decay estimates for certain excess functionals, which measure the oscillations of the solution or its gradient in a suitable sense. In this paper, we are working with a combination of a zero-order excess functional C_y and a first-order excess functional Ψ. For the case $p \ge 2$, the functional Ψ is defined by

$$\Psi(\xi_0, \rho, l) = \fint_{B_\rho(\xi_0)} \left[\left| \frac{u - l}{\rho(1 + |Xl|)} \right|^2 + \left| \frac{u - l}{\rho(1 + |Xl|)} \right|^p \right] d\xi,$$

with the horizontal affine functions $l : \mathbb{R}^{2n} \to \mathbb{R}^N$ defined in the subsection 2.2 below. It is straightforward to adapt the standard $\mathcal{A}$-harmonic approximation lemma by utilizing L^2-theory combined with the standard Sobolev inequality; see Wang and Manfredi [32] for the super-quadratic natural growth case. However, in the present situation, we treat the case of sub-quadratic controllable growth, and sub-quadratic natural growth, respectively. So one should establish the decay estimate for the following excess functional

$$\Psi(\xi_0, \rho, l) = \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u - l}{\rho} \right) \right|^2 d\xi, \tag{1.9}$$

where $V(A) = (1 + |A|^2)^{\frac{4}{p-2}} A$ for $A \in \mathbb{R}^k$, $k \in \mathbb{N}_+$. On the other hand, we define the Campanato type excess functional C_y by

$$C_y(\xi_0, \rho) = \rho^{-py} \fint_{B_\rho(\xi_0)} |u - u_{\xi_0,\rho}|^p d\xi, \quad 1 < p < 2, \ 0 < y < 1,$$

which provides a measure of the oscillations in the weak solutions u itself. It is remarkable that the excess functionals defined above involve only u, which simplifies the proofs of our partial regularity results. It is shown that if Ψ is small enough on a ball $B_{\xi_0}(\rho) \subset\subset U$, then, for some fixed $\theta \in (0,1)$, one obtain an excess improvement $\Psi(\xi_0, \theta r, l_{\xi_0,\theta r}) \le C_4\theta^2\Psi_\star(\xi_0, r, l_{\xi_0,r})$ under smallness condition assumptions; see for example, Lemma 4.3. At this point, one has to assume smallness on the $\Psi_\star$-excess. Also we note that such an excess improvement estimate has two different quantities Ψ and $\Psi_\star$ on the left, and the right hand side, respectively. Therefore, in contrast to the standard proof of partial regularity, the excess improvement cannot be iterated directly to yield an excess-decay estimate for Ψ-excess. In the present situation, however, iteration of the excess improvement yields that the Ψ-excess in (1.9) and also the C_y-excess remain bounded. Finally, the boundedness of the C_y-excess on any scale leads immediately to desired Hölder continuity of weak solutions u via the integral characterization of continuity by Campanato. We point out that the idea of such a combination of two excess functionals has its origin by Foss and Mingione [17] for continuous vector fields and integrands, and then, adapted to discontinuous problems with VMO coefficients for $p \ge 2$ by Bögelein-Duzaar-Habermann-Scheven [1]. It is worth mentioning that we obviously do not have access to use L^2-theory for functions in the horizontal Sobolev space $HW^{1,p}$ with $1 < p < 2$. Therefore, we have to establish the following Sobolev-Poincaré inequality with the function V (see Lemma 3.1 below),

$$\left(\fint_{B_\rho(\xi_0)}\left|V\left(\frac{u-u_{\xi_0,\rho}}{\rho}\right)\right|^{\frac{2Q}{Q-p}}d\xi\right)^{\frac{Q-p}{2Q}} \le C_P\left(\fint_{B_\rho(\xi_0)}|V(Xu)|^2\,d\xi\right)^{\frac{1}{2}},$$

with the constant C_P dependence only on N, p, Q. This inequality is an essential tool in order to get the regularity result. It is also one technique point where our case differs from the case $p \ge 2$ in [32].

Under the previous assumptions **(H1)**- **(H4)** and **(HC)**, and **(H1)**- **(H4)** and **(HN)**, respectively, we establish the following two partial Hölder continuity results.

Theorem 1.1. *Assume that coefficients $A_i^\alpha(\xi, u, Xu)$ and $B^\alpha(\xi, u, Xu)$ satisfy the assumptions **(H1)**- **(H4)** and **(HC)**. Let $u \in HW^{1,p}(\Omega, \mathbb{R}^N)$ with $1 < p < 2$ be weak solutions to the systems (1.1), i.e.,*

$$\int_\Omega A_i^\alpha(\xi, u, Xu)\cdot X\varphi\, d\xi = \int_\Omega B^\alpha(\xi, u, Xu)\cdot\varphi\, d\xi, \quad \forall\varphi \in C_0^\infty(\Omega, \mathbb{R}^N). \tag{1.10}$$

Then, there exists a relatively closed singular set $\Omega_0 \subset \Omega$ such that $u \in C^{0,y}_{loc}(\Omega\backslash\Omega_0,\ \mathbb{R}^N)$ for every $y \in (0,1)$. Moreover, for any $\lambda \in (0, Q)$ we have $Xu \in L^{p,\lambda}_{loc}(\Omega\backslash\Omega_0,\ \mathbb{R}^{2n\times N})$ with the Morrey parameter $\lambda = Q - p(1-y)$. Finally, we have that the singular set satisfies $\Omega_0 \subset \Sigma_1 \cup \Sigma_2$, where

$$\Sigma_1 = \left\{\xi_0 \in \Omega : \limsup_{r\to 0}\left|(Xu)_{\xi_0,r}\right| = \infty\right\},$$
$$\Sigma_2 = \left\{\xi_0 \in \Omega : \liminf_{r\to 0}\fint_{B_r(\xi_0)}\left|V(Xu) - V\left((Xu)_{\xi_0,r}\right)\right|^2 d\xi > 0\right\}$$

with the functional V defined in (2.3), and the singular set has $(2n+1)$-Lebesgue measure zero $|\Omega_0| = 0$ and its complement $\Omega \setminus \Omega_0$ is a set of full measure in Ω.

Theorem 1.2. *Assume that coefficients $A_i^\alpha(\xi, u, Xu)$ and $B^\alpha(\xi, u, Xu)$ satisfy the assumptions **(H1)**- **(H4)** and **(HN)**. Let $u \in HW^{1,p}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ be weak solutions to the system (1.1). Then, we have the same results that $u \in C^{0,y}_{loc}(\Omega\backslash\Omega_0,\ \mathbb{R}^N)$ and $Xu \in L^{p,\lambda}_{loc}(\Omega\backslash\Omega_0,\ \mathbb{R}^{2n\times N})$ as Theorem 1.*

Remark 1.3. *It is worth noting that the choice*

$$A_i^\alpha(\xi, u, P) = a(\xi)\left(1+|P|^2\right)^{\frac{p-2}{2}}P_i^\alpha \quad \text{for } i \in \{1,\cdots,2n\}, \alpha \in \{1,\cdots,N\}$$

makes the sub-elliptic p-Laplacian system with VMO-coefficients, involving sub-quadratic growth terms

$$-\sum_{i=1}^{2n} X_i\left(A_i^\alpha(\xi)\left(1+|Xu|^2\right)^{\frac{p-2}{2}}X_i u^\alpha\right) = B^\alpha(\xi, u, Xu)$$

just as a special case of (1.1), where $A_i^\alpha(\xi) \in VMO$, and $1 < p < 2$. So, combining the result for $2 \le p < \infty$ established by Wang and Manfredi in [32], our partial Hölder continuity results covers the model case of sub-elliptic p-Laplacian system with $1 < p < \infty$. It is remarkable that Zheng and Feng [36] showed everywhere regularity for weak solutions of sub-elliptic p-harmonic systems while p is very closed to 2.

The organization of this paper is as follows. In Section 2, we collect some basic notions and facts associated to Heisenberg groups, involving quasi-distance, horizontal Sobolev spaces, and horizontal affine function and some estimates. In Section 3, firstly an appropriate Sobolev-poincaré inequality which plays an important part on proving Hölder regularity is established. Then, an $\mathcal{A}$-harmonic approximation lemma, and a prior estimate for weak solution $h \in HW^{1,1}$ to the constant coefficient homogeneous sub-elliptic systems are given. In Section 4, we prove partial regularity results of Theorem 1.1 under sub-quadratic controllable structure assumptions **(H1)-(H4)** and **(HC)** by several steps. Step 1 is to gain a suitable Caccioppoli-type inequality which is an essential tool to get partial regularity. An appropriate linearization strategy is given in the second step. Then, one can achieve that solutions are approximately $\mathcal{A}$-harmonic by the linearization procedure, and an excess improvement estimate for the functional Ψ is obtained under two smallness condition assumptions, by combining with $\mathcal{A}$-harmonic approximation lemma in the third steps. Once the excess improvement is established, the iteration for the Ψ-excess and the C_y-excess can be acquired in Step 4. Finally, we show boundedness of the Campanato-type excess which leads immediately to desired Hölder continuity and Morrey regularity of Theorem 1.1. The last section shows the results of Theorem 1.2 under sub-quadratic natural structure assumptions **(H1)-(H4)** and **(HN)**. In such a case, we establish appropriate estimates just for the natural growth term, and the rest procedure is similar to the proof of Theorem 1.1.

2 Preliminaries

In this section, we will give introduction of the Heisenberg group $\mathbb{H}^n$ and definitions of several function spaces, and some elementary estimates which will be used later.

2.1 Introduction of the Heisenberg group $\mathbb{H}^n$

The Heisenberg group $\mathbb{H}^n$ is defined as $\mathbb{R}^{2n+1}$ endowed with the following group multiplication:

$$(\bar{\xi}, t) \cdot (\bar{\eta}, \tilde{t}) = \left(\bar{\xi} + \bar{\eta}, t + \tilde{t} + \frac{1}{2} \sum_{i=1}^{n} \left(x^i \tilde{y}^i - \tilde{x}^i y^i \right) \right),$$

for all $\xi = (\bar{\xi}, t) = (x^1, x^2, \cdots, x^n, y^1, y^2, \cdots, y^n, t)$, $\tilde{\xi} = (\bar{\eta}, \tilde{t}) = (\tilde{x}^1, \tilde{x}^2, \cdots, \tilde{x}^n, \tilde{y}^1, \tilde{y}^2, \cdots, \tilde{y}^n, \tilde{t})$. Its neutral element is $\mathbf{0}$, and its inverse to $(\bar{\xi}, t)$ is given by $(-\bar{\xi}, -t)$.

The basic vector fields corresponding to its Lie algebra can be explicitly calculated, and are given by

$$X_i \equiv X_i(\xi) = \frac{\partial}{\partial x_i} - \frac{y_i}{2} \frac{\partial}{\partial t}, \quad X_{n+i} \equiv X_{n+i}(\xi) = \frac{\partial}{\partial y_i} + \frac{x_i}{2} \frac{\partial}{\partial t}, \quad T \equiv T(\xi) = \frac{\partial}{\partial t}$$

for $i = 1, 2, \cdots, n$, and note that the special structure of the commutators:

$$[X_i, X_{i+n}] = -[X_{i+n}, X_i] = T, \quad \text{else} \quad [X_i, X_j] = 0, \quad \text{and} \quad [T, T] = [T, X_i] = 0,$$

that is, $(\mathbb{H}^n, \cdot)$ is a nilpotent Lie group of step 2. $X = (X_1, \cdots, X_{2n})$ is said to be the horizontal gradient, and T vertical derivative.

The homogeneous norm is defined by $\|(\bar{\xi}, t)\|_{\mathbb{H}^n} = \left(\|\bar{\xi}\|^4 + 16t^2 \right)^{1/4}$, and the metric induced by this homogeneous norm is given by

$$d(\tilde{\xi}, \xi) = \left\| \xi^{-1} \cdot \tilde{\xi} \right\|_{\mathbb{H}^n}.$$

The measure used on $\mathbb{H}^n$ is the Haar measure (Lebesgue measure in $\mathbb{R}^{2n+1}$), and the volume of the homogeneous ball $B_R(\xi_0) = \{\xi \in \mathbb{H}^n : d(\xi_0, \xi) < R\}$ is given by $\left|B_R(\xi_0)\right|_{\mathbb{H}^n} = R^{2n+2}\left|B_1(\xi_0)\right|_{\mathbb{H}^n} \triangleq \omega_n R^Q$, where the number

$$Q = 2n + 2 \tag{2.1}$$

is called the homogeneous dimension of $\mathbb{H}^n$, and the quantity ω_n is the volume of the homogeneous ball of radius 1.

Let $1 \le p \le +\infty$. We denote by

$$HW^{1,p}(\Omega) = \left\{u \in L^p(\Omega) | X_i u \in L^p(\Omega), i = 1, \cdots, k\right\}$$

the horizontal Sobolev space. Then, $HW^{1,p}(\Omega)$ is a Banach space under the norm

$$\|u\|_{HW^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{i=1}^{k} \|X_i u\|_{L^p(\Omega)}.$$

For $u \in HW^{1,q}(B_R(\xi_0))$, $1 < q < Q$ and $1 \le p \le \frac{qQ}{Q-q}$, Lu [29] showed the following Poincaré type inequality associated with Hörmander vector fields, which is naturally valid for $\mathbb{H}^n$:

$$\left(\fint_{B_R(\xi_0)} \left|u - u_{\xi_0,R}\right|^p d\xi\right)^{\frac{1}{p}} \le C_p R \left(\fint_{B_R(\xi_0)} |Xu|^q d\xi\right)^{\frac{1}{q}}. \tag{2.2}$$

The inequality (2.2) is valid for $p = q$ (≥ 2).

Throughout the paper, we shall use the functions V, $W : \mathbb{R}^k \to \mathbb{R}^k$ defined by

$$V(\varsigma) = (1 + |\varsigma|^2)^{\frac{p-2}{4}}\varsigma, \quad W(\varsigma) = \varsigma/(1 + |\varsigma|^{2-p})^{\frac{1}{2}} \tag{2.3}$$

for each $\varsigma \in \mathbb{R}^k$, $k \in \mathbb{N}$ and $p > 1$. The functions V and W are locally bi-Lipschitz bijection on $\mathbb{R}^k$.

The following inequality

$$\left(1 + |\varsigma|^2\right)^{\frac{2-p}{2}} \leqslant 1 + |\varsigma|^{2-p} \leqslant 2^{\frac{p}{2}}\left(1 + |\varsigma|^2\right)^{\frac{2-p}{2}},$$

immediately yields

$$\left|W(\varsigma)\right| \leqslant \left|V(\varsigma)\right| \leqslant 2^{\frac{p}{4}}\left|W(\varsigma)\right|. \tag{2.4}$$

The purpose of introducing W is the fact that in contrast to $|V|^{\frac{2}{m}}$, the function $|W|^{\frac{2}{m}}$ is convex. In fact, firstly by direct computation yields that $W^{\frac{2}{p}}(t) = t^{\frac{2}{p}}(1 + t^{2-p})^{\frac{-1}{p}}$ is a convex and monotone increasing function on $[0, \infty)$ with $W^{\frac{2}{p}}(0) = 0$; secondly we have

$$\left|W\left(\frac{\varsigma_1 + \varsigma_2}{2}\right)\right|^{\frac{2}{p}} \le W\left(\frac{|\varsigma_1| + |\varsigma_2|}{2}\right)^{\frac{2}{p}} \le \frac{|W(\varsigma_1)|^{\frac{2}{p}} + |W(\varsigma_2)|^{\frac{2}{p}}}{2}, \quad \varsigma_1, \varsigma_2 \in \mathbb{R}^n.$$

The following lemma includes some useful properties of the function V. The proof can be found in Lemma 2.1 of [4]. For simplicity, here, we replace the coefficient $2^{(p-2)/4}$ with $\frac{1}{\sqrt{2}}$ in the left of the first inequality (1) below, since the fact that $2^{-1/2} < 2^{(p-2)/4}$ for $p > 0$.

Lemma 2.1. *Let $p \in (1, 2)$ and $V : \mathbb{R}^k \to \mathbb{R}^k$ be the functions defined in (2.3). Then, for any $\varsigma_1, \varsigma_2 \in \mathbb{R}^k$ and $t > 0$, the following inequalities hold:*

(1). $\frac{1}{\sqrt{2}}\min\left(|\varsigma_1|, |\varsigma_1|^{\frac{p}{2}}\right) \le 2^{(p-2)/4}\min\left(|\varsigma_1|, |\varsigma_1|^{\frac{p}{2}}\right) \leqslant |V(\varsigma_1)| \leqslant \min\left(|\varsigma_1|, |\varsigma_1|^{\frac{p}{2}}\right)$;

(2). $\left|V(t\varsigma_1)\right| \leqslant \max\left(t, t^{\frac{p}{2}}\right)\left|V(\varsigma_1)\right|$;

(3). $\left|V(\varsigma_1 + \varsigma_2)\right| \leqslant C(p)(\left|V(\varsigma_1)\right| + \left|V(\varsigma_2)\right|)$;

(4). $\frac{p}{2}|\varsigma_1 - \varsigma_2| \leqslant \left|V(\varsigma_1) - V(\varsigma_2)\right|/(1 + |\varsigma_1|^2 + |\varsigma_2|^2)^{\frac{p-2}{4}} \leqslant C(p, k)|\varsigma_1 - \varsigma_2|$;

(5). $\left|V(\varsigma_1) - V(\varsigma_2)\right| \leqslant C(p, k)\left|V(\varsigma_1 - \varsigma_2)\right|$;

(6). $\left|V(\varsigma_1 - \varsigma_2)\right| \leqslant C(p, M)\left|V(\varsigma_1) - V(\varsigma_2)\right|$ *for all ς_2 with $|\varsigma_2| \le M$.*

2.2 Horizontal affine function and estimates in $\mathbb{H}^n$

Let $u \in L^2(B_\rho(\xi_0), \mathbb{R}^N)$, $\xi_0 \in \mathbb{R}^{2n+1}$, and consider the horizontal components

$$\bar{\xi} = (x^1, \dots, x^n, y^1, \dots, y^n) \quad \text{and} \quad \bar{\xi_0} = (x_0^1, \dots, x_0^n, y_0^1, \dots, y_0^n).$$

If the function

$$l_{\xi_0,\rho}(\bar{\xi}) = l_{\xi_0,\rho}(\bar{\xi_0}) + Xl_{\xi_0,\rho}(\bar{\xi} - \bar{\xi_0}),$$

minimizes the functional

$$l \mapsto \fint_{B_\rho(\xi_0)} |u - l|^2 d\xi,$$

among horizontal affine function $l : \mathbb{R}^{2n} \to \mathbb{R}^N$, then, we have

$$l_{\xi_0,\rho}(\bar{\xi_0}) = u_{\xi_0,\rho} = \fint_{B_\rho(\xi_0)} u d\xi,$$

and

$$Xl_{\xi_0,\rho} = \frac{Q-2}{c_0 Q} \frac{Q+2}{\rho^2} \fint_{B_\rho(\xi_0)} u \otimes (\bar{\xi} - \bar{\xi_0}) d\xi, \tag{2.5}$$

where the vector $u \otimes (\bar{\xi} - \bar{\xi_0})$ has components $u^\alpha(x^1 - x_0^1, x^2 - x_0^2, \dots, x^{2n} - x_0^{2n})$ with $\alpha = 1, 2, \dots, N$, and c_0 is a positive constant defined by

$$c_0 = \frac{\int_0^\pi (\sin\theta)^n d\theta}{\int_0^\pi (\sin\theta)^{n-1} d\theta} = \begin{cases} \dfrac{[(2k-2)!!]^2}{(2k-1)!!(2k-3)!!} \dfrac{2}{\pi}, n = 2k-1, \\ \dfrac{[(2k-1)!!]^2}{(2k)!!(2k-2)!!} \dfrac{\pi}{2}, n = 2k. \end{cases} \tag{2.6}$$

Here, we use the notation $(2k-2)!! = (2k-2)(2k-4)\cdots 4 \times 2$ and $(2k-1)!! = (2k-1)(2k-3)\cdots 3 \times 1$. The proof of the results above can be found in [32] by Wang and Manfredi. On the basis of this formula, elementary calculations yield the following estimates.

Lemma 2.2. *Let $u \in L^2(B_\rho(\xi_0), \mathbb{R}^N)$, $\theta \in (0,1)$.We denote by $l_{\xi_0,\rho}$ and $l_{\xi_0,\theta\rho}$, the horizontal affine function defined as above for the radii ρ and $\theta\rho$. Then, we have*

$$|Xl_{\xi_0,\rho} - Xl_{\xi_0,\theta\rho}|^p \le \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{(\theta\rho)}\right)^p \fint_{B_{\theta\rho}(\xi_0)} \left|u - l_{\xi_0,\rho}\right|^p d\xi, \tag{2.7}$$

and, more generally,

$$\left|Xl_{\xi_0,\rho} - Xl\right|^p \le \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{\rho}\right)^p \fint_{B_\rho(\xi_0)} |u - l|^p d\xi, \tag{2.8}$$

for every horizontal affine function $l : \mathbb{R}^{2n} \to \mathbb{R}^N$.

Proof. By the identity (2.5) and Hölder's inequality, we obtain

$$\begin{aligned} \left|Xl_{\xi_0,\rho} - Xl_{\xi_0,\theta\rho}\right|^p &= \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{(\theta\rho)^p}\right)^p \left| \fint_{B_{\theta\rho}(\xi_0)} \left(u - l_{\xi_0,\rho}(\bar{\xi_0}) - Xl_{\xi_0,\rho}(\bar{\xi} - \bar{\xi_0})\right) \otimes (\bar{\xi} - \bar{\xi_0}) d\xi \right|^p \\ &\le \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{(\theta\rho)^2}\right)^p \fint_{B_{\theta\rho}(\xi_0)} \left|u - l_{\xi_0,\rho}(\bar{\xi_0}) - Xl_{\xi_0,\rho}(\bar{\xi} - \bar{\xi_0})\right|^p d\xi \left(\fint_{B_{\theta\rho}(\xi_0)} \left|\bar{\xi} - \bar{\xi_0}\right|^{\frac{p}{p-1}} d\xi \right)^{p-1} \\ &\le \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{(\theta\rho)}\right)^p \fint_{B_{\theta\rho}(\xi_0)} \left|u - l_{\xi_0,\rho}\right|^p d\xi, \end{aligned} \tag{2.9}$$

where we have used the fact that $\left(\fint_{B_{\theta\rho}(\xi_0)} \left|\bar{\xi} - \bar{\xi_0}\right|^{\frac{p}{p-1}} d\xi\right)^{p-1} \le (\theta\rho)^{-p}$.

For (2.8), it follows

$$\begin{aligned}\left|Xl_{\xi_0,\rho} - Xl\right|^p &\le \left(\frac{Q-2}{c_0 Q}\frac{Q+2}{\rho^2}\right)^p \fint_{B_\rho(\xi_0)} \left|u - l(\bar{\xi}_0) - Xl(\bar{\xi} - \bar{\xi}_0)\right|^p d\xi \left(\fint_{B_\rho(\xi_0)} \left|\bar{\xi} - \bar{\xi}_0\right|^{\frac{p}{p-1}} d\xi\right)^{p-1} \\ &= \left(\frac{Q-2}{c_0 Q}\frac{Q+2}{\rho}\right)^p \fint_{B_{\theta\rho}(\xi_0)} |u - l|^p \, d\xi.\end{aligned}$$

□

According to the definition of the function $l_{\xi_0,\rho}$, the following version of the Poincaré inequality (2.2) is true, that is,

$$\left(\fint_{B_\rho(\xi_0)} \left|u - l_{\xi_0,\rho}(\bar{\xi})\right|^p d\xi\right)^{\frac{1}{p}} \le C_p \rho \left(\fint_{B_\rho(\xi_0)} \left|Xu - Xl_{\xi_0,\rho}\right|^q d\xi\right)^{\frac{1}{q}},$$

where $1 < q < Q$, $1 \le p \le \frac{qQ}{Q-q}$.

3 Sobolev-Poincaré type inequality and $\mathcal{A}$-harmonic approximation

We know that L^2-theory cannot be directly used to obtain appropriate estimates for solutions $u \in HW^{1,p}$ with $1 < p < 2$, so in this section, we first establish a suitable version of Sobolev-Poincaré type inequality with functions V. This inequality is an essential tool in proving partial regularity. Then, we give a prior estimate for $\mathcal{A}$-harmonic functions $h \in HW^{1,1}$, and introduce an $\mathcal{A}$-harmonic approximation lemma which plays an important part in getting excess improvement estimates.

Lemma 3.1. *(Sobolev-Poincaré type inequality). Let $p \in (1, 2)$ and $u \in HW^{1,p}(B_\rho(\xi_0), \mathbb{R}^N)$ with $B_\rho(\xi_0) \subset \Omega$, then, it follows*

$$\left(\fint_{B_\rho(\xi_0)} \left|V\left(\frac{u - u_{\xi_0,\rho}}{\rho}\right)\right|^{\frac{2p^*}{p}} d\xi\right)^{\frac{p}{2p^*}} \le C_P \left(\fint_{B_\rho(\xi_0)} |V(Xu)|^2 \, d\xi\right)^{\frac{1}{2}}, \tag{3.1}$$

with $p^* = \frac{pQ}{Q-p}$ the Sobolev critical exponent of p; here the constant C_P depends only on Q, N and p. In particular, the inequality also holds if we substitute 2 for $\frac{2p^*}{p}$.

Proof. We introduce the operator of fractional integration on Ω of order 1 as follows

$$I_1(f)(\xi) = \int_\Omega |f(\eta)| \frac{d(\xi,\eta)}{B\left(\xi, d(\xi,\eta)\right)} d\eta, \quad \xi \in B_\rho(\xi_0).$$

Based on Theorem 2.7 in [3] by Capogna, Danielli and Garofalo, we deduce for $1 < p < +\infty$

$$\left(\fint_{B_\rho(\xi_0)} |I_1(f)(\xi)|^{p^*} d\xi\right)^{\frac{1}{p^*}} = C\rho \left(\fint_{B_\rho(\xi_0)} |f(\xi)|^p d\xi\right)^{\frac{1}{p}}, \tag{3.2}$$

where we denote by $p^* = \frac{pQ}{Q-p}$ the Sobolev critical exponent of p, and the number Q the homogeneous dimension in $\mathbb{H}^n$.

Lu [21] gave a representation formula for a function on graded nilpotent Lie groups for the left invariant vector fields; see Lemma 3.1 there. One form of the representation states that there exist constants $c > 1$ and $C > 1$ such that

$$\left|u(\xi) - u_{\xi_0,\rho}\right| \le C \int_{B_{c\rho}(\xi_0)} |Xu(\eta)| \frac{d(\xi,\eta)}{B\left(\xi, d(\xi,\eta)\right)} d\eta, \quad \xi \in B_\rho(\xi_0).$$

Noting that $W^{2/p}(t)$ is monotone increasing and convex, we apply $W^{2/p}(t)$ to both sides of the last inequality and have by Jensen's inequality

$$W^{2/p}\left(\left|\frac{u(\xi)-u_{\xi_0,\rho}}{\rho}\right|\right) \le \frac{C}{\rho}\int_{\mathbb{R}^{2n+1}} \widetilde{W}^{2/p}\left(Xu(\eta)\right)\frac{d(\xi,\eta)}{B\left(\xi,d(\xi,\eta)\right)}d\eta,$$

and

$$\widetilde{W}\left(\left|Xu(\eta)\right|\right) = \begin{cases} 0, & \eta \notin B_{c\rho}(\xi_0), \\ W\left(\left|Xu(\eta)\right|\right), & \eta \in B_{c\rho}(\xi_0). \end{cases}$$

One can check that $W\left(\left|Xu(\eta)\right|\right) \in L^p\left(B_\rho(\xi_0)\right)$, which implies $\widetilde{W}\left(\left|Xu(\eta)\right|\right) \in L^p\left(\mathbb{R}^{2n+1}\right)$. Then, applying the inequality (3.2) yields

$$\begin{aligned}
\left[\int_{B_\rho(\xi_0)}\left|W\left(\frac{u(\xi)-u_{\xi_0,\rho}}{\rho}\right)\right|^{\frac{2Q}{Q-p}}d\xi\right]^{\frac{Q-p}{pQ}} &= \left[\int_{B_\rho(\xi_0)}\left[W^{2/p}\left(\frac{\left|u(\xi)-u_{\xi_0,\rho}\right|}{\rho}\right)\right]^{\frac{pQ}{Q-p}}d\xi\right]^{\frac{Q-p}{pQ}} \\
&\le \frac{C}{\rho}\left[\fint_{B_\rho(\xi_0)}\left(\int_{\mathbb{R}^{2n+1}}\widetilde{W}^{2/p}\left(Xu(\eta)\right)\frac{d(\xi,\eta)}{B\left(\xi,d(\xi,\eta)\right)}d\eta\right)^{\frac{pQ}{Q-p}}d\xi\right]^{\frac{Q-p}{pQ}} \\
&\le \frac{C}{\rho}\left[\fint_{B_\rho(\xi_0)}\left|I_1\left(W^{2/p}(Xu)\right)(\xi)\right|^{\frac{pQ}{Q-p}}d\xi\right]^{\frac{Q-p}{pQ}} \\
&\le C\left[\fint_{B_\rho(\xi_0)}W^2\left(\left|Xu(\xi)\right|\right)d\xi\right]^{\frac{1}{p}},
\end{aligned}$$

or

$$\left[\fint_{B_\rho(\xi_0)}\left|W\left(u(\xi)-u_{\xi_0,\rho}\right)\right|^{\frac{2Q}{Q-p}}d\xi\right]^{\frac{Q-p}{2Q}} \le C\left[\fint_{B_\rho(\xi_0)}W^2\left(\left|Xu(\xi)\right|\right)d\xi\right]^{\frac{1}{2}}.$$

We obtain the assertion of the theorem, first for W, and then, also for V by (2.4). □

Let $\mathcal{A} \in Bil(\Omega\times\mathbb{R}^N\times\mathbb{R}^{2n\times N},\mathbb{R}^{2n\times N})$ be a bilinear form with constant tensorial coefficients. We say that a map $h \in C^\infty(B_\rho(\xi_0),\mathbb{R}^N)$ is $\mathcal{A}$-harmonic if and only if

$$\fint_{B_\rho(\xi_0)}\mathcal{A}(Xh,X\varphi)d\xi = 0$$

holds for all testing function $\varphi \in C_0^\infty(B_\rho(\xi_0),\mathbb{R}^N)$.

Shores in [25] showed that weak solutions $h \in HW^{1,2}(\Omega,\mathbb{R}^N)$ of the constant coefficient homogeneous sub-elliptic systems belongs to C^∞ in the subset $\Omega_0 \subset \Omega$. Then, the following estimate holds for the solution $h \in HW^{1,2}(\Omega,\mathbb{R}^N)$,

$$\sup_{B_{\rho/2}(\xi_0)}\left(|Xh|^2+\left|X^2h\right|^2\right) \le C_0\rho^{-2}\fint_{B_\rho(\xi_0)}|Xh|^2\,d\xi.$$

Therefore, we can argue as the proof of Proposition 2.10 in [4] to obtain the same estimate for any $\mathcal{A}$-harmonic function $h \in HW^{1,1}(\Omega,\mathbb{R}^N)$.

Lemma 3.2. *Let $h \in HW^{1,1}(\Omega,\mathbb{R}^N)$ be weak solutions of the constant coefficient systems. Then, h is smooth and there exists $C_0 \ge 1$ such that for any $B_\rho(\xi_0) \subset \Omega$*

$$\sup_{B_{\rho/2}(\xi_0)}\left(|Xh|^2+\left|X^2h\right|^2\right) \le C_0\rho^{-2}\fint_{B_\rho(\xi_0)}|Xh|^2\,d\xi. \tag{3.3}$$

Similarly to [10], one can establish the following version of $\mathcal{A}$-harmonic approximation for the case $1 < p < 2$ in Heisenberg groups.

Lemma 3.3. *Let $0 < \nu \le L$ and $1 < p < 2$ be given. For every $\varepsilon > 0$, there is a constant $\delta = \delta(Q, N, p, \nu, L, \varepsilon) \in (0, 1]$ assume that $y \in [0, 1]$ and that $\mathcal{A}$ is a bilinear form on $\mathbb{R}^{2n\times N}$ with the properties*

$$\mathcal{A}(P, P) \ge \nu|P|^2 \quad \text{and} \quad \mathcal{A}(P, \bar{P}) \le L|P||\bar{P}|, \quad P, \bar{P} \in \mathbb{R}^{2n\times N}.$$

Furthermore, let $w \in HW^{1,p}(B_\rho(\xi_0), \mathbb{R}^N)$ be an approximate $\mathcal{A}$-harmonic map in the sense that the following estimate holds

$$\fint_{B_\rho(\xi_0)} \mathcal{A}(Xw, X\varphi)d\xi \le \delta y \sup_{B_\rho(\xi_0)} |X\varphi|, \quad \forall \varphi \in C_0^\infty(B_\rho(\xi_0), \mathbb{R}^N),$$

and

$$\fint_{B_\rho(\xi_0)} |V(Xw)|^2 d\xi \le y^2.$$

Then, there exists an $\mathcal{A}$-harmonic map $h \in C^\infty(B_\rho(\xi_0), \mathbb{R}^N)$ which satisfies

$$\fint_{B_\rho(\xi_0)} \left| V\left(\frac{w - yh}{\rho}\right)\right|^2 d\xi \le y^2\varepsilon \quad \text{and} \quad \fint_{B_\rho(\xi_0)} |V(Xh)|^2 d\xi \le 1.$$

4 Partial Hölder continuity for sub-quadratic controllable growth

In this section, we prove the partial regularity result of Theorem 1 under the assumptions of sub-quadratic controllable structure conditions. Now we begin with the following.

4.1 Caccioppoli-type inequality

We know that Caccioppoli-type inequality is a preliminary tool to prove partial regularity for systems. So in this subsection, we shall prove a Caccioppoli-type inequality for weak solutions to the sub-elliptic systems (1.1) with sub-quadratic controllable growth conditions.

Lemma 4.1. *(Caccioppoli-type inequality). Let $u \in HW^{1,p}(\Omega, \mathbb{R}^N)$ be weak solutions of the nonlinear sub-elliptic systems (1.1) under the assumptions **(H1)-(H4)-(HC)**. Then, for any $\xi_0 = (x^1, \cdots, x^n, y^1, \cdots, y^n, t) \in \Omega$ with $B_r(\xi_0) \subset\subset \Omega$, and any horizontal affine functions $l : \mathbb{R}^{2n} \to \mathbb{R}^N$ with $|l(\bar{\xi_0})| + |Xl| \le M_0$, we have the estimate*

$$\begin{aligned}\fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xu - Xl)|^2 d\xi \le & C_c \left[\fint_{B_r(\xi_0)} \left| V\left(\frac{u - l}{r}\right)\right|^2 d\xi + \omega\left(\fint_{B_r(\xi_0)} (|u - l(\bar{\xi_0})|^p)d\xi\right) + \boldsymbol{V}(r)\right.\\ & \left. + (r^2 + r^{p'})\left[\fint_{B_r(\xi_0)} (|Xu|^p + |u|^{p^*} + 1)d\xi\right]^{\frac{p'}{(p^*)'}}\right],\end{aligned}$$

where C_c is a positive constants depending only on Q, p, ν, L, M_0, and the exponent $p' = \frac{p}{p-1}$, and $(p^*)' = \frac{p^*}{p^*-1}$ with $p^* = \frac{pQ}{Q-p}$.

Proof. We choose a standard cut-off function $\phi \in C_0^\infty(B_r(\xi_0), [0, 1])$ with $\phi \equiv 1$ on $B_{\frac{r}{2}}(\xi_0)$ and $|X\phi| \le \frac{4}{r}$. Then, $\varphi = \phi^2(u - l)$ can be taken as a testing function for sub-elliptic systems (1.1). Hence, we have

$$\fint_{B_r(\xi_0)} A_i^\alpha(\xi, u, Xu)\phi^2(Xu - Xl)d\xi = -2\fint_{B_r(\xi_0)} A_i^\alpha(\xi, u, Xu)\phi(u - l)X\phi d\xi + \fint_{B_r(\xi_0)} B^\alpha(\xi, u, Xu)\phi^2(u - l)d\xi,$$

where we have used the fact that $X\varphi = \phi^2(Xu - Xl) + 2\phi(u - l)X\phi$.

In view of the identities $\fint_{B_r(\xi_0)} \left(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl)\right)_{\xi_0,r} X\varphi d\xi = 0$, and

$$-\fint_{B_r(\xi_0)} A_i^\alpha(\xi, u, Xl)\phi^2(Xu - Xl)d\xi = 2\fint_{B_r(\xi_0)} A_i^\alpha(\xi, u, Xl)\phi(u-l)X\phi d\xi - \fint_{B_r(\xi_0)} A_i^\alpha(\xi, u, Xl)X\varphi d\xi.$$

It follows for weak solutions u of systems (1.1) that

$$\begin{aligned} I_0 :=& \fint_{B_r(\xi_0)} [A_i^\alpha(\xi, u, Xu) - A_i^\alpha(\xi, u, Xl)]\phi^2(Xu - Xl)d\xi \\ =& 2\fint_{B_r(\xi_0)} [A_i^\alpha(\xi, u, Xl) - A_i^\alpha(\xi, u, Xu)]\phi(u-l)X\phi d\xi \\ &+ \fint_{B_r(\xi_0)} [A_i^\alpha(\xi, l(\bar{\xi_0}), Xl) - A_i^\alpha(\xi, u, Xl)]X\varphi d\xi \\ &+ \fint_{B_r(\xi_0)} [(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl))_{\xi_0,r} - A_i^\alpha(\xi, l(\bar{\xi_0}), Xl)]X\varphi d\xi \\ &+ \fint_{B_r(\xi_0)} B^\alpha(\xi, u, Xu)\phi^2(u-l)d\xi \\ =:& 2I_1 + I_2 + I_3 + I_4, \end{aligned} \tag{4.1}$$

with the obvious labelling for $I_0 - I_4$.

We first estimate the left-hand side of (4.1). By the first inequality of (1.2), Young's inequality and definition of the function V (2.3), we have

$$\begin{aligned} I_0 =& \fint_{B_r(\xi_0)} \int_0^1 \langle D_P A_i^\alpha(\xi, u, Xl + s(Xu - Xl))(Xu - Xl), (Xu - Xl)\rangle \phi^2 ds\, d\xi \\ \geq& \fint_{B_r(\xi_0)} \int_0^1 \nu(1 + |Xl + s(Xu - Xl)|)^{p-2}|Xu - Xl|^2\phi^2 ds\, d\xi \\ \geq& \fint_{B_r(\xi_0)} \nu\left[3(1 + |Xl|^2 + |Xu - Xl|^2)\right]^{\frac{p-2}{2}} |Xu - Xl|^2\phi^2 d\xi \\ \geq& \fint_{B_r(\xi_0)} \nu\left[3(1 + M_0^2)\right]^{\frac{p-2}{2}} \left(1 + |Xu - Xl|^2\right)^{\frac{p-2}{2}} |Xu - Xl|^2\phi^2 d\xi \\ =& \nu\left[3(1 + M_0^2)\right]^{\frac{p-2}{2}} \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi, \end{aligned} \tag{4.2}$$

where we have used the elementary inequality $1 + |a|^2 + |b-a|^2 \leqslant 3\left(1 + |a|^2 + |b|^2\right)$, and $1 < p < 2$.

Now, we are going to estimate the terms $I_1 - I_4$ on the right-hand side of (4.1). For small positive $\epsilon < 1$ appearing in lines, it will be fixed later.

Estimate for I_1. We shall decompose the ball $B_r(\xi_0)$ into four subsets: $\Omega_1 := B_r(\xi_0) \cap \{|Xu - Xl| \leq 1\} \cap \left\{\left|\frac{u-l}{r}\right| \leq 1\right\}$, $\Omega_2 := B_r(\xi_0) \cap \{|Xu - Xl| \leq 1\} \cap \left\{\left|\frac{u-l}{r}\right| \geq 1\right\}$, $\Omega_3 := B_r(\xi_0) \cap \{|Xu - Xl| \geq 1\} \cap \left\{\left|\frac{u-l}{r}\right| \leq 1\right\}$, and $\Omega_4 := B_r(\xi_0) \cap \{|Xu - Xl| \geq 1\} \cap \left\{\left|\frac{u-l}{r}\right| \geq 1\right\}$.

Case 1: Using the second inequality of (1.2), $|X\phi| \leq \frac{4}{r}$, Young's inequality, and Lemma 2.1, we derive the following bound for I_1 on the subset Ω_1.

$$\begin{aligned} &\fint_{\Omega_1} [A_i^\alpha(\xi, u, Xu) - A_i^\alpha(\xi, u, Xl)]\phi(u-l)X\phi d\xi \\ &\leq L\fint_{\Omega_1} \int_0^1 (1 + |Xl + s(Xu - Xl)|)^{p-2} ds\, |Xu - Xl||u-l||X\phi|\phi d\xi \\ &\leq 4L\fint_{\Omega_1} \phi|\sqrt{2}V(Xu - Xl)|\left|\frac{u-l}{r}\right| d\xi \\ &\leq 2\varepsilon\fint_{\Omega_1} \phi^2|V(Xu - Xl)|^2 d\xi + 32L^2\varepsilon^{-1}\fint_{\Omega_1} \left|\left(\frac{u-l}{r}\right)\right|^2 d\xi, \end{aligned} \tag{4.3}$$

where we have used the inequality $(1+|Xl+s(Xu-Xl)|)^{p-2}\le 1$ for $1<p<2$.

Case 2: Similarly to the case 1, there is

$$\begin{aligned}
&\fint_{\Omega_2}[A_i^\alpha(\xi,u,Xu)-A_i^\alpha(\xi,u,Xl)]\phi(u-l)X\phi d\xi\\
&\le 2^{\frac{p}{2(p-1)}}\varepsilon\fint_{\Omega_2}\phi^{\frac{p}{p-1}}|V(Xu-Xl)|^{\frac{p}{p-1}}d\xi+(4L)^p\varepsilon^{1-p}\fint_{\Omega_2}\left|\frac{u-l}{r}\right|^p d\xi\\
&\le 2\varepsilon\fint_{\Omega_2}\phi^2|V(Xu-Xl)|^2d\xi+2(4L)^p\varepsilon^{1-p}\fint_{\Omega_2}\left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi,
\end{aligned}\tag{4.4}$$

where we have used Lemma 2.1, and the inequality $|V(Xu-Xl)|^{\frac{p}{p-1}}\le|V(Xu-Xl)|^2\le|Xu-Xl|^2\le 1$ for $1<p<2$ on the set Ω_2.

Case 3: By Young's inequality and Lemma 2.1, it follows that on the subset Ω_3,

$$\begin{aligned}
&\fint_{\Omega_3}[A_i^\alpha(\xi,u,Xu)-A_i^\alpha(\xi,u,Xl)]\phi(u-l)X\phi d\xi\\
&\le\varepsilon\fint_{\Omega_3}\phi^p|Xu-Xl|^p d\xi+(4L)^{\frac{p}{p-1}}\varepsilon^{\frac{1}{1-p}}\fint_{\Omega_3}\left|\frac{u-l}{r}\right|^2 d\xi\\
&\le 2\varepsilon\fint_{\Omega_3}\phi^2|V(Xu-Xl)|^2d\xi+2(4L)^{\frac{p}{p-1}}\varepsilon^{\frac{1}{1-p}}\fint_{\Omega_3}\left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi,
\end{aligned}\tag{4.5}$$

where we have used the fact that $\left|\frac{u-l}{r}\right|^{\frac{p}{p-1}}\le\left|\frac{u-l}{r}\right|^2$ as $\frac{p}{p-1}\ge 2$ and $\left|\frac{u-l}{r}\right|\le 1$ on the subset Ω_3.

Case 4: On the subset Ω_4, it holds that by the assumption $|l(\bar{\xi}_0)|+|Xl|\le M_0$

$$\begin{aligned}
&\fint_{\Omega_4}[A_i^\alpha(\xi,u,Xu)-A_i^\alpha(\xi,u,Xl)]\phi(u-l)X\phi d\xi\\
&\le 2^{2+\frac{1}{p}}L\fint_{\Omega_4}\phi\left|V(Xu-Xl)\right|^{\frac{2}{p}}\left|\frac{u-l}{r}\right|d\xi\\
&\le\fint_{\Omega_4}\varepsilon\phi^{\frac{p}{p-1}}\left|V(Xu-Xl)\right|^{\frac{2}{p-1}}d\xi+2^{2p+1}L^p\fint_{\Omega_4}\varepsilon^{1-p}\left|\frac{u-l}{r}\right|^p d\xi\\
&\le\varepsilon\fint_{\Omega_4}\phi^2|V(Xu-Xl)|^2d\xi+C(p,L)\varepsilon^{1-p}\fint_{\Omega_4}\left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi,
\end{aligned}\tag{4.6}$$

here, we have used the smallness assumption $\Phi(\xi_0,r,l):=\fint_{B_r(\xi_0)}|V(Xu-Xl)|^2d\xi\le 1$ and $\phi^{\frac{p}{p-1}}\le\phi^2$.

From (4.3), (4.4), (4.5) and (4.6), we have the estimate for the term I_1 as follows

$$I_1\le 2\varepsilon\fint_{B_r(\xi_0)}\phi^2|V(Xu-Xl)|^2d\xi+C(p,L,M_0)\varepsilon^{\frac{1}{1-p}}\fint_{B_r(\xi_0)}\left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi,\tag{4.7}$$

where we have used the inequality $\varepsilon^{\frac{1}{1-p}}\ge\varepsilon^{-1}\ge\varepsilon^{1-p}$ for small positive constant $\varepsilon<1$.

Estimate for I_2. By the first inequality of (1.3), we get

$$\begin{aligned}
I_2\le&L\fint_{B_r(\xi_0)}\omega\left(|u-l(\bar{\xi}_0)|^p\right)(1+|Xl|)^{p-1}|X\varphi|d\xi\\
\le&C(p,L,M_0)\fint_{B_r(\xi_0)}\omega\left(|u-l(\bar{\xi}_0)|^p\right)\phi^2|Xu-Xl|d\xi\\
&+C(p,L,M_0)\fint_{B_r(\xi_0)}\omega\left(|u-l(\bar{\xi}_0)|^p\right)\phi|u-l||X\phi|d\xi\\
=:&I_{21}+I_{22}.
\end{aligned}\tag{4.8}$$

To estimate the term I_{21}, we divide the domain of integration into two parts $\Omega_5:=B_r(\xi_0)\cap\{|Xu-Xl|\le 1\}$ and $\Omega_6:=B_r(\xi_0)\cap\{|Xu-Xl|>1\}$.

Case 1: On the set Ω_5 where $|Xu - Xl| \le 1$, it holds

$$\begin{aligned} I_{21}\,(\text{on }\Omega_5) \le & 2\varepsilon \fint_{\Omega_5} \phi^4 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{-1} \fint_{\Omega_5} \omega^2 \left(|u - l(\bar{\xi}_0)|^p\right) d\xi \\ \le & 2\varepsilon \fint_{\Omega_5} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{-1} \omega \left(\fint_{\Omega_5} |u - l(\bar{\xi}_0)|^p d\xi \right), \end{aligned} \tag{4.9}$$

where we have used in turn Young's inequality, $\omega^2 \le \omega$, the concavity of ω and Jensen's inequality.

Case 2: On the part Ω_6 where $|Xu - Xl| > 1$, we find

$$\begin{aligned} I_{21}\,(\text{on }\Omega_6) \le & \varepsilon \fint_{\Omega_6} \phi^{2p} |Xu - Xl|^p d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \fint_{\Omega_6} \omega^{\frac{p}{p-1}} \left(|u - l(\bar{\xi}_0)|^p\right) d\xi \\ \le & 2\varepsilon \fint_{\Omega_6} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega \left(\fint_{\Omega_6} |u - l(\bar{\xi}_0)|^p d\xi \right), \end{aligned} \tag{4.10}$$

where we have used the inequality $\omega^{\frac{p}{p-1}} \le \omega$.

Combining (4.9) with (4.10) leads to

$$I_{21} \le 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega \left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi \right), \tag{4.11}$$

where we have use the fact $\varepsilon^{\frac{1}{1-p}} \ge \varepsilon^{-1}$ for $0 < \varepsilon < 1$.

The term I_{22} can be estimated similarly as I_{21} above. Here, we split the ball $B_r(\xi_0)$ into two subsets $\Omega_7 := B_r(\xi_0) \cap \{|\frac{u-l}{r}| \le 1\}$ and $\Omega_8 := B_r(\xi_0) \cap \{\frac{u-l}{r}| > 1\}$.

Case 1: On the subset Ω_7, it yields

$$I_{22}\,(\text{on }\Omega_7) \le 2\varepsilon \fint_{\Omega_7} \phi^2 \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{-1} \omega \left(\fint_{\Omega_7} |u - l(\bar{\xi}_0)|^p d\xi \right). \tag{4.12}$$

Case 2: We deduce on Ω_8

$$I_{22}\,(\text{on }\Omega_8) \le 2\varepsilon \fint_{\Omega_8} \phi^p \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega \left(\fint_{\Omega_8} |u - l(\bar{\xi}_0)|^p d\xi \right). \tag{4.13}$$

From (4.12) and (4.13), it follows

$$I_{22} \le 2\varepsilon \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega \left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi \right). \tag{4.14}$$

Joining (4.8), (4.11) and (4.14), we obtain

$$I_2 \le 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + 2\varepsilon \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega \left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi \right). \tag{4.15}$$

We are now in the position to handle the term I_3. By VMO-condition (1.5), the term I_3 can be estimated as follows

$$\begin{aligned} I_3 \le & \fint_{B_r(\xi_0)} \mathbf{v}_{\xi_0} (1 + |Xl|)^{p-1} |X\varphi| d\xi \\ \le & C(p, M_0) \fint_{B_r(\xi_0)} \mathbf{v}_{\xi_0} \phi^2 |Xu - Xl| d\xi + C(p, M_0) \fint_{B_r(\xi_0)} \mathbf{v}_{\xi_0} |u - l| \phi |X\phi| d\xi \\ =: & I_{31} + I_{32}. \end{aligned} \tag{4.16}$$

We can argue the terms I_{31} and I_{32} as the same way treating the terms I_{21} and I_{22}.

Case 1: On the set Ω_5 where $|Xu - Xl| \le 1$, we use $\mathbf{v}_{\xi_0} \le 2L$ and (1.6) to infer the following estimate

$$\begin{aligned} I_{31}\ (\text{on } \Omega_5) \le & \fint_{\Omega_5} \varepsilon\phi^4 \left|\sqrt{2}V(Xu - Xl)\right|^2 d\xi + C(p, M_0) \fint_{\Omega_5} \varepsilon^{-1}\mathbf{v}_{\xi_0}^2 d\xi \\ \le & 2\varepsilon \fint_{\Omega_5} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{-1}\mathbf{V}(r). \end{aligned} \tag{4.17}$$

Case 2: On the part Ω_6 where $|Xu - Xl| > 1$, we use (1.6) and the fact that $\mathbf{v}_{\xi_0}^{\frac{p}{p-1}} = \mathbf{v}_{\xi_0} \cdot \mathbf{v}_{\xi_0}^{\frac{1}{p-1}}$, $\mathbf{v}_{\xi_0} \le 2L$, to deduce

$$\begin{aligned} I_{31}\ (\text{on } \Omega_6) \le & \fint_{\Omega_6} \varepsilon\phi^{2p} |Xu - Xl|^p d\xi + C(p, M_0) \fint_{\Omega_6} \varepsilon^{\frac{1}{1-p}}\mathbf{v}_{\xi_0}^{\frac{p}{1-p}} d\xi \\ \le & 2\varepsilon \fint_{\Omega_6} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\mathbf{V}(r). \end{aligned} \tag{4.18}$$

Using (4.17) and (4.18), we get

$$I_{31} \le 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\mathbf{V}(r). \tag{4.19}$$

Similarly, the term I_{32} can be estimated as follows

$$I_{32} \le 2\varepsilon \fint_{B_r(\xi_0)} \left|V\left(\frac{u - l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\mathbf{V}(r). \tag{4.20}$$

Joining (4.16), (4.19) with (4.20), we have

$$I_3 \le 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + 2\varepsilon \fint_{B_r(\xi_0)} \left|V\left(\frac{u - l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\mathbf{V}(r). \tag{4.21}$$

Estimate for I_4. Using Hölder inequality, one has

$$\begin{aligned} I_4 \le & C \fint_{B_r(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right)^{1-\frac{1}{p^*}} \left|\phi^2(u - l)\right| d\xi \\ \le & C \left[\fint_{B_r(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{1-\frac{1}{p^*}} \left(\fint_{B_r(\xi_0)} \left|\phi^2(u - l)\right|^{p^*} d\xi\right)^{\frac{1}{p^*}}. \end{aligned} \tag{4.22}$$

To obtain an appropriate estimate for I_4, we take the domain $B_r(\xi_0)$ into four parts as the same way of I_1.

Case 1: For the case of $\Omega_1 = B_r(\xi_0) \cap \{|Xu - Xl| \le 1\} \cap \left\{\left|\frac{u-l}{r}\right| \le 1\right\}$, by Sobolev type inequality, Hölder's inequality, Young's inequality and Lemma 2.1, it follows that

$$\begin{aligned} & \left[\fint_{\Omega_1} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{1-\frac{1}{p^*}} \left[\left(\fint_{\Omega_1} \left|\phi^2(u - l)\right|^{p^*} d\xi\right)^{\frac{1}{p^*}}\right] \\ \le & C_p \left[\fint_{\Omega_1} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{1-\frac{1}{p^*}} r \left[\fint_{\Omega_1} \left(\left|\frac{u - l}{r}\right| + \phi^2 \left|\sqrt{2}V(Xu - Xl)\right|\right)^p d\xi\right]^{\frac{1}{p}} \\ \le & C(C_p, p, \varepsilon) r^2 \left[\fint_{\Omega_1} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{2\left(1-\frac{1}{p^*}\right)} + C_p\varepsilon \fint_{\Omega_1} \left|\frac{u - l}{r}\right|^2 \\ & + 2C_p\varepsilon \fint_{\Omega_1} \phi^4 \left|V(Xu - Xl)\right|^2 d\xi \\ \le & C(C_p, p, \varepsilon) r^2 \left[\fint_{\Omega_1} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{2\left(1-\frac{1}{p^*}\right)} + 2C_p\varepsilon \fint_{\Omega_1} \left|V\left(\frac{u - l}{r}\right)\right|^2 \\ & + 2C_p\varepsilon \fint_{\Omega_1} \phi^2 \left|V\left(Xu - Xl\right)\right|^2 d\xi. \end{aligned} \tag{4.23}$$

Case 2: On the part $\Omega_2 = B_r(\xi_0) \cap \{|Xu - Xl| \le 1\} \cap \left\{\left|\frac{u-l}{r}\right| \ge 1\right\}$, the following estimate holds

$$\begin{aligned}
&\left[\fint_{\Omega_2}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}}\left[\left(\fint_{\Omega_2}\left|\phi^2(u-l)\right|^{p^*}d\xi\right)^{\frac{1}{p^*}}\right]\\
\le& C(C_p,p,\varepsilon)r^{\frac{p}{p-1}}\left[\fint_{\Omega_2}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} + C_p\varepsilon\fint_{\Omega_2}\left|\frac{u-l}{r}\right|^p\\
&+ C(C_p,p,\varepsilon)r^2\left[\fint_{\Omega_2}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{2\left(1-\frac{1}{p^*}\right)} + C_p\varepsilon\fint_{\Omega_2}\phi^4\left|V(Xu-Xl)\right|^2 d\xi\\
\le& C(C_p,p,\varepsilon)\left(r^{\frac{p}{p-1}} + r^2\right)\left[\fint_{\Omega_2}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} + 2C_p\varepsilon\fint_{\Omega_2}\left|V\left(\frac{u-l}{r}\right)\right|^2\\
&+ 2C_p\varepsilon\fint_{\Omega_2}\phi^2\left|V\left(Xu-Xl\right)\right|^2 d\xi,
\end{aligned} \tag{4.24}$$

where we have used the fact that $\fint_{B_r(\xi_0)}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi \ge 1$, and $\phi \le 1$.

Case 3: On the part $\Omega_3 = B_r(\xi_0) \cap \{|Xu - Xl| \ge 1\} \cap \left\{\left|\frac{u-l}{r}\right| \le 1\right\}$, it yields

$$\begin{aligned}
&\left[\fint_{\Omega_3}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}}\left(\fint_{\Omega_3}\left|\phi^2(u-l)\right|^{p^*}d\xi\right)^{\frac{1}{p^*}}\\
\le& C_p\left[\fint_{\Omega_3}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}} r\left[\fint_{\Omega_3}\left(\left|\frac{u-l}{r}\right| + \phi^2\left|Xu-Xl\right|\right)^p d\xi\right]^{\frac{1}{p}}\\
\le& C(C_p,p)\left[\fint_{\Omega_3}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}}\left[r\left(\fint_{\Omega_3}\left|\frac{u-l}{r}\right|^2 d\xi\right)^{\frac{1}{2}} + r\left(\fint_{\Omega_3}\phi^{2p}\left|Xu-Xl\right|^p\right)^{\frac{1}{p}}d\xi\right]\\
\le& C(C_p,p,\varepsilon)\left(r^{\frac{p}{p-1}} + r^2\right)\left[\fint_{\Omega_3}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} + 2C_p\varepsilon\fint_{\Omega_3}\left|V\left(\frac{u-l}{r}\right)\right|^2\\
&+ 2C_p\varepsilon\fint_{\Omega_3}\phi^2\left|V\left(Xu-Xl\right)\right|^2 d\xi,
\end{aligned} \tag{4.25}$$

Case 4: For the last case of $\Omega_4 = B_r(\xi_0) \cap \{|Xu - Xl| \ge 1\} \cap \left\{\left|\frac{u-l}{r}\right| \ge 1\right\}$, we get

$$\begin{aligned}
&\left[\fint_{\Omega_4}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}}\left[\left(\fint_{\Omega_4}\left|\phi^2(u-l)\right|^{p^*}d\xi\right)^{\frac{1}{p^*}}\right]\\
\le& 2^{\frac{p-1}{p}}C_p\left[\fint_{\Omega_4}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{1-\frac{1}{p^*}}\left[r\left(\fint_{\Omega_4}\left|\frac{u-l}{r}\right|^p d\xi\right)^{\frac{1}{p}} + r\left(\fint_{\Omega_4}\phi^{2p}\left|Xu-Xl\right|^p\right)^{\frac{1}{p}}d\xi\right]\\
\le& C(C_p,p,\varepsilon)r^{\frac{p}{p-1}}\left[\fint_{\Omega_4}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} + 2C_p\varepsilon\fint_{\Omega_4}\left|\frac{u-l}{r}\right|^p\\
&+ 2C_p\varepsilon\fint_{\Omega_4}\phi^{2p}|Xu-Xl|^p\,d\xi\\
\le& C(C_p,p,\varepsilon)r^{\frac{p}{p-1}}\left[\fint_{\Omega_4}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} + 4C_p\varepsilon\fint_{\Omega_4}\left|V\left(\frac{u-l}{r}\right)\right|^2\\
&+ 4C_p\varepsilon\fint_{\Omega_4}\phi^2\left|V\left(Xu-Xl\right)\right|^2 d\xi.
\end{aligned} \tag{4.26}$$

Combining the estimates (4.22)-(4.26), we find that

$$I_4 \le 4C_p\varepsilon\fint_{B_r(\xi_0)}\phi^2\left|V\left(Xu-Xl\right)\right|^2 d\xi + C(C_p,p,\varepsilon)\left(r^2 + r^{\frac{p}{p-1}}\right)\left[\fint_{B_r(\xi_0)}\left(|Xu|^p + |u|^{p^*} + 1\right)d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}}$$

$$+ 4C_p\varepsilon \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2. \tag{4.27}$$

Joining the estimates(4.2),(4.7),(4.15),(4.21), (4.27) with (4.1), we arrive at

$$\begin{aligned}
&\left(3(1+M_0^2)\right)^{\frac{p-2}{2}} \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi \\
&\le (6\varepsilon + 4C_p\varepsilon) \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + \left[C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} + 4\varepsilon(1 + C_p)\right] \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi \\
&\quad + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\, \omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi_0})|^p d\xi\right) + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}}\, \mathbf{V}(r) \\
&\quad + C(C_p, p, \varepsilon)\left(r^2 + r^{\frac{p}{p-1}}\right)\left[\fint_{B_r(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}}.
\end{aligned}$$

Here, choosing $\varepsilon < \frac{\left(3(1+M_0^2)\right)^{\frac{p-2}{2}}}{6+4C_p}$, we can absorb the first integral of the right-hand side into the left. Keeping in mind the properties of ϕ, we have thus shown

$$\begin{aligned}
\fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xu - Xl)|^2 d\xi &\le 2^Q \fint_{B_r(\xi_0)} |V(Xu - Xl)|^2 \phi^2 d\xi \\
&\le C_c \left[\fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + \omega\left(\fint_{B_r(\xi_0)} (|u - l(\bar{\xi_0})|^p) d\xi\right) + \mathbf{V}(r)\right. \\
&\quad \left. + (r^2 + r^{p'})\left[\fint_{B_r(\xi_0)} (|Xu|^p + |u|^{p^*} + 1) d\xi\right]^{\frac{p'}{(p^*)'}}\right],
\end{aligned}$$

with a constant $C_c = C_c(Q, p, \nu, L, M_0)$, $p' = \frac{p}{p-1}$, and $(p^*)' = \frac{p^*}{p^*-1}$. This proves the claim. □

For sake of simplicity, we motivated the form of the Caccioppoli inequalities in Lemma 4.1. We set

$$\Phi(\xi_0, r, l) := \fint_{B_r(\xi_0)} |V(Xu - Xl)|^2 d\xi,$$

$$\Psi(\xi_0, r, l) := \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi,$$

$$\begin{aligned}
\Psi_\star(\xi_0, r, l) :=& \Psi(\xi_0, r, l) + \omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi_0})|^p d\xi\right) + \mathbf{V}(r) \\
&+ \left(r^2 + r^{p'}\right)\left[\fint_{B_r(\xi_0)} (|Xu|^p + |u|^{p^*} + 1) d\xi\right]^{\frac{p'}{(p^*)'}}.
\end{aligned}$$

In the sequel, when the choice of ξ_0 or l is clear, we frequently write $\Phi(r, l)$ or $\Phi(r)$ respectively, as a replacement of $\Phi(\xi_0, r, l)$.

4.2 Approximate $\mathcal{A}$-harmonicity of weak solutions

To apply $\mathcal{A}$-harmonic approximation lemma, we need to establish the following lemma, which provides a linearization strategy for non-linear sub-elliptic systems (1.1).

Lemma 4.2. *Under the assumptions of Theorem 1.1 are satisfied, $B_{2\rho}(\xi_0) \subseteq \Omega$ with $\rho \le \rho_0$ and an arbitrary horizontal function $l : \mathbb{R}^{2n} \to \mathbb{R}^N$, we define*

$$\mathcal{A} = \left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl)\right)_{\xi_0,\rho} \quad \text{and} \quad w = u - l,$$

then, w is approximately $\mathcal{A}$–harmonic in the sense that

$$\left| \fint_{B_\rho(\xi_0)} \mathcal{A}(Xw, X\varphi) d\xi \right| \le C_1 \left[\Psi_\star(2\rho) + \mu\left(\Psi_\star^{\frac{1}{2}}(2\rho)\right) + \mu\left(\Psi_\star^{\frac{1}{p}}(2\rho)\right) \right] \sup_{B_\rho(\xi_0)} |X\varphi|$$

for all $\varphi \in C_0^\infty(B_\rho(\xi_0), \mathbb{R}^N)$, and the positive constant $C_1 = C(p, M_0, L, C_c)$.

Proof. Without loss of generality, we assume that $\sup_{B_\rho(\xi_0)} |X\varphi| \le 1$. Noting that $w = u - l$, we compute

$$\begin{aligned}
&\fint_{B_\rho(\xi_0)} \mathcal{A}(Xw, X\varphi) d\xi \\
&\quad = \fint_{B_\rho(\xi_0)} \int_0^1 \left[\left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl)\right)_{\xi_0,\rho} - \left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl + sXw)\right)_{\xi_0,\rho} \right] Xw ds\, d\xi \sup_{B_\rho(\xi_0)} |X\varphi| \\
&\qquad + \fint_{B_\rho(\xi_0)} \int_0^1 \left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl + sXw)\right)_{\xi_0,\rho} Xw ds\, d\xi \sup_{B_\rho(\xi_0)} |X\varphi| \\
&\quad =: (J_1 + J_2) \sup_{B_\rho(\xi_0)} |X\varphi|,
\end{aligned} \tag{4.28}$$

with obvious labelling of J_1 and J_2.

In order to get the bound for the first term J_1, we first use the inequality (1.4) to obtain

$$\begin{aligned}
&\int_0^1 \left| \left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl)\right)_{\xi_0,\rho} - \left(D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl + sXw)\right)_{\xi_0,\rho} \right| ds \\
&\quad = \int_0^1 \left| \fint_{B_\rho(\xi_0)} D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl) - D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl + sXw) d\xi \right| ds \\
&\quad \le \int_0^1 \fint_{B_\rho(\xi_0)} \left| D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl) - D_P A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl + sXw) \right| d\xi ds \\
&\quad \le L \fint_{B_\rho(\xi_0)} \mu\left(\frac{|Xu - Xl|}{1 + |Xl|}\right) (1 + 2|Xl|)^{p-2} d\xi.
\end{aligned}$$

By the monotonicity of μ and the inequality above, it yields

$$\begin{aligned}
J_1 &\le L \fint_{B_\rho(\xi_0)} \mu\left(\frac{|Xu - Xl|}{1 + |Xl|}\right) (1 + 2|Xl|)^{p-2} |Xu - Xl| d\xi \\
&\le C(p, L, M_0) \fint_{B_\rho(\xi_0)} \mu(|Xu - Xl|) |Xu - Xl| d\xi.
\end{aligned}$$

Here, we decompose the ball $B_\rho(\xi_0)$ into two parts Ω_5 and Ω_6.

Case 1: On the domain Ω_5 where $|Xu - Xl| \le 1$, it follows by Lemma 2.1 Young's inequality, Jensen's inequality, and Hölder's inequality in turn

$$\begin{aligned}
J_1 \text{ (on } \Omega_5) &\le C(p, L, M_0) \fint_{\Omega_5} \mu\left(|V(Xu - Xl)|\right) |V(Xu - Xl)| d\xi \\
&\le C(p, L, M_0) \left[\mu^2 \left(\fint_{\Omega_5} |V(Xu - Xl)| d\xi \right) + \fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi \right] \\
&\le C(p, L, M_0) \left[\mu\left(\left(\fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi \right)^{\frac{1}{2}} \right) + \fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi \right],
\end{aligned}$$

where we have used the inequality $\mu^2 \le \mu$.

Case 2: On the set Ω_6 where $|Xu - Xl| > 1$, we have the following bound

$$\begin{aligned}
J_1 \text{ (on } \Omega_6) \le & C(p, L, M_0) \fint_{\Omega_6} \mu(|Xu - Xl|)|Xu - Xl| d\xi \\
\le & C(p, L, M_0) \left[\mu^{\frac{p}{p-1}} \left(\fint_{\Omega_6} |Xu - Xl| d\xi \right) + \fint_{\Omega_6} |Xu - Xl|^p d\xi \right] \\
\le & C(p, L, M_0) \left[\mu \left(\left(\fint_{\Omega_6} |Xu - Xl|^p d\xi \right)^{\frac{1}{p}} \right) + \fint_{\Omega_6} |V(Xu - Xl)|^2 d\xi \right] \\
\le & C(p, L, M_0) \left[\mu \left(\left(\fint_{\Omega_6} |V(Xu - Xl)|^2 d\xi \right)^{\frac{1}{p}} \right) + \fint_{\Omega_6} |V(Xu - Xl)|^2 d\xi \right],
\end{aligned}$$

where we have used $\mu^{\frac{p}{p-1}} \le \mu$ and Lemma 2.1. Then, we get the following estimate

$$J_1 \le C(p, L, M_0) \left[\mu(\Phi^{\frac{1}{2}}(\rho)) + \mu(\Phi^{\frac{1}{p}}(\rho)) + \Phi(\rho) \right]. \tag{4.29}$$

Based on the following facts

$$\fint_{B_\rho(\xi_0)} \langle A_i^\alpha(\xi, u, Xu), X\varphi \rangle d\xi - \fint_{B_\rho(\xi_0)} \langle B^\alpha(\xi, u, Xu), \varphi \rangle d\xi = 0 \quad \text{and} \quad \fint_{B_\rho(\xi_0)} \langle \left(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl) \right)_{\xi_0, \rho}, X\varphi \rangle d\xi = 0,$$

the integral J_2 can be rewritten as

$$\begin{aligned}
J_2 = & \fint_{B_\rho(\xi_0)} \left[\left(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xu) \right)_{\xi_0, \rho} - \left(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl) \right)_{\xi_0, \rho} \right] X\varphi d\xi \\
= & \fint_{B_\rho(\xi_0)} \left[\left(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xu) \right)_{\xi_0, \rho} - A_i^\alpha(\xi, l(\bar{\xi_0}), Xu) \right] X\varphi d\xi \\
& + \fint_{B_\rho(\xi_0)} \left[A_i^\alpha(\xi, l(\bar{\xi_0}), Xu) - A_i^\alpha(\xi, u, Xu) \right] X\varphi d\xi \\
& + \fint_{B_\rho(\xi_0)} B^\alpha(\xi, u, Xu)\varphi d\xi \\
=: & J_{21} + J_{22} + J_{23},
\end{aligned} \tag{4.30}$$

with the obvious meaning of $J_{21} + J_{22} + J_{23}$.

Using the assumption of $|l(\bar{\xi_0})| + |Xl| \le M_0$ and VMO-condition (1.5), We find that

$$\begin{aligned}
J_{21} \le & \fint_{B_\rho(\xi_0)} \mathbf{v}_{\xi_0} (1 + |Xu|)^{p-1} d\xi \\
\le & \fint_{B_\rho(\xi_0)} \mathbf{v}_{\xi_0} (1 + |Xl|^{p-1} + |Xu - Xl|^{p-1}) d\xi \\
\le & (1 + M_0^{p-1}) \fint_{B_\rho(\xi_0)} \left(\mathbf{v}_{\xi_0} + \mathbf{v}_{\xi_0} |Xu - Xl|^{p-1} \right) d\xi,
\end{aligned}$$

where we have used the inequality $0 < p - 1 < 1$ in the second line.

Now, we discuss it on the domain Ω_5 and Ω_6, respectively.

Case 1: On the set Ω_5 where $|Xu - Xl| \le 1$, the following estimate holds

$$\begin{aligned}
J_{21} \text{ (on } \Omega_5) \le & (1 + M_0^{p-1}) \fint_{\Omega_5} \left(\mathbf{v}_{\xi_0} + \mathbf{v}_{\xi_0} |\sqrt{2} V(Xu - Xl)|^{p-1} \right) d\xi \\
\le & (1 + M_0^{p-1}) \left(\fint_{\Omega_5} \mathbf{v}_{\xi_0} d\xi + \fint_{\Omega_5} \mathbf{v}_{\xi_0}^{\frac{2}{2-p}} d\xi + 2 \fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi \right) \\
\le & C(p, L, M_0) \left(\fint_{\Omega_5} \mathbf{v}_{\xi_0} d\xi + \fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi \right),
\end{aligned}$$

where we have used $\mathbf{v}_{\xi_0}^{\frac{2}{p-2}} = \mathbf{v}_{\xi_0} \cdot \mathbf{v}_{\xi_0}^{\frac{p}{2-p}}$, $\mathbf{v}_{\xi_0} \le 2L$ and Lemma 2.1.

Case 2: On the part Ω_6 where $|Xu - Xl| > 1$, it yields

$$\begin{aligned} J_{21}\ (\text{on } \Omega_6) \le & (1+M_0)\left(\fint_{\Omega_6} \mathbf{v}_{\xi_0} d\xi + \int_{\Omega_6} \mathbf{v}_{\xi_0}^{p} d\xi + \int_{\Omega_6} |Xu - Xl|^p d\xi\right) \\ \le & C(p, L, M_0)\left(\fint_{\Omega_6} \mathbf{v}_{\xi_0} d\xi + \int_{\Omega_6} |V(Xu - Xl)|^2 d\xi\right), \end{aligned}$$

where we have used the assumption $\mathbf{v}_{\xi_0} \le 2L$ and Lemma 2.1.

Then, we get the following estimate for J_{21}

$$J_{21} \le C(p, L, M_0)\left(\mathbf{V}(\rho) + \Phi(\rho)\right). \tag{4.31}$$

By first inequality of (1.3), the term J_{22} can be estimated as follows

$$\begin{aligned} J_{22} \le & L \fint_{B_\rho(\xi_0)} \omega\left(|u - l(\bar{\xi}_0)|^p\right)(1 + |Xu|)^{p-1} d\xi \\ \le & L(1 + M_0^{p-1}) \fint_{B_\rho(\xi_0)} \left[\omega\left(|u - l(\bar{\xi}_0)|^p\right) + \omega\left(|u - l(\bar{\xi}_0)|^p\right)|Xu - Xl|^{p-1}\right] d\xi. \end{aligned}$$

Similarly, for the case of $|Xu - Xl| \le 1$ on Ω_5, applying Young's inequality, Jensen's inequality and Lemma 2.1, we deduce that

$$J_{22}\ (\text{on } \Omega_5) \le C(L, p, M_0)\left[\omega\left(\fint_{\Omega_5} |u - l(\bar{\xi}_0)|^p d\xi\right) + \fint_{\Omega_5} |V(Xu - Xl)|^2 d\xi\right],$$

where we have used $\omega^{\frac{2}{2-p}} \le \omega \le 1$.
For the other case of $|Xu - Xl| > 1$ on Ω_6, it follows

$$J_{22}\ (\text{on } \Omega_6) \le C(L, p, M_0)\left[\omega\left(\fint_{\Omega_6} |u - l(\bar{\xi}_0)|^p d\xi\right) + \fint_{\Omega_6} |V(Xu - Xl)|^2 d\xi\right],$$

where we have used $\omega^p \le \omega \le 1$, Jensen's inequality and Lemma 2.1.

Thus, we get the following estimate for J_{22}

$$J_{22} \le C(L, p, M_0)\left[\omega\left(\fint_{B_\rho(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi\right) + \Phi(\rho)\right]. \tag{4.32}$$

Finally, we handle the term J_{23} by the same as the way for I_4 to obtain

$$\begin{aligned} J_{23} \le & 4C_p\varepsilon \fint_{B_\rho(\xi_0)} \phi^2 \left|V\left(Xu - Xl\right)\right|^2 d\xi + 4C_p\varepsilon \fint_{B_\rho(\xi_0)} \left|V\left(\frac{u-l}{\rho}\right)\right|^2. \\ & + C(C_p, \varepsilon)\left(\rho^2 + \rho^{\frac{p}{p-1}}\right)\left[\fint_{B_\rho(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}} \\ \le & C(C_p, \varepsilon)\left[\Phi(\rho) + \Psi(\rho) + \left(\rho^2 + \rho^{\frac{p}{p-1}}\right)\left[\fint_{B_\rho(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right]^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}}\right] \end{aligned} \tag{4.33}$$

Joining the estimates (4.31)-(4.33) with (4.30), we have

$$\begin{aligned} J_2 \le & C(p, L, M_0, C_p)\left[\Phi(\rho) + \Psi(\rho) + \omega\left(\fint_{B_\rho(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi\right) + \mathbf{V}(\rho)\right. \\ & \left. + \left(\rho^2 + \rho^{\frac{p}{p-1}}\right)\left(\fint_{B_\rho(\xi_0)} \left(|Xu|^p + |u|^{p^*} + 1\right) d\xi\right)^{\left(1-\frac{1}{p^*}\right)\frac{p}{p-1}}\right] \end{aligned}$$

$$= C(p, L, M_0, C_p)\left[\Phi(\rho) + \Psi_\star(\rho)\right]. \tag{4.34}$$

Plugging (4.29) and (4.34) into (4.28), we finally arrive at

$$\begin{aligned}\left|\fint_{B_\rho(\xi_0)} \mathcal{A}(Xw, X\varphi)d\xi\right| &\le C(p, L, M_0, C_p)\left[\mu(\Phi^{\frac{1}{2}}(\rho)) + \mu(\Phi^{\frac{1}{p}}(\rho)) + \Phi(\rho) + \Psi_\star(\rho)\right]\sup_{B_\rho(\xi_0)}|X\varphi| \\ &\le C(p, L, M_0, C_c, C_p)\left[\mu(\Psi_\star^{\frac{1}{2}}(2\rho)) + \mu(\Psi_\star^{\frac{1}{p}}(2\rho)) + \Psi_\star(2\rho)\right]\sup_{B_\rho(\xi_0)}|X\varphi|,\end{aligned}$$

where we have employed the Caccioppoli-type inequality from Lemma 4.1, $\Psi_\star(\rho) \le C(n, p)\Psi_\star(2\rho)$ in the last step. This yields the claim. □

4.3 Excess improvement

The strategy of our proof is to approximate the given solution in the sense of L^2 by $\mathcal{A}$-harmonic functions. Now we are in the position to establish the excess improvement.

Lemma 4.3. *Suppose that the assumptions of Theorem 1.1 are satisfied and consider a ball $B_r(\xi_0) \subseteq \Omega$ with $r \le \rho_0$. For the constants $\delta = \delta(Q, N, p, L, \nu, \theta^{Q+4}) \in (0, 1]$ and $y \in (0, 1]$ from the $\mathcal{A}$-harmonic approximation lemma 3.3, we let $0 < \theta \le \frac{1}{4}$ be arbitrary and also impose the following smallness conditions:*

(i). $\Psi_\star^{\frac{1}{2}}(r) < \frac{\delta}{2}$;

(ii). $y := \sqrt{\Psi_\star(r) + \left(\frac{\delta}{2}\right)^{-2}\left[\mu\left(\Psi_\star^{\frac{1}{2}}(r)\right) + \mu\left(\Psi_\star^{\frac{1}{p}}(r)\right)\right]^2} \le 1.$

Then, there holds an excess improvement estimate

$$\Psi(\xi_0, \theta r, l_{\xi_0,\theta r}) \le C_4\theta^2\Psi_\star(\xi_0, r, l_{\xi_0,r})$$

with some constants C_4 that depend only on n, N, p, ν, δ and L. Here, $l_{\xi_0,\theta r}$ and $l_{\xi_0,r}$ denote the minimizing affine functions introduced in Lemma 2.2.

Proof. We denote $\Psi_\star(r) = \Psi_\star(\xi_0, r, l_{\xi_0,r})$, and take

$$\tilde{w} = \frac{u - l_{\xi_0,r}}{C_2}$$

with $l_{\xi_0,r} = u_{\xi_0,r} + Xl_{\xi_0,r}(\bar{\xi} - \bar{\xi}_0)$ and $C_2 = \max\{C_1, \sqrt{C_c}\}$. We claim that $\tilde{w}$ satisfies the assumptions of $\mathcal{A}$-harmonic approximation lemma 3.3.

First note that, for our choice of the bilinear form

$$\mathcal{A} = \left(D_P A_i^\alpha(\cdot, l_{\xi_0,r}(\bar{\xi}_0), Xl_{\xi_0,r})\right)_{\xi_0,\rho}.$$

Next by Lemma 4.2 with $\rho = \frac{r}{2}$ and $l = l_{\xi_0,r}$, and the assumptions (i) and (ii), we find the map $\tilde{w}$ is approximately $\mathcal{A}$-harmonic in the sense that

$$\begin{aligned}\left|\fint_{B_{\frac{r}{2}}(\xi_0)} \mathcal{A}(X\tilde{w}, X\varphi)d\xi\right| &\le y\frac{C_1}{C_2}\left[\frac{\Psi_\star(r) + \mu\left(\Psi_\star^{\frac{1}{2}}(r)\right) + \mu\left(\Psi_\star^{\frac{1}{p}}(r)\right)}{y}\right]\sup_{B_{\frac{r}{2}}(\xi_0)}|X\varphi| \\ &\le y\left[\Psi_\star^{\frac{1}{2}}(r) + \frac{\delta}{2}\right]\sup_{B_{\frac{r}{2}}(\xi_0)}|X\varphi| \\ &\le y\delta\sup_{B_{\frac{r}{2}}(\xi_0)}|X\varphi|\end{aligned} \tag{4.35}$$

for all $\varphi \in C_0^\infty(B_{\frac{r}{2}}(\xi_0), \mathbb{R}^N)$, and

$$\fint_{B_{\frac{r}{2}}(\xi_0)} |V(X\tilde{w})|^2 d\xi = \frac{1}{C_2^2} \fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xu - Xl_{\xi_0,r})|^2 d\xi \le \frac{\Phi(r/2, l_{\xi_0,r})}{C_2^2} \le \frac{C_c}{C_2^2} \Psi_\star(r) \le y^2. \tag{4.36}$$

The estimates (4.35) and (5.2) tell us that the conditions of Lemma 3.3 are satisfied. So, there exists an $\mathcal{A}$-harmonic $h \in C_0^\infty(B_{\frac{r}{2}}(\xi_0), \mathbb{R}^N)$ such that

$$\fint_{B_{\frac{r}{2}}(\xi_0)} \left| V\left(\frac{\tilde{w} - yh}{r}\right) \right|^2 d\xi \le y^2\varepsilon, \quad \text{and} \quad \fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xh)|^2 d\xi \le 1.$$

In order to estimate excess functional

$$\Psi(\xi_0, \theta r, l_{\xi_0,\theta r}) = \fint_{B_{\theta r}(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\theta r}}{\theta r}\right) \right|^2 d\xi,$$

we now have to handle the integral $\fint_{B_{\theta r}(\xi_0)} \left|V(X^2h(\xi))\right|^2 d\xi$. Since the function $h(\xi)$ is $\mathcal{A}$-harmonic, we know that $h(\xi) \in C^\infty(\Omega)$ by Lemma 3.2. Noting that the boundedness $|Xh(\xi)| \le M$ in the ball $B_{\frac{r}{2}}(\xi_0) \subset\subset \Omega$, and using Hölder's inequality, we have the estimate for $\theta \in (0, \frac{1}{4})$

$$\begin{aligned}
\fint_{B_{\theta r}(\xi_0)} \left|V(X^2h(\xi))\right|^2 d\xi \le \sup_{B_{\frac{r}{4}}(\xi_0)} \left|X^2h(\xi)\right|^2 &\le C_0 r^{-2} \fint_{B_{\frac{r}{2}}(\xi_0)} \left|Xh(\xi)\right|^2 d\xi \\
&\le C_0 M r^{-2} \left[\left(\fint_{B_{\frac{r}{2}}(\xi_0)} \left|Xh(\xi)\right|^2 d\xi \right)^{\frac{1}{2}} + \left(\fint_{B_{\frac{r}{2}}(\xi_0)} \left|Xh(\xi)\right|^p d\xi \right)^{\frac{1}{p}} \right] \\
&\le 2C_0 M r^{-2} \left[\left(\fint_{B_{\frac{r}{2}}(\xi_0)} \left|V(Xh(\xi))\right|^2 d\xi \right)^{\frac{1}{2}} + \left(\fint_{B_{\frac{r}{2}}(\xi_0)} \left|V(Xh(\xi))\right|^2 d\xi \right)^{\frac{1}{p}} \right] \\
&\le C r^{-2},
\end{aligned} \tag{4.37}$$

where we have used the estimate (3.3) and Lemma 2.1.

We write $l^h(\xi) = h_{\xi_0,\theta r} + (Xh)_{\xi_0,\theta r}(\bar{\xi} - \bar{\xi}_0)$. Based on (3.1) and (4.37), it follows that

$$\begin{aligned}
&\fint_{B_{\theta r}(\xi_0)} \left| V\left(\frac{\tilde{w} - yl^h}{\theta r}\right) \right|^2 d\xi \\
&\le C \left[\fint_{B_{\theta r}(\xi_0)} \left| V\left(\frac{\tilde{w} - yh}{\theta r}\right) \right|^2 d\xi + \fint_{B_{\theta r}(\xi_0)} \left| yV\left(\frac{h - h_{\xi_0,\theta r} - (Xh)_{\xi_0,\theta r}(\bar{\xi} - \bar{\xi}_0)}{\theta r}\right) \right|^2 d\xi \right] \\
&\le C\theta^{-2}(2\theta)^{-Q} \fint_{B_{\frac{r}{2}}(\xi_0)} \left| V\left(\frac{\tilde{w} - yh}{r}\right) \right|^2 d\xi + CC_P y^2 \fint_{B_{\theta r}(\xi_0)} \left| V\left(Xh - (Xh)_{\xi_0,\theta r}\right) \right|^2 d\xi \\
&\le C\theta^{-2}(2\theta)^{-Q} y^2\varepsilon + CC_P^2(\theta r)^2 y^2 \fint_{B_{\theta r}(\xi_0)} \left| V(X^2h) \right|^2 d\xi \\
&\le C(C_P, C_0) y^2 \left(\theta^{-2-Q}\varepsilon + \theta^2\right) \le C(C_P, C_0)(1 + 16\delta^{-2})\theta^2 \Psi_\star(r),
\end{aligned}$$

where we have taken $\varepsilon = \theta^{Q+4}$.

Scaling back to u, we infer

$$\fint_{B_{\theta r}(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,r} - C_2 yl^h}{\theta r}\right) \right|^2 d\xi \le C_2^2 C(C_P, C_0)(1 + 16\delta^{-2})\theta^2 \Psi_\star(r).$$

In view of the defining property of $l_{\xi_0,\theta r}$, we arrive at

$$\fint_{B_{\theta r}(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\theta r}}{\theta r}\right)\right|^2 d\xi \le C_2^2 C(C_P, C_0)(1 + 16\delta^{-2})\theta^2 \Psi_\star(r) \le C_4 \theta^2 \Psi_\star(r),$$

here, we have denoted $C_4 = C_2^2 C(C_P, C_0)(1 + 16\delta^{-2})$. Then, it implies excess improvement estimate

$$\Psi(\xi_0, \theta r, l_{\xi_0,\theta r}) \le C_4 \theta^2 \Psi_\star(\xi_0, r, l_{\xi_0,r}).$$

□

4.4 Iteration

First, let $y \in (0, 1)$ be an fixed Hölder exponent. We define the Campanato-type excess

$$C_y(\xi_0, \rho) = \rho^{-py} \fint_{B_\rho(\xi_0)} |u - u_{\xi_0,\rho}|^p d\xi, \quad 1 < p < 2.$$

Next, we iterate the excess improvement estimate from Lemma 4.3.

Lemma 4.4. *Suppose that the assumptions of Theorem 1.1 are satisfied. For every $y \in (0, 1)$, there are constant $\varepsilon_\star, \kappa_\star, \rho_\star$ and $\theta \in (0, \frac{1}{8}]$, if*

$$\Psi(\xi_0, r, l_{\xi_0,r}) < \varepsilon_\star \quad \text{and} \quad C_y(\xi_0, r) < \kappa_\star, \tag{A_0}$$

for $r \in (0, \rho_\star)$ with $B_r(\xi_0) \subset \Omega$, then,

$$\Psi(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) < \varepsilon_\star \quad \text{and} \quad C_y(\xi_0, \theta^k r) < \kappa_\star, \tag{A_k}$$

respectively, for every $k \in \mathbb{N}$.

Proof. We begin by choosing the constants. First, we let

$$\theta = \min\left\{\left[\frac{c_0 Q}{\sqrt[p]{2}(Q-2)(Q+2)}\right]^{\frac{1}{1-y}}, \frac{1}{2C_4}\right\} \le \frac{1}{8}, \tag{4.38}$$

where c_0 is defined in (2.6), and C_4 is determined in Lemma 4.3, respectively. We note that the choice of θ fixes the constant $\delta = \delta(Q, N, p, \nu, L, \theta^{Q+4})$ from Lemma 3.3. Next, we fix an $\varepsilon_\star$ small sufficiently to ensure

$$\varepsilon_\star \le \min\left\{\left(\frac{\theta^{Q+py}}{8}\right)^{\frac{2}{p}}, \frac{\delta^2}{3\delta^2 + 48}\right\} \quad \text{and} \quad \mu(\varepsilon_\star) \le \varepsilon_\star. \tag{4.39}$$

Then, we choose $\kappa_\star > 0$ so small that

$$\omega(\kappa_\star) \le \varepsilon_\star.$$

Finally, we fix $\rho_\star > 0$ small enough to guarantee

$$\rho_\star \le \min\left\{\rho_0, \kappa_\star^{\frac{1}{p(1-y)}}, 1\right\}, \quad \mathbf{V}(\rho_\star) \le \varepsilon_\star, \quad \text{and} \quad \mathcal{F}(\rho_\star) \le \varepsilon_\star, \tag{4.40}$$

here we have abbreviated $\mathcal{F}(r) =: (r^2 + r^{p'})\left[\fint_{B_r(\xi_0)}(|Xu|^p + |u|^{p^*} + 1)d\xi\right]^{\frac{p'}{(p^*)'}}$.

Now, we are in the position to prove the assertion (A_k) by induction. We assume that (A_k) is true for up to some $k \in \mathbb{N}$. Then, we prove the first part of the assertion (A_{k+1}), that is, the one concerning $\Psi(\theta^{k+1} r, l_{\xi_0,\theta^{k+1} r})$. For this we are going to prove that the small assumptions for the excess improvement in Lemma 4.3 are satisfied. Firstly, by the assumptions of (A_k) and the choices of $\varepsilon_\star$, $\kappa_\star$ and $\rho_\star$, we deduce

$$\Psi_\star(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) \le \Psi(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) + \omega(C_y(\xi_0, \theta^k r)) + \mathbf{V}(\theta^k r) + \mathcal{F}(\theta^k r)$$

$$\begin{aligned}&\leq \varepsilon_\star + \omega(\kappa_\star) + \mathbf{V}(\rho_\star) + \mathcal{F}(\rho_\star)\\ &\leq 4\varepsilon_\star.\end{aligned}$$

Now it is easy to check that our choice of $\varepsilon_\star$ implies that the smallness condition assumptions (i)-(ii) in Lemma 4.3 are satisfied on the level $\theta^k r$, that is, we have

$$\Psi_\star^{\frac{1}{2}}(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) \leq 2\sqrt{\varepsilon_\star} \leq \frac{\delta}{2},$$

where we have used $\varepsilon_\star \leq \frac{\delta^2}{16}$ due to the choice of $\varepsilon_\star \leq \frac{\delta^2}{3\delta^2+48}$ in (4.39), and

$$\begin{aligned}y(\theta^k r) :=& \sqrt{\Psi_\star(\theta^k r) + \left(\frac{\delta}{2}\right)^{-2}\left[\mu\left(\sqrt{\Psi_\star(\theta^k r)}\right) + \mu\left(\sqrt[p]{\Psi_\star(\theta^k r)}\right)\right]^2}\\ \leq& \sqrt{3\varepsilon_\star + \left(\frac{\delta}{2}\right)^{-2}\left(\sqrt{3\varepsilon_\star} + \sqrt[p]{3\varepsilon_\star}\right)^2}\\ \leq& \sqrt{3\varepsilon_\star + \left(\frac{4}{\delta^2}\right)\left(2\sqrt{3\varepsilon_\star}\right)^2} = \sqrt{\varepsilon_\star\left(\frac{3\delta^2+48}{\delta^2}\right)} \leq 1.\end{aligned}$$

Consequently, we apply Lemma 4.3 with the radius $\theta^k r$ instead of r, this leads to

$$\Psi(\xi_0, \theta^{k+1} r, l_{\xi_0,\theta^{k+1}r}) \leq C_4\theta^2\Psi_\star(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) \leq 4C_4\theta^2\varepsilon_\star < \varepsilon_\star,$$

by the choice of θ in (4.38). This is the result for the first part of (A_{k+1}).

Now it remains to show the second part, that is, the one concerning $C_y(\xi_0, \theta^{k+1}r)$. Since $l_{\xi_0,\theta^k r} = u_{\xi_0,\theta^k r} + Xl_{\xi_0,\theta^k r}(\bar{\xi} - \bar{\xi}_0)$, we can estimate

$$\begin{aligned}C_y(\xi_0, \theta^{k+1}r) =& (\theta^{k+1}r)^{-py} \fint_{B_{\theta^{k+1}r}(\xi_0)} |u - u_{\xi_0,\theta^{k+1}r}|^p d\xi\\ \leq& (\theta^{k+1}r)^{-py} \fint_{B_{\theta^{k+1}r}(\xi_0)} |u - u_{\xi_0,\theta^k r}|^p d\xi\\ \leq& 2^{p-1}(\theta^{k+1}r)^{-py}\left[\fint_{B_{\theta^{k+1}r}(\xi_0)} |u - l_{\xi_0,\theta^k r}|^p d\xi + |Xl_{\xi_0,\theta^k r}|^p(\theta^{k+1}r)^p\right]\\ \leq& 2(\theta^{k+1}r)^{-py}\left[\theta^{-Q}\fint_{B_{\theta^k r}(\xi_0)} |u - l_{\xi_0,\theta^k r}|^p d\xi + |Xl_{\xi_0,\theta^k r}|^p(\theta^{k+1}r)^p\right]\\ =& 2(\theta^k r)^{p(1-y)}\left[\theta^{-Q-py}\fint_{B_{\theta^k r}(\xi_0)}\left|\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right|^p d\xi + |Xl_{\xi_0,\theta^k r}|^p\theta^{p(1-y)}\right].\end{aligned}$$

Now we are going to estimate the term $\fint_{B_{\theta^k r}(\xi_0)}\left|\frac{u-l_{\xi_0,\theta^k r}}{\theta^k r}\right|^p d\xi$. Similarly, we divide the domain of integration into two subsets $\Omega_9 := B_{\theta^k r}(\xi_0) \cap \left\{\left|\frac{u-l_{\xi_0,\theta^k r}}{\theta^k r}\right| > 1\right\}$ and $\Omega_{10} := B_{\theta^k r}(\xi_0) \cap \left\{\left|\frac{u-l_{\xi_0,\theta^k r}}{\theta^k r}\right| \leq 1\right\}$.

On the subset Ω_9, we get

$$\fint_{\Omega_9}\left|\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right|^p d\xi \leq 2\fint_{\Omega_9}\left|V\left(\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right)\right|^2 d\xi.$$

For the case of $\left|\frac{u-l_{\xi_0,\theta^k r}}{\theta^k r}\right| \leq 1$ on Ω_{10}, noting the fact of $1 < p < 2$, we find

$$\fint_{\Omega_{10}}\left|\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right|^p d\xi \leq \left(\fint_{\Omega_{10}}\left|\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right|^2 d\xi\right)^{\frac{p}{2}} \leq \left(2\fint_{\Omega_{10}}\left|V\left(\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r}\right)\right|^2 d\xi\right)^{\frac{p}{2}}.$$

Therefore, we deduce the following estimate

$$\fint_{B_{\theta^k r}(\xi_0)} \left| \frac{u - l_{\xi_0,\theta^k r}}{\theta^k r} \right|^p d\xi \le 2 \left(\fint_{B_{\theta^k r}(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\theta^k r}}{\theta^k r} \right) \right|^2 d\xi \right)^{\frac{p}{2}} \le 2\Psi^{\frac{p}{2}}(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) \le 2\varepsilon_\star^{\frac{p}{2}}.$$

By means of (2.8) with the choice of $l \equiv u_{\xi_0,\theta^k r}$, and the assumption A_k, we obtain

$$\begin{aligned} |Xl_{\xi_0,\theta^k r}|^p &\le \left(\frac{Q-2}{c_0 Q} \frac{Q+2}{\theta^k r} \right)^p \fint_{B_{\theta^k r}(\xi_0)} |u - u_{\xi_0,\theta^k r}|^p d\xi \\ &\le \left(\frac{(Q-2)(Q+2)}{c_0 Q} \right)^p (\theta^k r)^{p(\gamma-1)} C_\gamma(\xi_0, \theta^k r) \\ &\le \left[\frac{(Q-2)(Q+2)}{c_0 Q} \right]^p (\theta^k r)^{p(\gamma-1)} \kappa_\star. \end{aligned}$$

Recalling that $r \in (0, \rho_\star)$, we deduce that

$$C_\gamma(\xi_0, \theta^{k+1} r) \le 4\rho_\star^{p(1-\gamma)} \theta^{-Q-p\gamma} \varepsilon_\star^{\frac{p}{2}} + \left[\frac{(Q-2)(Q+2)}{c_0 Q} \right]^p \theta^{p(1-\gamma)} \kappa_\star \le \frac{\kappa_\star}{2} + \frac{\kappa_\star}{2} \le \kappa_\star,$$

where we have used the choice of $\varepsilon_\star$ in (4.39), the choice of $\rho_\star$ from (4.40) and θ in (4.38).

This proves the second part of the assertion (A_{k+1}) and finally complete the proof of the lemma. □

4.5 Proof of Theorem 1.

Proof. By Lebesgue's differentiation theorem, we obtain $|\Sigma_1 \cup \Sigma_2| = 0$. So our aim is to show that every $\xi_0 \in \Omega \setminus (\Sigma_1 \cup \Sigma_2)$ is a regular point. For every $0 < \rho < dist(\xi_0, \partial\Omega)$, by Sobolev-Poincaré type inequality, it follows

$$\begin{aligned} &\Psi(\xi_0, \rho, l_{\xi_0,\rho}) = \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\rho}}{\rho} \right) \right|^2 d\xi \\ &\le C_P^2 \fint_{B_\rho(\xi_0)} \left| V\left(Xu - (Xu)_{\xi_0,\rho} \right) \right|^2 d\xi \\ &\le C_P^2 C(p, M) \fint_{B_\rho(\xi_0)} \left| V(Xu) - V\left((Xu)_{\xi_0,\rho} \right) \right|^2 d\xi. \end{aligned} \tag{4.41}$$

Moreover, for any $\gamma \in (0, 1)$ and $\rho \le 1$, the following estimate holds

$$\begin{aligned} C_\gamma(\xi_0, \rho) &= \rho^{-p\gamma} \fint_{B_\rho(\xi_0)} |u - u_{\xi_0,\rho}|^p d\xi \\ &\le \rho^{p-p\gamma} \fint_{B_\rho(\xi_0)} \left| \frac{u - l_{\xi_0,\rho}}{\rho} \right|^p d\xi. \end{aligned}$$

If $\left| \frac{u - l_{\xi_0,\rho}}{\rho} \right| > 1$, we have

$$\fint_{B_\rho(\xi_0)} \left| \frac{u - l_{\xi_0,\rho}}{\rho} \right|^p d\xi \le 2 \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\rho}}{\rho} \right) \right|^2 d\xi.$$

If $\left| \frac{u - l_{\xi_0,\rho}}{\rho} \right| \le 1$, we obtain $\left| \frac{u - l_{\xi_0,\rho}}{\rho} \right|^p \le \left| \frac{u - l_{\xi_0,\rho}}{\rho} \right|^2 + 1$. Then, it implies

$$\fint_{B_\rho(\xi_0)} \left| \frac{u - l_{\xi_0,\rho}}{\rho} \right|^p d\xi \le 2 \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\rho}}{\rho} \right) \right|^2 d\xi + 1.$$

So, it yields

$$C_\gamma(\xi_0, \rho) \le 2\rho^{p-p\gamma} \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u - l_{\xi_0,\rho}}{\rho} \right) \right|^2 d\xi + \rho^{p-p\gamma} \le \rho^{p-p\gamma} \left[2\Psi(\xi_0, \rho, l_{\xi_0,\rho}) + 1 \right]. \tag{4.42}$$

Keeping in mind the definitions of Σ_1, Σ_2, from the estimates (4.41) and (4.42), we know that there exists a radius $\rho : 0 < \rho < min\{\rho_\star, dist(\xi_0, \partial\Omega)\}$ such that

$$\Psi(\xi_0, \rho, l_{\xi_0,\rho}) < \varepsilon_\star \quad \text{and} \quad C_y(\xi_0, \rho) < \kappa_\star.$$

Using the absolute continuity of the integral, there exists a neighborhood $U \subseteq \Omega$ of ξ_0 with

$$\Psi(\xi, \rho, l_{\xi_0,\rho}) < \varepsilon_\star \quad \text{and} \quad C_y(\xi, \rho) < \kappa_\star, \quad \forall \xi \in U.$$

Applying Lemma 4.4 in any point $\xi \in U$, then, we get

$$\Psi(\xi, \theta^k\rho, l_{\xi_0,\theta^k\rho}) < \varepsilon_\star \quad \text{and} \quad C_y(\xi, \theta^k\rho) < \kappa_\star, \quad \forall \xi \in U, k \in \mathbb{N}. \tag{4.43}$$

This together with Campanato's characterization of Hölder continuous functions imply that

$$\sup_{\xi\in U,\sigma\in(0,\rho)} \sigma^{-py} \fint_{B_\sigma(\xi)} |u - u_{\xi,\sigma}|^p d\eta = \sup_{\xi\in U,\sigma\in(0,\rho)} C_y(\xi, \sigma) < \kappa_\star < \infty.$$

Hence $u \in C^{0,y}_{loc}(U, \mathbb{R}^N)$.

Furthermore, it holds for $|Xu - Xl_{\xi,\sigma}| > 1$

$$\fint_{B_\sigma(\xi)} |Xu - Xl_{\xi,\sigma}|^p d\xi \le 2 \fint_{B_\sigma(\xi)} |V(Xu - Xl_{\xi,\sigma})|^2 d\xi, \tag{4.44}$$

and we have if $|Xu - Xl_{\xi,\sigma}| \le 1$

$$\fint_{B_\sigma(\xi)} |Xu - Xl_{\xi,\sigma}|^p d\xi \le 2 \fint_{B_\sigma(\xi)} |V(Xu - Xl_{\xi,\sigma})|^2 d\xi + 1. \tag{4.45}$$

Combining (4.44) and (4.45) with (4.43) and (1.6), we get for $y \in (0, 1)$

$$\begin{aligned}
&\sup_{\xi\in U,\sigma\in(0,\rho)} \sigma^{p(1-y)} \fint_{B_\sigma(\xi)} |Xu - Xl_{\xi,\sigma}|^p d\xi \\
&\le \sup_{\xi\in U,\sigma\in(0,\rho)} \sigma^{p(1-y)} \left[2 \fint_{B_\sigma(\xi)} |V(Xu - Xl_{\xi,\sigma})|^2 d\xi + 1 \right] \\
&\le \sup_{\xi\in U,\sigma\in(0,\rho)} \sigma^{p(1-y)} \left[2C_c \left(\Psi(\xi, \sigma, l_{\xi,\sigma}) + \omega\big(C_y(\xi, \sigma)\big) + \mathbf{V}(\sigma) + \mathcal{F}(\sigma) \right) + 1 \right] \le \infty,
\end{aligned}$$

with abbreviation of $\mathcal{F}(r) =: (r^2 + r^{p'}) \left[\fint_{B_r(\xi_0)} (|Xu|^p + |u|^{p^*} + 1) d\xi \right]^{\frac{p'}{(p^*)'}}$.

In view of the well known equivalence of Campanato and Morrey spaces for parameters $\lambda \in (0, Q)$, it yields $Xu \in L^{p,\lambda}(U, \mathbb{R}^{2n\times N})$ with $\lambda = Q - p(1 - y)$. In particular, the parameter λ can be chosen arbitrary chose to Q. This concludes the proof of Theorem 1.1. □

5 Partial Hölder continuity for sub-quadratic Natural growth

In this section, we prove the partial regularity result of Theorem 1.2 under the assumptions of sub-quadratic natural structure conditions **(H1)-(H4)** and **(HN)**. In this case, we will need to restrict ourselves to bounded solution of (1.1), where the bound $M = \sup_\Omega |u|$ satisfies the smallness assumption

$$2a(M)(M + M_0) \left(3(1 + M_0^2) \right)^{\frac{2-p}{2}} < \nu$$

in our present situation with $a(M)$ defined in (1.8). Such a similar smallness condition is necessary for a partial regularity result even in the elliptic case with quadratic growth ($p = 2$); for example, see [18].

We first introduce an elementary inequality showed by Kanazawa in [26]. It is useful to get suitable estimates for the natural growth term in proving Caccioppoli-type inequality.

Lemma 5.1. *Consider fixed a, $b \geq 0$, $p \geq 1$. Then, for any $\varepsilon > 0$, there exists $K = K(p, \varepsilon) \geq 0$ satisfying*

$$(a+b)^p \leq (1+\varepsilon)a^p + Kb^p. \tag{5.1}$$

Lemma 5.2. *(Caccioppoli-type inequality). Let $u \in HW^{1,p}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ be weak solutions of the systems (1.1) under the assumptions **(H1)-(H4)-(HN)** with $\nu > 2a(M)(M + M_0)\left(3(1+M_0^2)\right)^{\frac{2-p}{2}}$. Then, for any $\xi_0 = (x^1, \cdots, x^n, y^1, \cdots, y^n, t) \in \Omega$ and $r \leq 1$ with $B_r(\xi_0) \subset\subset \Omega$, and any horizontal affine functions $l : \mathbb{R}^{2n} \to \mathbb{R}^N$ with $|l(\bar{\xi_0})| + |Xl| \leq M_0$, we have the estimate*

$$\fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xu - Xl)|^2 d\xi \leq C_c \left[\fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + \omega\left(\fint_{B_r(\xi_0)} (|u - l(\bar{\xi_0})|^p) d\xi \right) + \boldsymbol{V}(r) + r^2 + r^{p'} \right],$$

where C_c is some positive constants depending only on $Q, N, p, a, b, L, \nu, M, M_0$, and the exponent $p' = \frac{p}{p-1}$.

Proof. Let $\varphi = \phi^2(u-l)$ be a testing function for sub-elliptic systems (1.1), where the standard cut-off function $\phi \in C_0^\infty(B_r(\xi_0), [0,1])$ with $\phi \equiv 1$ on $B_{\frac{r}{2}}(\xi_0)$ and $|X\phi| \leq \frac{4}{r}$. By the same way as the case of controllable growth, we have for weak solutions u of the systems (1.1)

$$\begin{aligned}
I_0' :=& \fint_{B_r(\xi_0)} [A_i^\alpha(\xi, u, Xu) - A_i^\alpha(\xi, u, Xl)]\phi^2(Xu - Xl)d\xi \\
=& 2\fint_{B_r(\xi_0)} [A_i^\alpha(\xi, u, Xl) - A_i^\alpha(\xi, u, Xu)]\phi(u-l)X\phi d\xi \\
&+ \fint_{B_r(\xi_0)} [A_i^\alpha(\xi, l(\bar{\xi_0}), Xl) - A_i^\alpha(\xi, u, Xl)]X\varphi d\xi \\
&+ \fint_{B_r(\xi_0)} [(A_i^\alpha(\cdot, l(\bar{\xi_0}), Xl))_{\xi_0, r} - A_i^\alpha(\xi, l(\bar{\xi_0}), Xl)]X\varphi d\xi \\
&+ \fint_{B_r(\xi_0)} B^\alpha(\xi, u, Xu)\phi^2(u-l)d\xi \\
=:& 2I_1' + I_2' + I_3' + I_4',
\end{aligned} \tag{5.2}$$

with the obvious meaning for $I_0' - I_4'$.

With respect to the terms $I_0' - I_3'$, here, we choose the same estimates as (4.2), (4.7), (4.15) and (4.21), that is,

$$I_0' \geq \nu\left[3(1+M_0^2)\right]^{\frac{p-2}{2}} \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi, \tag{5.3}$$

$$I_1' \leq 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi, \tag{5.4}$$

$$\begin{aligned}
I_2' \leq& 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + 2\varepsilon \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi \\
&+ C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi_0})|^p d\xi \right),
\end{aligned} \tag{5.5}$$

$$I_3' \leq 2\varepsilon \fint_{B_r(\xi_0)} \phi^2 |V(Xu - Xl)|^2 d\xi + 2\varepsilon \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi + C(p, L, M_0)\varepsilon^{\frac{1}{1-p}} \mathbf{V}(r). \tag{5.6}$$

Now we are in the position to get an appropriate bound for the term I_4'. By **(H4)**, elementary inequality (5.1) and Young's inequality, it yields

$$\begin{aligned}
I_4' \leq& \fint_{B_r(\xi_0)} (a|Xu|^p + b)\phi^2 |u - l| d\xi \\
\leq& a \fint_{B_r(\xi_0)} (|Xu - Xl| + |Xl|)^p \phi^2 |u - l| d\xi + b \fint_{B_r(\xi_0)} \phi^2 |u - l| d\xi
\end{aligned}$$

$$\leq a \fint_{B_r(\xi_0)} \left[(1+\epsilon)|Xu - Xl|^p + (1+K)\,|Xl|^p\right] \phi^2 |u-l| d\xi + b \fint_{B_r(\xi_0)} r\phi^2 \left|\frac{u-l}{r}\right| d\xi. \tag{5.7}$$

We denote by I'_{41} the first term of the right-hand side of (5.7). If $|Xu - Xl| \geq 1$, the following estimate holds

$$I'_{41} \leq 2a(M_0+M)(1+\epsilon) \fint_{B_r(\xi_0)} |V(Xu-Xl)|^2 \phi^2 d\xi + a\,(1+K)\,|Xl|^p \fint_{B_r(\xi_0)} \phi r \left|\frac{u-l}{r}\right| d\xi.$$

If $|Xu - Xl| \leq 1$, we have $|Xu - Xl|^p \leq |Xu - Xl|^2 + 1$. Then, it follows

$$I'_{41} \leq 2a(M_0+M)(1+\epsilon) \fint_{B_r(\xi_0)} |V(Xu-Xl)|^2 \phi^2 d\xi + a\,(1+K)\,(1+|Xl|^p) \fint_{B_r(\xi_0)} \phi r \left|\frac{u-l}{r}\right| d\xi.$$

Combining these estimates above, we have

$$I'_4 \leq 2a(M_0+M)(1+\epsilon) \fint_{B_r(\xi_0)} |V(Xu-Xl)|^2 \phi^2 d\xi + \left[a\,(1+K)\,(1+|Xl|^p) + b\right] \fint_{B_r(\xi_0)} \phi r \left|\frac{u-l}{r}\right| d\xi. \tag{5.8}$$

We denote by I_{42} the second term of the right-hand side of (5.8). If $\left|\frac{u-l}{r}\right| \geq 1$, it leads to

$$\begin{aligned} I'_{42} &\leq \left[a\,(1+K)\,(1+|Xl|^p) + b\right]^p \fint_{B_r(\xi_0)} \phi^p \left|\frac{u-l}{r}\right|^p d\xi + r^{p'} \\ &\leq \left[a\,(1+K)\,(1+|Xl|^p) + b\right]^p \fint_{B_r(\xi_0)} \phi^p \left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi + r^{p'}. \end{aligned}$$

If $\left|\frac{u-l}{r}\right| \leq 1$, it yields

$$\begin{aligned} I'_{42} &\leq \left[a\,(1+K)\,(1+|Xl|^p) + b\right]^2 \fint_{B_r(\xi_0)} \phi^2 \left|\frac{u-l}{r}\right|^2 d\xi + r^2 \\ &\leq \left[a\,(1+K)\,(1+|Xl|^p) + b\right]^2 \fint_{B_r(\xi_0)} \phi^2 \left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi + r^2. \end{aligned}$$

So, we finally arrive at

$$\begin{aligned} I'_4 \leq\, & 2a(M_0+M)(1+\epsilon) \fint_{B_r(\xi_0)} |V(Xu-Xl)|^2 \phi^2 d\xi \\ & + \left[a\,(1+K)\,(1+M_0^p) + b\right]^2 \fint_{B_r(\xi_0)} \left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi + r^2 + r^{p'}. \end{aligned} \tag{5.9}$$

Combining (5.3)-(5.6), (5.9) and (5.2), we have

$$\begin{aligned} & \left[\nu\left(3(1+M_0^2)\right)^{\frac{p-2}{2}} - 6\varepsilon - 2a(M_0+M)(1+\epsilon)\right] \fint_{B_r(\xi_0)} |V(Xu-Xl)|^2 \phi^2 d\xi \\ \leq\, & C(p,L,M_0)\left\{\varepsilon^{\frac{1}{1-p}} + 4\varepsilon + \left[a\,(1+K)\,(1+M_0^p) + b\right]^2\right\} \fint_{B_r(\xi_0)} \left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi \\ & + C(p,L,M_0)\varepsilon^{\frac{1}{1-p}}\omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi\right) + C(p,L,M_0)\varepsilon^{\frac{1}{1-p}}\mathbf{V}(r) + r^2 + r^{p'}. \end{aligned} \tag{5.10}$$

Noting that the smallness condition $2a(M+M_0)\left(3(1+M_0^2)\right)^{\frac{2-p}{2}} < \nu$, we fix the constant $\varepsilon > 0$ small sufficiently such that the coefficient $\left[\nu\left(3(1+M_0^2)\right)^{\frac{p-2}{2}} - 6\varepsilon - 2a(M+M_0)(1+\epsilon)\right] > 0$. Dividing the inequality (5.10) by the positive constant, finally we deduce

$$\fint_{B_{\frac{r}{2}}(\xi_0)} |V(Xu-Xl)|^2 d\xi \leq C_c \left[\fint_{B_r(\xi_0)} \left|V\left(\frac{u-l}{r}\right)\right|^2 d\xi + \omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi\right) + \mathbf{V}(r) + r^2 + r^{p'}\right],$$

where $C_c = C(Q,p,a,b,L,\nu,M_0,M)$. This yields the claim. □

For sake of simplicity, we motivated the form of the Caccioppoli inequalities in Lemma 5.2. We write

$$\overline{\Phi}(\xi_0, r, l) := \fint_{B_r(\xi_0)} |V(Xu - Xl)|^2 d\xi,$$

$$\overline{\Psi}(\xi_0, r, l) := \fint_{B_r(\xi_0)} \left| V\left(\frac{u-l}{r}\right)\right|^2 d\xi,$$

$$\overline{\Psi}_\star(\xi_0, r, l) :=\Psi(\xi_0, r, l) + \omega\left(\fint_{B_r(\xi_0)} |u - l(\bar{\xi}_0)|^p d\xi\right) + \mathbf{V}(r) + \left(r^2 + r^{p'}\right).$$

Lemma 5.3. *Under the assumptions of Theorem 1.1 are satisfied, $B_{2\rho}(\xi_0) \subseteq \Omega$ with $\rho \le \rho_0$ and an arbitrary horizontal function $l : \mathbb{R}^{2n} \to \mathbb{R}^N$, we define*

$$\mathcal{A} = \left(D_P A_i^\alpha(\cdot, l(\bar{\xi}_0), Xl)\right)_{\xi_0,\rho} \quad \text{and} \quad w = u - l,$$

then, w is approximately $\mathcal{A}$−harmonic in the sense that

$$\left|\fint_{B_\rho(\xi_0)} \mathcal{A}(Xw, X\varphi) d\xi\right| \le C_1' \left[\overline{\Psi}_\star(2\rho) + \mu\left(\overline{\Psi}_\star^{\frac{1}{2}}(2\rho)\right) + \mu\left(\overline{\Psi}_\star^{\frac{1}{p}}(2\rho)\right)\right] \sup_{B_\rho(\xi_0)} |X\varphi|$$

for all $\varphi \in C_0^\infty(B_\rho(\xi_0), \mathbb{R}^N)$, and the positive constant $C_1' = C(p, a, b, M_0, L, C_c)$.

Proof. The proof is similar as the case of controllable growth. Here, we just give the different estimate for the natural growth term, that is,

$$\begin{aligned} J_{23} \le& 2a(M_0 + M)(1+\epsilon) \fint_{B_\rho(\xi_0)} |V(Xu - Xl)|^2 \phi^2 d\xi \\ &+ \left[a\,(1+K)\,(1 + M_0^p) + b\right]^2 \fint_{B_\rho(\xi_0)} \left| V\left(\frac{u-l}{\rho}\right)\right|^2 d\xi + \rho^2 + \rho^{p'} \\ \le& C(p, a, b, M_0, M)\left(\overline{\Phi}(\rho) + \overline{\Psi}(\rho) + \rho^2 + \rho^{p'}\right) \\ \le& C(p, a, b, M_0, M, C_c)\left(\overline{\Phi}(\rho) + \overline{\Psi}_\star(\rho)\right), \end{aligned}$$

where we have used the bound for the natural growth term I_4' in (5.9). The rest procedure is very similar as the proof in Lemma 4.2, and we omit them. So we obtain the claim. □

Applying Lemma 5.2 and Lemma 5.3, we can establish the improvement estimate for Excess functional $\overline{\Psi}$ with the same form as Lemma 4.3, that is,

Lemma 5.4. *Suppose that the assumptions of Theorem 1.2 are satisfied and consider a ball $B_r(\xi_0) \subseteq \Omega$ with $r \le \rho_0$. For the constants $\delta = \delta(Q, N, p, L, \nu, \theta^{Q+4}) \in (0, 1]$ and $\gamma \in (0, 1]$ from the $\mathcal{A}$-harmonic approximation lemma 3.3, we let $0 < \theta \le \frac{1}{4}$ be arbitrary and also impose the following smallness conditions:*

(i). $\overline{\Psi}_\star^{\frac{1}{2}}(r) < \frac{\delta}{2}$;

(ii). $\gamma := \sqrt{\overline{\Psi}_\star(r) + \left(\frac{\delta}{2}\right)^{-2}\left[\mu\left(\overline{\Psi}_\star^{\frac{1}{2}}(r)\right) + \mu\left(\overline{\Psi}_\star^{\frac{1}{p}}(r)\right)\right]^2} \le 1.$

Then, the following excess improvement estimate holds

$$\overline{\Psi}(\xi_0, \theta r, l_{\xi_0,\theta r}) \le C_4' \theta^2 \overline{\Psi}_\star(\xi_0, r, l_{\xi_0,r}),$$

where constants C_4' depend only on $Q, N, p, a, b, \nu, \delta$ and L.

By Lemma 5.4, the iteration for the $\overline{\Psi}$-excess and the C_γ-excess can be obtained as follows,

Lemma 5.5. *Suppose that the assumptions of Theorem 1.2 are satisfied. For every $y \in (0, 1)$, there are constant $\varepsilon_\star$, $\kappa_\star$, $\rho_\star$ and $\theta \in (0, \frac{1}{8}]$, if*

$$\overline{\Psi}(\xi_0, r, l_{\xi_0,r}) < \varepsilon_\star \quad and \quad C_y(\xi_0, r) < \kappa_\star, \tag{A_0}$$

for $r \in (0, \rho_\star)$ with $B_r(\xi_0) \subset \Omega$, then,

$$\overline{\Psi}(\xi_0, \theta^k r, l_{\xi_0,\theta^k r}) < \varepsilon_\star \quad and \quad C_y(\xi_0, \theta^k r) < \kappa_\star, \tag{A_k}$$

respectively, for every $k \in \mathbb{N}$.

Proof of Theorem1.2. It is enough to use Lemma 5.5, and repeat the procedure for the proof of Theorem 1.1 in the previous Subsection 4.5.

Acknowledgement: The authors wish to thank the referees for their careful reading of my manuscript and valuable suggestions.

Funding: The research is supported by the National Natural Science Foundation of China (No.11661006), and the Science and Technology Planning Project of Jiangxi Province, China (No. GJJ190741).

References

[1] V. Bögelein, F. Duzaar, J. Habermann, C. Scheven, Partial Hölder continuity for discontinuous elliptic problems with VMO-coefficients, Proc. Lond. Math. Soc. **103** (2011), 371-404.
[2] M. Bramanti, An Invitation to Hypoelliptic Operators and Hörmander's Vector Fields, Springer, 2014.
[3] L. Capogna, D. Danielli, N. Garofalo, An embedding theorem and the Harnack inequality for nonlinear sub-elliptic equations, Comm. Partial Differential Equations **18** (1993), 1765-1794.
[4] M. Carozza, N. Fusco, G. Mingione, Partial regularity of minimizers of quasiconvex integrals with sub-quadratic growth, Ann. Mat. Pura Appl. **175** (1998), 141-164.
[5] L. Capogna, N. Garofalo, Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type, J. European Math. Society **5** (2003), 1-40.
[6] S. Chen, Z. Tan, The method of A-harmonic approximation and optimal interior partial regularity for nonlinear elliptic systems under the controllable growth condition, J. Math. Anal. Appl. **335** (2007), 20-42.
[7] G. Di Fazio, M. Fanciullo, Gradient estimates for elliptic systems in Carnot-Carathéodory spaces. Comment. Math. Univ. Caroline. **43** (2002), 605-618.
[8] F. Duzaar, A. Gastel, Nonlinear elliptic systems with Dini continuous coefficients, Arch. Math. **78** (2002), 58-73.
[9] F. Duzaar, J. F. Grotowski, Partial regularity for nonlinear elliptic systems: The method of A-harmonic approximation, Manuscripta Math. **103** (2000), 267-298.
[10] F. Duzaar, J. F. Grotowski, M. Kronz, Regularity of almost minimizers of quasi-convex variational integrals with sub-quadratic growth, Ann. Mat. Pura Appl. **184** (2005), 421-448.
[11] F. Duzaar, A. Gastel, G. Mingione, Elliptic systems, singular sets and Dini continuity, Comm. Partial Differential Equations **29** (2004), 1215-1240.
[12] F. Duzaar, G. Mingione, The p-harmonic approximation and the regularity of p-harmonic maps, Calc. Var. Partial Differention Equations **20** (2004), 235-256.
[13] F. Duzaar, G. Mingione, Regularity for degenerate elliptic problems via p-harmonic approximation, Ann. Inst. H. Poincaré Anal. Non Linèaire **21** (2004), 735-766.
[14] Y. Dong, P. Niu, Estimates in Morrey spaces and Hölder continuity for weak solutions to degenerate elliptic systems, Manuscripta Math. **138** (2012), 419-437.
[15] F. Duzaar, K. Steffen, Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals, J. Reine Angew. Math. **546** (2002), 73-138.
[16] A. Föglein, Partial regularity results for sub-elliptic systems in the Heisenberg group, Calc. Var. Partial Differential Equations **32** (2008), 25-51.
[17] M. Foss, G. Mingione, Partial continuity for elliptic problems , Ann. Inst. H. Poincaré Anal. Non Linéaire **25** (2008), 471-503.
[18] M. Giaquinta, Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems, Princeton University Press, Princeton, NJ, 1983.
[19] D. Gao, P. Niu, J. Wang, Partial regularity for degenerate sub-elliptic systems associated with Hörmander's vector fields, Nonlinear Anal. **73** (2010), 3209-3223.

[20] C. S. Goodrich, M. A. Ragusa, A. Scapellato, Partial regularity of solutions to $p(x)$-Laplacian PDEs with discontinuous coefficients, J. Differential Equations **268** (2020), 5440-5468.
[21] G. Lu, Embedding theorems on Campanato-Morrey space for vector fields on Hörmander type, Approx. Theory Appl. **14** (1998) 69-80.
[22] S. Polidoro, M. A. Ragusa, Harnack inequality for hypoelliptic ultraparabolic equations with a singular lower order term, Revista Matematica Iberoamericana **24** (2008), 1011-1046.
[23] M. A. Ragusa, A. Tachikawa, Regularity of minimizers of some variational integrals with discontinuity, Z. Anal. Anwend. **27** (2008), 469-482.
[24] A. Scapellato, New perspectives in the theory of some function spaces and their applications, AIP Conference Proceedings **1978**, 140002(2018); https://doi.org/10.1063/1.5043782.
[25] E. Shores, Regularity theory for weak solutions of systems in Carnot groups, Ph. D. Thesis, University of Arkansas. 2005.
[26] T. Kanazawa, Partial regularity for elliptic systems with VMO-coefficients, Riv. Math. Univ. Parma (N.S.) **5** (2014), 311-333.
[27] Z. Tan, Y. Wang, S. Chen, Partial regularity in the interior for discontinuous inhomogeneous elliptic system with VMO-coefficients, Ann. Mat. Pura Appl. **196** (2017), 85-105.
[28] Z. Tan, Y. Wang, S. Chen, Partial regularity up to the boundary for solutions of sub-quadratic elliptic systems, Adv. Nonlinear Anal. **7** (2018), 469-483.
[29] G. Lu, The sharp Poincaré inequality for free vector fields: an endpoint result, Revista Matematica Iberoamericana **10** (1994), 453-466.
[30] J. Wang, D. Liao, Optimal partial regularity for sub-elliptic systems with sub-quadratic growth in Carnot groups, Nonlinear Anal. **75** (2012), 2499-2519.
[31] J. Wang, D. Liao, S. Gao, Z. Yu, Optimal partial regularity for sub-elliptic systems with Dini continuous coefficients under the superquadratic natural growth, Nonlinear Anal. **114** (2015), 13-25.
[32] J. Wang, Juan J. Manfredi, Partial Hölder continuity for nonlinear sub-elliptic systems with VMO-coefficients in the Heisenberg group, Adv. Nonlinear Anal. **7** (2018), 96-114.
[33] J. Wang, S. Zhang, Q. Yang, Partial regularity for discontinuous sub-elliptic systems with sub-quadratic growth in the Heisenberg group, Nonlinear Anal. **195** (2020), 111719; https://doi.org/10.1016/j.na.2019.111719.
[34] C. Xu, C. Zuily, Higher interior regularity for quasilinear sub-elliptic systems, Calc. Var. Partial Differential Equations **5** (1997), 323-343.
[35] S. Zheng, Partial regularity for quasi-linear elliptic systems with VMO coefficients under the natural growth, Chinese Ann. Math. Ser. A **29** (2008), 49-58.
[36] S. Zheng, Z. Feng, Regularity of sub-elliptic p-harmonic systems with subcritical growth in Carnot group, J. Differential Equations **258** (2015), 2471-2494.