Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(1): 1407–1437.
DOI:10.3934/math.2025065
Received: 05 November 2024
Revised: 09 January 2025
Accepted: 13 January 2025
Published: 22 January 2025

Research article

Gradient regularity for nonlinear sub-elliptic systems with the drift term: sub-quadratic growth case

Beibei Chen, Jialin Wang* and Dongni Liao*

School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, China

* **Correspondence:** Email: wangjialin@gnnu.edu.cn, liaodongni@gnnu.edu.cn.

Abstract: This paper focuses on nonlinear sub-elliptic systems with drift terms in divergence form, under Dini continuity conditions, where the growth rate satisfies $\frac{2Q}{Q+2} < m < 2$, and Q represents the homogeneous dimension in the Heisenberg group. By generalizing the $\mathcal{A}$-harmonic approximation technique to accommodate sub-quadratic growth, we establish the C^1 regularity associated with the horizontal gradient of weak solutions away from a negligible set.

Keywords: gradient regularity; Dini continuous coefficient; Heisenberg group; drift term; sub-quadratic growth
Mathematics Subject Classification: 35B65, 35H20

1. Introduction

In this paper, we consider the following nonlinear sub-elliptic systems with the drift term Tu under sub-quadratic natural growth conditions in the Heisenberg group $\mathbb{H}^n$

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

where Ω is a bounded domain, and the horizontal gradient $X = \{X_1, \cdots, X_{2n}\}$ with the horizontal vector fields X_i ($i = 1, \cdots, 2n$) and the vertical vector fields T is defined (2.1) in the next section, $u = (u^1, \cdots, u^N) : \Omega \to \mathbb{R}^N$, $A_i^\alpha(\xi, u, Xu) : \Omega\times\mathbb{R}^N\times\mathbb{R}^{2n\times N} \to \mathbb{R}^{2n\times N}$, and $B^\alpha(\xi, u, Xu) : \Omega\times\mathbb{R}^N\times\mathbb{R}^{2n\times N} \to \mathbb{R}^N$.

As is well known, operators with drift terms possess significant importance for research and application. For instance, the Kolmogorov–Fokker–Planck operator frequently arises in transport diffusion equations within physical science, natural science, and statistical models. Following the publication of Lanconelli and Polidoro's fundamental work [25], this type of operator has garnered increasing attention. For a comprehensive understanding of the Kolmogorov–Fokker–Planck operator, the readers may refer to [26], which reviews the class of Kolmogorov operators with constant coefficients. Within the family of Kolmogorov operators, homogeneous ones occupy a central role. Indeed, any Kolmogorov operator can be approximated, in an appropriate sense, by a homogeneous operator. For more regularity results concerning operators with drift terms, the readers may refer to previous studies [5, 14, 20, 21, 24] and the references therein. In particular, Austin and Tyson [1] achieved C^∞-smoothness by the geometric analysis method for the following operator

$$L = -\frac{1}{4}\sum_{j=1}^{2n} X_j^2 \pm \sqrt{3}T \tag{1.2}$$

in the Heisenberg groups. Recently, Zhang and Niu [35] treated a quasi-linear sub-elliptic equation with drift in the Heisenberg group. For nonlinear discontinuous sub-elliptic systems with drift, Zhang and Wang [36, 37] proved the partial $C^{0,\gamma}(0 < \gamma < 1)$ Hölder regularity of weak solutions.

The findings in the study of weak solutions for sub-elliptic equations and systems without the drift term include several notable regularity results. These results are significant because they provide insights into the behavior of solutions under various conditions. For a comprehensive understanding, the readers are encouraged to consult the works of Domokos [17]; Capogna [6, 7]; Manfredi and Mingione [27]; Mingione, Zatorska-Goldstein, and Zhong [29]; Mukherjee and Zhong [28]; and Citti and Mukherjee [9] for sub-elliptic equations, as well as the studies [8, 19, 30, 33, 34] for sub-elliptic systems. Among these contributions, a particularly noteworthy development is the extension of the $\mathcal{A}$-harmonic approximation technique to noncommutative nil-potent Lie groups. This technique involves constructing approximate solutions that satisfy certain harmonic-like properties, which can then be used to deduce the regularity properties of the original solutions. By applying this method in the context of noncommutative nil-potent Lie groups, researchers have been able to establish optimal partial regularity for nonlinear sub-elliptic systems, involving different growth rates and variant structure coefficients. It is worth pointing out that the $\mathcal{A}$-harmonic approximation method was introduced by Simon [31], and developed by Duzaar and Steffen [18] in the Euclidean space, and we refer the readers to [3,4,11,15,16] and the references therein for more results concerning nonlinear elliptic and parabolic systems.

Therefore, we examine the technique of $\mathcal{A}$-harmonic approximation to achieve C^1 regularity for nonlinear sub-elliptic systems with the drift term Tu in the Heisenberg group. The primary novel aspect of this paper is our capacity to tackle the systems (1.1) that incorporate the drift term Tu, featuring a sub-quadratic growth rate, while relaxing the assumption on the principal coefficients to Dini continuity. We note that the first new challenge emerges due to the presence of the drift term Tu without any assumption of integrability. Then, we adopt a clever approach to avoid the requirement of integrability. In fact, we subtly employ the relationship

$$T = X_i X_{n+i} - X_{n+i} X_i, \quad i = 1, 2, \cdots, n,$$

and introduce a horizontal affine function l defined in (2.6) to derive a suitable estimate for the drift term Tu. In contrast to the sub-quadratic sub-elliptic systems without any drift term, as examined in [33] for the range $1 < m < 2$, our scenario requires more stringent constraints $\frac{2Q}{Q+2} < m < 2$. This is primarily attributed to the fact that the estimates stemming from the drift term Tu cannot be incorporated into the existing estimates during the formulation of Caccioppoli-type inequalities; see Lemma 3.2. The second challenge arises from the sub-quadratic growth rate, which prevents us from utilizing L^2-theory for functions in the horizontal Sobolev space $HW^{1,m}$ with $\frac{2Q}{Q+2} < m < 2$. For this reason, we choose the following excess functional

$$\Phi(\xi_0, \rho, Xl) = \fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 d\xi,$$

with $\fint_{B_\rho(\xi_0)} u(\xi)d\xi = \left|B_\rho(\xi_0)\right|_{\mathbb{H}^n}^{-1} \int_{B_\rho(\xi_0)} u(\xi)d\xi$ and $V(A) = (1 + |A|^2)^{\frac{4}{m-2}}$, and establish decay estimates for Φ by a generalization of the $\mathcal{A}$-harmonic approximation (Lemma 3.1) with the auxiliary function V in the Heisenberg groups.

Now we are in the position to introduce the following structural assumptions for the coefficients A_i^α and B^α that are essential for our analysis throughout the paper.

(H1) The leading coefficient $A_i^\alpha(\xi, u, p)$ is differentiable in p, and there exists a constant C such that

$$\left|A^\alpha_{i,p^j_\beta}(\xi, u, p)\right| \le C\left(1 + |p|^2\right)^{\frac{m-2}{2}}, \quad (\xi, u, p) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^{2n\times N}, \quad \frac{2Q}{Q+2} < m < 2, \tag{1.3}$$

where $A^\alpha_{i,p^j_\beta}(\xi, u, p) = \frac{\partial A_i^\alpha(\xi,u,p)}{\partial p^j_\beta}$.

(H2) The term $A_i^\alpha(\xi, u, p)$ satisfies the following ellipticity condition

$$A^\alpha_{i,p^j_\beta}(\xi, u, p)\eta_i^\alpha \eta_j^\beta \ge \lambda\left(1 + |p|^2\right)^{\frac{m-2}{2}} |\eta|^2, \quad \forall \eta \in \mathbb{R}^{2n\times N}, \tag{1.4}$$

where λ is a positive constant.

(H3) There exists a modulus of continuity $\mu : (0, \infty) \to [0, \infty)$ such that

$$\left|A_i^\alpha(\xi, u, p) - A_i^\alpha(\tilde{\xi}, \tilde{u}, p)\right| \le K(|u|)\mu\left((d^m(\xi, \tilde{\xi}) + |u - \tilde{u}|^m)^{\frac{1}{m}}\right)(1 + |p|)^{\frac{m}{2}}, \tag{1.5}$$

where $K(\cdot) : [0, \infty) \to [0, \infty)$ is monotonously nondecreasing. Without loss of generality, it is convenient to take $K(\cdot) \ge 1$.

(HN) (Natural growth condition) For $|u| \le M = \sup_\Omega |u|$, the nonhomogeneous term $B^\alpha(\xi, u, p)$ satisfies the following growth condition

$$|B^\alpha(\xi, u, p)| \le a|p|^m + b, \tag{1.6}$$

where the positive constants $a = a(M)$ and $b = b(M)$ possibly depend on $M > 0$.

Without loss of generality, we can assume that

(μ1) μ is nondecreasing with $\mu(0+) = 0,\ \mu(1) = 1$;

(μ2) μ is concave, and $r \longrightarrow r^{-\gamma}\mu(r)$ is nonincreasing for some exponent $\gamma \in (0, 1)$;

(μ3) Dini's condition $M(r) = \int_0^r \frac{\mu(\rho)}{\rho} d\rho < \infty$ holds for some $r > 0$.

Furthermore, (H1) implies that

$$\left|A_i^\alpha(\xi, u, p) - A_i^\alpha(\xi, u, \tilde{p})\right| \le C(L)\left(1 + |p|^2 + |\tilde{p}|^2\right)^{\frac{m-2}{2}} |p - \tilde{p}|.$$

In addition, there exists a continuously non-negative and bounded function $\omega(s, t) : [0, \infty) \times [0, \infty) \to [0, \infty)$, satisfying $\omega(s, 0) = 0$ for all s. Furthermore $\omega(s, t)$ is monotonously nondecreasing in s for a fixed t and monotonously nondecreasing in t for a fixed s such that

$$\left|A^\alpha_{i,p^j_\beta}(\xi, u, p) - A^\alpha_{i,p^j_\beta}(\xi, u, \tilde{p})\right| \le C\left(1 + |p|^2 + |\tilde{p}|^2\right)^{\frac{m-2}{2}} \omega\left(|p|, |p - \tilde{p}|^2\right). \tag{1.7}$$

By the method of $\mathcal{A}$-harmonic approximation to establish C^1 regularity, the key point is to establish a certain excess decay estimate for the excess functional Φ. In the case where $m \geq 2$, this functional is given by

$$\Phi(\xi_0, \rho, Xl) = \fint_{B_\rho(\xi_0)} \left[|Xu - Xl|^2 + |Xu - Xl|^m \right] d\xi. \tag{1.8}$$

However, in the case of the sub-quadratic $\frac{2Q}{Q+2} < m < 2$, one should establish the excess decay estimate for the following functional:

$$\Phi(\xi_0, \rho, Xl) = \fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 d\xi.$$

It is shown that if $\Phi(\xi_0, \rho, Xl)$ is small enough on a ball $B_\rho(\xi_0)$, then for some fixed $\theta \in (0, 1)$, one has the excess improvement

$$\Phi(\xi_0, \theta\rho, (Xu)_{\xi_0,\theta\rho}) \leq \theta^{2\tau} \Phi(\xi_0, \rho, (Xu)_{\xi_0,\rho}) + K^*(|u_{\xi_0,\rho}|, |(Xu)_{\xi_0,\rho}|)\mu^2(\rho^\sigma),$$

where $\sigma = \min\{(2 - m)(m - 1)/m, (m - 1)/2\}$ and $K^*(s, t) = C_7 H^{2/(m-1)^2(2-m)}(s, M + t)$ with positive constants C_7. Iteration of this result leads to the excess decay estimate, which implies the regularity result.

The main result in this paper is as follows:

Theorem 1.1. *Assume that the coefficients A_i^α and B^α satisfy (H1)–(H3) and (HN) with $(\mu 1) - (\mu 3)$. Let $u \in HW^{1,m}\left(\Omega, \mathbb{R}^N\right) \cap L^\infty\left(\Omega, \mathbb{R}^N\right)$ be a weak solution to the system (1.1) with $2a(1 + 3M)/[3^{m-2}C(M, m, n)] < \lambda$ and the constant $C(M, m, n)$ in line with Lemma 2.2, i.e., for $\forall \varphi \in C_0^\infty\left(\Omega, \mathbb{R}^N\right)$,*

$$\int_\Omega A_i^\alpha(\xi, u, Xu) \cdot X_i\varphi^\alpha d\xi - \int_\Omega X_i u \cdot X_{n+i}\varphi^\alpha d\xi + \int_\Omega X_{n+i} u \cdot X_i\varphi^\alpha d\xi = \int_\Omega B^\alpha(\xi, u, Xu) \cdot \varphi^\alpha d\xi. \tag{1.9}$$

Then, there exists an open subset $\Omega_0 \subset \Omega$, such that $u \in C^1\left(\Omega_0, \mathbb{R}^N\right)$. Moreover, $\Omega \backslash \Omega_0 = \Sigma_1 \cup \Sigma_2$ and the Haar measure $(\Omega \backslash \Omega_0) = 0$, where

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

In addition, for $\tau \in [\gamma, 1)$ and $\xi_0 \in \Omega_0$, the derivative Xu has the modulus of continuity $r \longrightarrow r^\tau + M(r)$ in the neighborhood of ξ_0.

It is worth pointing out that the Haar measure in the Heisenberg groups with the underlying manifold $\mathbb{R}^{2n+1}$ is just the Lebesgue measure in $\mathbb{R}^{2n+1}$. Our result is optimal in the sense that when $\mu(\rho) = \rho^\gamma, 0 < \gamma < 1$, we have $M(r) = \int_0^r \frac{\mu(\rho)}{\rho} d\rho = \gamma^{-1} r^\gamma$, and we obtain $C^{1,\gamma}$ optimal Hölder regularity of weak solutions in that case; see [32] by Wang, Liao, and Yu.

Remark 1.1. *When we replace the condition in (1.6) with*

$$|B^\alpha(\xi, u, p)| \le a|p|^{m-\varepsilon} + b$$

with $\varepsilon > 0$ and omit the requirement $2a(1+3M)/[3^{m-2}C(M,m,n)] < \lambda$ in Theorem 1.1, the conclusion of Theorem 1.1 remains valid.

2. Preliminaries

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

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

$$\begin{aligned} \cdot \ : \ & \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{H}^n \\ & (\xi^1, t) \cdot (\tilde{\xi}^1, \tilde{t}) \mapsto \left(\xi^1 + \tilde{\xi}^1, t + \tilde{t} + \frac{1}{2} \sum_{i=1}^{n} (x_i \tilde{y}_i - \tilde{x}_i y_i) \right), \end{aligned}$$

for all $\xi = (\xi^1, t) = (x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n, t)$, $\tilde{\xi} = (\tilde{\xi}^1, \tilde{t}) = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n, \tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_n, \tilde{t})$. Its neutral element is 0, and its inverse of $(\tilde{\xi}^1, t)$ is given by $(-\tilde{\xi}^1, -t)$. In particular, the mapping $(\xi, \tilde{\xi}) \mapsto \xi \cdot \tilde{\xi}^{-1}$ is smooth; therefore $(\mathbb{H}^n, \cdot)$ is a Lie group.

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

$$X_i = \frac{\partial}{\partial x_i} - \frac{y_i}{2}\frac{\partial}{\partial t}, \quad X_{n+i} = \frac{\partial}{\partial y_i} + \frac{x_i}{2}\frac{\partial}{\partial t}, \quad T = \frac{\partial}{\partial t}, \ i = 1, 2, \cdots, n. \tag{2.1}$$

The special structure of the commutators is

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

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

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

$$d(\tilde{\xi}, \xi) = \left\| \xi^{-1} \cdot \tilde{\xi} \right\|_{\mathbb{H}^n}. \tag{2.2}$$

The measure used on $\mathbb{H}^n$ is the Haar measure (the 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

$$|B_R(\xi_0)|_{\mathbb{H}^n} = R^{2n+2} \, |B_1(\xi_0)|_{\mathbb{H}^n} \triangleq \omega_n R^Q, \tag{2.3}$$

where the number

$$Q = 2n + 2 \tag{2.4}$$

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

Definition 2.1. *Let $\Omega \subset \mathbb{H}^n$ be an open set, where the horizontal Sobolev space $HW^{1,m}(\Omega)$ $(1 \leq m < \infty)$ is defined as:*

$$HW^{1,m}(\Omega) = \left\{ u \in L^m(\Omega) \middle| X_i u \in L^m(\Omega), i = 1, 2, \cdots, 2n \right\},$$

which is a Banach space under the norm

$$\|u\|_{HW^{1,m}(\Omega)} = \|u\|_{L^m(\Omega)} + \sum_{i=1}^{2n} \|X_i u\|_{L^m(\Omega)}, \tag{2.5}$$

and the spaces $HW_0^{1,m}(\Omega)$ is the completion of $C_0^\infty(\Omega)$ under the norm (2.5).

Definition 2.2. *(Horizontal affine function). Let $u \in L^2\left(B_\rho(\xi_0), \mathbb{R}^N\right)$, $\xi_0 \in \mathbb{R}^{2n+1}$, and denote the horizontal components as $\xi^1 = \left(x^1, \cdots, x^n, y^1, \cdots, y^n\right)$ and $\xi_0^1 = \left(x_0^1, \cdots, x_0^n, y_0^1, \cdots, y_0^n\right)$. We call*

$$l_{\xi_0,\rho}(\xi^1) = l_{\xi_0,\rho}(\xi_0^1) + Xl_{\xi_0,\rho}(\xi^1 - \xi_0^1) \tag{2.6}$$

the horizontal affine function.

If the horizontal affine function $l_{\xi_0,\rho}(\xi^1)$ is a minimizer of the functional

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

we then have

$$l_{\xi_0,\rho}(\xi_0^1) = u_{\xi_0,\rho} = \fint_{B_\rho(\xi_0)} u d\xi,$$

and

$$Xl_{\xi_0,\rho} = \frac{Q-2}{C_0 Q} \frac{Q+2}{\rho^2} \fint_{B_\rho(\xi_0)} u \otimes (\xi^1 - \xi_0^1) d\xi.$$

Lemma 2.1. *(from [2]) For every $1 < m < 2$, it holds that*

$$1 \leq \frac{\int_0^1 (1 + |p + s(\tilde{p} - p)|^2)^{\frac{m-2}{2}} \mathrm{d}s}{(1 + |p|^2 + |\tilde{p}|^2)^{\frac{m-2}{2}}} \leq \frac{8}{m-1}, \tag{2.7}$$

for any $p, \tilde{p} \in R^{2n \times N}$.

From the fact that μ is nondecreasing, we conclude that $s\mu(t) \leq s\mu(s)$ for all $0 \leq t \leq s$. We also note that $s\mu(t) \leq t\mu(s) \leq t$ for $0 \leq s \leq t$ and $0 \leq s \leq 1$ by the nonincreasing property of $r \longmapsto \frac{\mu(r)}{r}$ and $\mu(1) \leq 1$. Combining both cases, we get

$$s\mu(t) \leq s\mu(s) + t, \qquad s \in [0, 1], \ t > 0.$$

From ($\mu 2$), we deduce for $\theta \in (0, 1), t > 0$, and $j \in \mathbb{N}$ that

$$\frac{2}{\gamma}(1 - \theta^\gamma)\mu^{1/2}(\theta^{2j}t) = \int_{\theta^{2(j+1)}t}^{\theta^{2j}t} \tau^{\frac{\gamma}{2}-1} \frac{\mu^{1/2}(\theta^{2j}t)}{(\theta^{2j}t)^{\gamma/2}} d\tau \leq \int_{\theta^{2(j+1)}t}^{\theta^{2j}t} \frac{\mu^{1/2}(\tau)}{\tau} d\tau,$$

which implies

$$\sum_{j=0}^{\infty}\mu^{1/2}(\theta^{2j}t) \leq \frac{\gamma}{2(1-\theta^{\gamma})}\int_0^t \frac{\mu^{1/2}(\tau)}{\tau}d\tau := \frac{\gamma}{2(1-\theta^{\gamma})}H(t). \tag{2.8}$$

This shows particularly that $\mu(t) \leq \frac{\gamma^2}{4}H^2(t)$ for all $t \leq 0$, and $t \longmapsto t^{-\gamma}H(t)$ is also nonincreasing.

Throughout the paper, we use the functions $V = V_m, W = W_m : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by

$$V(\zeta) = \zeta/(1+|\zeta|^2)^{\frac{2-m}{4}}, \qquad W(\zeta) = \zeta/(1+|\zeta|^{2-m})^{\frac{1}{2}}$$

for each $\zeta \in \mathbb{R}^n$ and $m > 1$.

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, direct computation shows that $W^{\frac{2}{m}}(t) = t^{\frac{2}{m}}(1+t^{2-m})^{-\frac{1}{m}}$ is a convex and monotone increasing function on $[0,\infty)$ with $W^{\frac{2}{m}}(0) = 0$. Moreover, we have

$$\begin{aligned}\left|W\left(\frac{\zeta_1+\zeta_2}{2}\right)\right|^{\frac{2}{m}} &= W\left(\frac{|\zeta_1+\zeta_2|}{2}\right)^{\frac{2}{m}} \leq W\left(\frac{|\zeta_1|+|\zeta_2|}{2}\right)^{\frac{2}{m}} \\ &\leq \frac{W(|\zeta_1|)^{\frac{2}{m}} + W(|\zeta_2|)^{\frac{2}{m}}}{2} = \frac{|W(\zeta_1)|^{\frac{2}{m}} + |W(\zeta_2)|^{\frac{2}{m}}}{2},\end{aligned}$$

for any $\zeta_1, \zeta_2 \in \mathbb{R}^n$.

The following lemma includes some useful properties of the function V, as outlined in Lemma 2.1 from [22].

Lemma 2.2. *Let $m \in (1,2)$ and $V, W : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be the functions defined above. Then, the following statements hold true for any $\zeta_1, \zeta_2 \in \mathbb{R}^n$ and $t > 0$:*

(1) $\frac{1}{\sqrt{2}}\min\left(|\zeta_1|, |\zeta_1|^{\frac{m}{2}}\right) \leq |V(\zeta_1)| \leq \min\left(|\zeta_1|, |\zeta_1|^{\frac{m}{2}}\right)$;
(2) $|V(t\zeta_1)| \leq \max\left(t, t^{\frac{m}{2}}\right)|V(\zeta_1)|$;
(3) $|V(\zeta_1+\zeta_2)| \leq C(m)(|V(\zeta_1)| + |V(\zeta_2)|)$;
(4) $\frac{m}{2}|\zeta_1-\zeta_2| \leq |V(\zeta_1)-V(\zeta_2)|/(1+|\zeta_1|^2+|\zeta_2|^2)^{\frac{m-2}{4}} \leq C(m,n)|\zeta_1-\zeta_2|$;
(5) $|V(\zeta_1)-V(\zeta_2)| \leq C(m,n)|V(\zeta_1-\zeta_2)|$;
(6) $|V(\zeta_1-\zeta_2)| \leq C(m,M)|V(\zeta_1)-V(\zeta_2)|$ for all ζ_1 with $|\zeta_2| \leq M$.

The inequalities (1)–(3) also hold if V is replaced by W.

For later purposes, we state the following two simple estimates, which can be easily deduced from Lemma 2.2 (1) and (6). For $\zeta_1, \zeta_2 \in \mathbb{R}^n$, $|\zeta_2| \leq M$, and for $|\zeta_1-\zeta_2| \leq 1$,

$$|\zeta_1-\zeta_2|^2 \leq C(m,M)|V(\zeta_1)-V(\zeta_2)|^2. \tag{2.9}$$

When $|\zeta_1-\zeta_2| > 1$, it yields

$$|\zeta_1-\zeta_2|^m \leq C(m,M)|V(\zeta_1)-V(\zeta_2)|^2. \tag{2.10}$$

We introduce a Sobolev–Poincaré type inequality and a prior estimate specifically for the case of sub-quadratic growth of the Heisenberg groups. Detailed proofs for these assertions can be found in the work by Wang, Liao, and Yu [32].

Lemma 2.3. *(Sobolev–Poincaré type inequality) Let $m \in (1,2)$ and $u \in HW^{1,m}(B_\rho(\xi_0), \mathbb{R}^N)$ with $B_\rho(\xi_0) \subset \Omega$; then*

$$\left(\fint_{B_\rho(\xi_0)} \left| W\left(\frac{u - u_{\xi_0,\rho}}{\rho} \right) \right|^{\frac{2m^*}{m}} \mathrm{d}\xi \right)^{\frac{m}{2m^*}} \leq C_p \left(\fint_{B_\rho(\xi_0)} |W(Xu)|^2 \mathrm{d}\xi \right)^{\frac{1}{2}}, \tag{2.11}$$

where $m^* = \frac{mQ}{Q-m}$ is the Sobolev critical exponent of m. Furthermore, the analogous inequality is valid with W being replaced by V, as defined in (2.11). In particular, the inequality also holds if we substitute 2 for $\frac{2m^*}{m}$

Lemma 2.4. *Let $u \in HW^{1,1}(\Omega, \mathbb{R}^N)$ be a weak solution of*

$$\int_\Omega A_{i,j}^{\alpha\beta} X_j u^\beta X_i \phi^\alpha \mathrm{d}\xi = 0$$

for any $\phi \in C_0^1(\Omega, \mathbb{R}^N)$, where $A_{i,j}^{\alpha\beta}$ is a constant matrix satisfying the strong Legendre–Hadamard condition:

$$A_{i,j}^{\alpha\beta} \eta^\alpha \eta^\beta \mu_i \mu_j > c|\eta|^2 |\mu|^2, \ \ \eta \in \mathbb{R}^N, \mu \in \mathbb{R}^k.$$

Then u is smooth. $C_0 \geq 1$ exists such that for any $B_\rho(\xi_0) \subset \Omega$

$$\sup_{B_{\rho/2}(\xi_0)} (|u - u_{\xi_0,\rho}|^2 + \rho^2 |Xu|^2 + \rho^4 |X^2 u|^2) \leq C_0 \rho^2 \fint_{B_\rho(\xi_0)} |Xu|^2 \mathrm{d}\xi. \tag{2.12}$$

We will conclude this section with the following lemma from [23], which will be used to establish Caccioppoli-type inequality.

Lemma 2.5. *Let $f(t)$ be a non-negative bounded function defined for $0 \leq T_0 \leq t \leq T_1$. Suppose that for $T_0 \leq t < s \leq T_1$, we have*

$$f(t) \leq A(s-t)^{-\alpha} + B(s-t)^{-\beta} + C + \theta f(s),$$

where A, B, α, β, and θ are non-negative constants and $\theta < 1$. Then there exists a constant $\bar{C} = \bar{C}(\theta, \alpha, \beta)$ such that for every $\rho, R : T_0 \leq \rho < R \leq T_1$, we have

$$f(\rho) \leq \bar{C} \left[A(s-t)^{-\alpha} + B(s-t)^{-\beta} + C \right].$$

3. Partial C^1 regularity

In this section, we mainly prove the Caccioppoli-type inequality for weak solutions of the systems (1.1) with drift. First, we state the result of the $\mathcal{A}$-harmonic approximation lemma, specifically addressing the case of sub-quadratic growth in the Heisenberg group, as exemplified in [10] for more general Carnot groups. The proof is similar to that in the Euclidean space [13].

3.1. Caccioppoli-type inequality for sub-elliptic systems

Let $Bil(\mathbb{R}^{2n\times N})$ denote the collection of bilinear forms defined in $\mathbb{R}^{2n\times N}$, and suppose $\mathcal{A} \in Bil(\mathbb{R}^{2n\times N})$. We say that a function $h \in HW^{1,m}(\Omega, \mathbb{R}^N)$ is $\mathcal{A}$-harmonic if h satisfies:

$$\int_\Omega \mathcal{A}(Xh, X\varphi)\mathrm{d}\xi = 0, \quad \forall\varphi \in C_0^1(\Omega, \mathbb{R}^N). \tag{3.1}$$

Lemma 3.1. *Let λ and L be fixed positive numbers $1 < m < 2$, and $n, N \in \mathbb{N}$ with $n \geq 2$. If, for any given $\varepsilon > 0$, there exists $\delta = \delta(n, N, \lambda, L, \varepsilon) \in (0, 1]$ with the following properties:*

(I) For any $\mathcal{A} \in Bil(\mathbb{R}^{2n\times N})$ satisfying:

$$\mathcal{A}(v, v) \geq \lambda|v|^2, \quad A(v, \overline{v}) \leq L|v||\overline{v}|, \quad v, \overline{v} \in \mathbb{R}^{2n\times N}, \tag{3.2}$$

(II) For any $g \in HW^{1,m}(B_\rho(\xi_0), \mathbb{R}^N)$ satisfying:

$$⨍_{B_\rho(\xi_0)} | V(Xg) |^2 \,\mathrm{d}\xi \leq \Upsilon^2 \leq 1, \tag{3.3}$$

$$\left| ⨍_{B_\rho(\xi_0)} \mathcal{A}(Xg, X\varphi)\mathrm{d}\xi \right| \leq \Upsilon\delta \sup_{B_\rho(\xi_0)} |X\varphi|, \quad \forall\varphi \in C_0^1(B_\rho(\xi_0), \mathbb{R}^N). \tag{3.4}$$

There then exists an $\mathcal{A}$-harmonic function h

$$h \in H = \left\{ h \in HW^{1,m}(B_\rho(\xi_0), \mathbb{R}^N) | ⨍_{B_\rho(\xi_0)} | V(Xh) |^2 \,\mathrm{d}\xi \leq 1 \right\},$$

such that

$$⨍_{B_\rho(\xi_0)} \left| V\left(\frac{g - \Upsilon h}{\rho}\right) \right|^2 \mathrm{d}\xi \leq \Upsilon^2\varepsilon. \tag{3.5}$$

We point out that Föglein, in [19], gave another version of the $\mathcal{A}$-harmonic approximation lemma, which developed the case of quadratic growth in the Euclidean space [11] to super-quadratic growth in the Heisenberg group.

In what follows, we assume that $\rho_1(s, t) = (1 + s + t)^{-1}K(s + t)^{-1}$, and $K_1(s, t) = (1 + t)^{2m}K^4(s + t)$ for $s, t \geq 0$, and note that $\rho_1 \leq 1$ and that $s \longrightarrow \rho_1(s, t)$, $t \longrightarrow \rho_1(s, t)$ are nonincreasing functions, where $K(\cdot)$ comes from (H3).

To show Theorem 1.1, our first aim is to establish a suitable Caccioppoli-type inequality.

Lemma 3.2. *(Caccioppoli-type inequality) Let $u \in HW^{1,m}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ be a weak solution to the system in (1.1) under the conditions (H1)–(H3) and (HN) and $(\mu 1) - (\mu 3)$ with $2a(1 + 3M)/[3^{m-2}C(M, m, n)] < \lambda$. Then, for every $\xi_0 \in \Omega$, and $0 < \rho \leq \rho_1^{\frac{m}{(2-m)(m-1)}}(|u_0|, |Xl|)$, it holds that*

$$⨍_{B_{\rho/2}(\xi_0)} | V(Xu - Xl) |^2 \,\mathrm{d}\xi \leq C_c\left(⨍_{B_\rho(\xi_0)} \left| V\left(\frac{u - u_0 - Xl(\xi^1 - \xi_0^1)}{\rho}\right) \right|^2 \mathrm{d}\xi + ⨍_{B_\rho(\xi_0)} \left| \frac{u - u_0 - Xl(\xi^1 - \xi_0^1)}{\rho} \right|^2 \mathrm{d}\xi + F \right) \tag{3.6}$$

with

$$F = K(\cdot)(1+|Xl|)]^{2m/(m-1)}\mu^2(\rho^{(2-m)(m-1)/m}) + (1+2M+|Xl|\rho)^{m/(m-1)^2}\rho^{m-1} + [2a(2+|Xl|)+b]^{2/(m-1)(2-m)}\rho, \tag{3.7}$$

where we define $K(\cdot) = K(|u_0| + |Xl|)$, $\xi^1 = (x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n)$ is the horizontal component of $\xi = (x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n, t) \in \mathbb{H}^n$, and the constant $C_c = C_c(n, N, m, \lambda, M)$.

Proof. Let $\eta \in C_0^\infty(B_\rho(\xi_0))$ be a standard cut-off function satisfying $0 \le \eta \le 1, |X\eta| < \frac{c}{\rho}$, and $\eta \equiv 1$ on $B_{\rho/2}(\xi_0)$. We take $v = u(\xi) - u_0 - Xl(\xi^1 - \xi_0^1)$ and $l = u_0 + Xl(\xi^1 - \xi_0^1)$, and define the two functions

$$\varphi = \eta^2 v, \qquad \phi = (1-\eta^2)v.$$

Then, one has

$$X\varphi + X\phi = Xu - Xl,$$

and

$$|V(X\varphi)|, |V(X\phi)| \le C(m)\left(|V(Xv)| + \left|V\left(\frac{v}{\rho}\right)\right|\right). \tag{3.8}$$

Using (H2), Lemma 2.1, and the elementary inequality

$$1 + |a|^2 + |b-a|^2 \le 3(1 + |a|^2 + |b|^2), \tag{3.9}$$

we have

$$\begin{aligned}
&\fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi, u, Xl + X\varphi) - A_i^\alpha(\xi, u, Xl)]X_i\varphi^\alpha \mathrm{d}\xi \\
=&\fint_{B_\rho(\xi_0)} \left[\int_0^1 \frac{dA_i^\alpha(\xi, u, Xl + \theta X\varphi)}{d\theta}\right] X_i\varphi^\alpha \mathrm{d}\xi \\
=&\fint_{B_\rho(\xi_0)} \int_0^1 \frac{\partial A_i^\alpha(\xi, u, Xl + \theta X\varphi)}{\partial p_j^\beta} d\theta X_j\varphi^\beta X_i\varphi^\alpha \mathrm{d}\xi \\
\ge&\lambda \fint_{B_\rho(\xi_0)} \int_0^1 [1 + |Xl + \theta((X\varphi + Xl) - Xl)|^2]^{\frac{m-2}{2}} d\theta |X\varphi|^2 \mathrm{d}\xi \\
\ge&\lambda \fint_{B_\rho(\xi_0)} (1 + |Xl|^2 + |X\varphi - Xl|^2)^{\frac{m-2}{2}} |X\varphi|^2 \mathrm{d}\xi \\
\ge&3^{\frac{m-2}{2}}\lambda \fint_{B_\rho(\xi_0)} (1 + |Xl|^2 + |X\varphi|^2)^{\frac{m-2}{2}} |X\varphi|^2 \mathrm{d}\xi.
\end{aligned} \tag{3.10}$$

From (3.10), it follows that

$$3^{\frac{m-2}{2}}\lambda \fint_{B_\rho(\xi_0)} (1 + |Xl|^2 + |X\varphi|^2)^{\frac{m-2}{2}} |X\varphi|^2 \mathrm{d}\xi$$

$$
\begin{aligned}
&\le \fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi,u,Xl+X\varphi) - A_i^\alpha(\xi,u,Xl)]X_i\varphi^\alpha \mathrm{d}\xi\\
&= \fint_{B_\rho(\xi_0)} A_i^\alpha(\xi,u,Xu)X_i\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} A_i^\alpha(\xi,u,Xl)X_i\varphi^\alpha \mathrm{d}\xi\\
&\quad + \fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi,u,Xl+X\varphi) - A_i^\alpha(\xi,u,Xu)]X_i\varphi^\alpha \mathrm{d}\xi\\
&\le \fint_{B_\rho(\xi_0)} X_iuX_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i}uX_i\varphi^\alpha \mathrm{d}\xi\\
&\quad + \fint_{B_\rho(\xi_0)} B_\alpha(\xi,u,Xu)\varphi^\alpha \mathrm{d}\xi\\
&\quad - \fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi,u,Xl) - A_i^\alpha(\xi,u_0+Xl(\xi^1-\xi_0^1),Xl)]X_i\varphi^\alpha \mathrm{d}\xi\\
&\quad - \fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi,u_0+Xl(\xi^1-\xi_0^1),Xl) - A_i^\alpha(\xi_0,u_0,Xl)]X_i\varphi^\alpha \mathrm{d}\xi\\
&\quad - \fint_{B_\rho(\xi_0)} \int_0^1 \frac{\partial A_i^\alpha(\xi,u,Xu-\theta(Xu-X\varphi-Xl))}{\partial p_j^\beta} d\theta X_j\phi^\beta X_i\varphi^\alpha \mathrm{d}\xi\\
&= I + II + III + IV + V.
\end{aligned}
\tag{3.11}
$$

Noting that $A_i^\alpha(\xi_0,u_0,Xl)$ is a constant, we have

$$
\fint_{B_\rho(\xi_0)} A_i^\alpha(\xi_0,u_0,Xl)X\varphi^\alpha \mathrm{d}\xi = 0.
$$

By the condition $\varphi = v$ on $B_{\rho/2}(\xi_0)$, the elementary inequality $1+|a|^2+|b-a|^2 \le 3(1+|a|^2+|b|^2)$, and the fact that $\frac{m-2}{2} < 0$ holds true for $\frac{2Q}{Q+2} < m < 2$, the left-hand side of (3.11) can be estimated by

$$
\begin{aligned}
&3^{\frac{m-2}{2}}\lambda \fint_{B_{\rho/2}(\xi_0)} (1+|Xl|^2+|X\varphi|^2)^{\frac{m-2}{2}}|X\varphi|^2 \mathrm{d}\xi\\
=&3^{\frac{m-2}{2}}\lambda \fint_{B_{\rho/2}(\xi_0)} (1+|Xl|^2+|Xv|^2)^{\frac{m-2}{2}}|Xv|^2 \mathrm{d}\xi\\
=&3^{\frac{m-2}{2}}\lambda \fint_{B_{\rho/2}(\xi_0)} (1+|Xl|^2+|Xu-Xl|^2)^{\frac{m-2}{2}}|Xu-Xl|^2 \mathrm{d}\xi\\
\ge&3^{\frac{m-2}{2}}\lambda \fint_{B_{\rho/2}(\xi_0)} \left[3(1+|Xl|^2+|Xu|^2)\right]^{\frac{m-2}{2}} |Xu-Xl|^2 \mathrm{d}\xi\\
=&3^{m-2}\lambda \fint_{B_{\rho/2}(\xi_0)} \left[(1+|Xl|^2+|Xu|^2)^{\frac{m-2}{4}}|Xu-Xl|\right]^2 \mathrm{d}\xi\\
\ge&3^{m-2}\lambda C(m,n) \fint_{B_{\rho/2}(\xi_0)} |V(Xu)-V(Xl)|^2 \mathrm{d}\xi\\
\ge&3^{m-2}\lambda C(M,m,n) \fint_{B_{\rho/2}(\xi_0)} |V(Xu-Xl)|^2 \mathrm{d}\xi\\
=&3^{m-2}\lambda C(M,m,n) \fint_{B_{\rho/2}(\xi_0)} |V(Xv)|^2 \mathrm{d}\xi,
\end{aligned}
\tag{3.12}
$$

where we have applied the fact that $Xv = Xu - Xl$, Lemma 2.2 (4) in the third inequality from the end, Lemma 2.2 (6) in the penultimate inequality.

We are in the position to estimate the term I. By the fact that $Tl = X_iX_{n+i}l - X_{n+i}X_il = 0$ and the condition $|T\eta| \le \frac{C}{\rho^2}$, it leads to

$$\begin{aligned}
I =& \fint_{B_\rho(\xi_0)} X_iu \cdot X_{n+i}\left(\eta^2(u-l)\right)d\xi - \fint_{B_\rho(\xi_0)} X_{n+i}u \cdot X_i\left(\eta^2(u-l)\right)d\xi \\
=& \fint_{B_\rho(\xi_0)} X_i(u-l) \cdot X_{n+i}\left(\eta^2(u-l)\right)d\xi + \fint_{B_\rho(\xi_0)} X_il \cdot X_{n+i}\left(\eta^2(u-l)\right)d\xi \\
&- \fint_{B_\rho(\xi_0)} X_{n+i}(u-l) \cdot X_i\left(\eta^2(u-l)\right)d\xi - \fint_{B_\rho(\xi_0)} X_{n+i}l \cdot X_i\left(\eta^2(u-l)\right)d\xi \\
=& \fint_{B_\rho(\xi_0)} \eta^2X_i(u-l) \cdot X_{n+i}(u-l)d\xi + \fint_{B_\rho(\xi_0)} 2\eta X_{n+i}\eta X_i\left((u-l)^2\right)d\xi \\
&- \fint_{B_\rho(\xi_0)} X_{n+i}X_il \cdot \left(\eta^2(u-l)\right)d\xi - \fint_{B_\rho(\xi_0)} \eta^2X_{n+i}(u-l) \cdot X_i(u-l)d\xi \\
&- \fint_{B_\rho(\xi_0)} 2\eta X_i\eta X_{n+i}\left((u-l)^2\right)d\xi + \fint_{B_\rho(\xi_0)} X_iX_{n+i}l \cdot \left(\eta^2(u-l)\right)d\xi \\
\le& \fint_{B_\rho(\xi_0)} \eta X_{n+i}\eta X_i\left((u-l)^2\right)d\xi - \fint_{B_\rho(\xi_0)} \eta X_i\eta X_{n+i}\left((u-l)^2\right)d\xi \\
\le& - \fint_{B_\rho(\xi_0)} \eta X_iX_{n+i}\eta \cdot (u-l)^2d\xi + \fint_{B_\rho(\xi_0)} \eta X_{n+i}X_i\eta \cdot (u-l)^2d\xi \\
\le& - \fint_{B_\rho(\xi_0)} \eta T\eta \cdot (u-l)^2\mathrm{d}\xi \\
\le& C\fint_{B_\rho(\xi_0)} \left|\frac{u-l}{\rho}\right|^2 \mathrm{d}\xi.
\end{aligned} \tag{3.13}$$

Applying Hölder's inequality, we infer that

$$\begin{aligned}
II =& \fint_{B_\rho(\xi_0)} B_\alpha(\xi, u, Xu)\varphi^\alpha\mathrm{d}\xi \\
\le& \fint_{B_\rho(\xi_0)} (a|Xu|^m + b)|v|\eta^2\mathrm{d}\xi \\
\le& \fint_{B_\rho(\xi_0)} (a|Xu|^m|v|\eta^2 + b|v|\eta^2)\mathrm{d}\xi \\
\le& \fint_{B_\rho(\xi_0)} \left[\left(a(1+\mu)|Xu - Xl|^m + \left(1 + \frac{1}{\mu}\right)|Xl|^m\right)|v|\eta^2 + b\rho\eta^2\left|\frac{v}{\rho}\right|\right]\mathrm{d}\xi.
\end{aligned} \tag{3.14}$$

To obtain a suitable estimate for II, we need to split the domain $B_\rho(\xi_0)$ into four parts: $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|Xu - Xl| \le 1\}$, $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|Xu - Xl| > 1\}$, $B_\rho(\xi_0) \cap \{|v/\rho| \le 1\} \cap \{|Xu - Xl > 1\}$, and $B_\rho(\xi_0) \cap \{|v/\rho| \le 1\} \cap \{|Xu - Xl| \le 1\}$. We then use Young's inequality, and note that $|v| = |u - u_0 - Xl(\xi^1 - \xi_0^1)| \le 2M + |Xl|\rho$ on $B_\rho(\xi_0)$ to have the following estimates.

Case 1: For $B_\rho(\xi_0) \cap \{|Xu - Xl| > 1\} \cap \{|v/\rho| \le 1\}$, it follows that

$$\begin{aligned}&\left(a(1+\mu)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\right)|v|\eta^2 + (b\rho\eta^2)\left|\frac{v}{\rho}\right|\\ \le& a(1+\mu)(2M+Xl\rho)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\eta^2\rho\left|\frac{v}{\rho}\right| + b\rho\eta^2\\ \le& aC(m,M)(1+\mu)(2M+Xl\rho)|V(Xv)|^2 + a\left(1+\frac{1}{\mu}\right)|Xl|^m\rho + b\rho.\end{aligned}$$

Case 2: For the set $B_\rho(\xi_0) \cap \{|Xu - Xl| > 1\} \cap \{|v/\rho| > 1\}$, we have

$$\begin{aligned}&\left(a(1+\mu)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\right)|v|\eta^2 + (b\rho\eta^2)\left|\frac{v}{\rho}\right|\\ \le& a(1+\mu)(2M+Xl\rho)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\eta^2\rho\left|\frac{v}{\rho}\right| + \varepsilon(b\rho\eta^2)^{\frac{m}{m-1}} + C(\varepsilon)\left|\frac{v}{\rho}\right|^m\\ \le& aC(m,M)(1+\mu)(2M+Xl\rho)|V(Xv)|^2 + \varepsilon(b\rho)^{\frac{m}{m-1}}\\ &+ \varepsilon\left[a\left(1+\frac{1}{\mu}\right)|Xl|^m\rho\right]^{\frac{m}{m-1}} + C(\varepsilon,m,M)\left|V\left(\frac{v}{\rho}\right)\right|^2.\end{aligned}$$

Case 3: For $B_\rho(\xi_0) \cap \{|Xu - Xl| \le 1\} \cap \{|v/\rho| \le 1\}$, we get

$$\begin{aligned}&\left(a(1+\mu)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\right)|v|\eta^2 + (b\rho\eta^2)\left|\frac{v}{\rho}\right|\\ \le& \varepsilon|Xu-Xl|^2 + C(\varepsilon)[a(1+\mu)\rho]^{2/(2-m)} + a\left(1+\frac{1}{\mu}\right)|Xl|^m\eta^2\rho + b\rho\eta^2\\ \le& \varepsilon C(m,M)|V(Xv)|^2 + C(\varepsilon)[a(1+\mu)\rho]^{2/(2-m)} + a\left(1+\frac{1}{\mu}\right)|Xl|^m\rho + b\rho.\end{aligned}$$

Case 4: For the case where $B_\rho(\xi_0) \cap \{|Xu - Xl| \le 1\} \cap \{|v/\rho| > 1\}$, one obtains

$$\begin{aligned}&\left(a(1+\mu)|Xu-Xl|^m + a\left(1+\frac{1}{\mu}\right)|Xl|^m\right)|v|\eta^2 + (b\rho\eta^2)\left|\frac{v}{\rho}\right|\\ =& a(1+\mu)|Xu-Xl|^m|v|\eta^2\rho^{\frac{m(2-m)}{2}}\rho^{\frac{m(m-2)}{2}} + a\left(1+\frac{1}{\mu}\right)|Xl|^m|v|\eta^2 + (b\rho\eta^2)\left|\frac{v}{\rho}\right|\\ \le& \varepsilon\rho^{2-m}|Xu-Xl|^2 + C(\varepsilon)(a(1+\mu))^{\frac{2}{2-m}}(2M+Xl\rho)^{\frac{2-2m+m^2}{2-m}}\left|\frac{v}{\rho}\right|^m\\ &+ \varepsilon\left[a\left(1+\frac{1}{\mu}\right)|Xl|^m\rho\eta^2\right]^{\frac{m}{m-1}} + \varepsilon(b\rho\eta^2)^{\frac{m}{m-1}} + C(\varepsilon)\left|\frac{v}{\rho}\right|^m\\ \le& \varepsilon C(m,M)|V(Xv)|^2 + C(\varepsilon,m,M)(a(1+\mu))^{\frac{2}{2-m}}(2M+Xl\rho)^{\frac{2-2m+m^2}{2-m}}\left|V\left(\frac{v}{\rho}\right)\right|^2\\ &+ \varepsilon\left[a\left(1+\frac{1}{\mu}\right)|Xl|^m\rho\right]^{\frac{m}{m-1}} + \varepsilon(b\rho)^{\frac{m}{m-1}} + C(\varepsilon,m,M)\left|V\left(\frac{v}{\rho}\right)\right|^2.\end{aligned}$$

Combining these estimates in II'', we have

$$II \leq aC(m,M)(1+\mu)(1+3M) \fint_{B_\rho(\xi_0)} |V(Xv)|^2 d\xi + C(\varepsilon,\mu,m,M,a) \int_{B_\rho(\xi_0)} \left|V\left(\frac{v}{\rho}\right)\right|^2 d\xi$$
$$+ C \max\left\{\left[a\left(1+\frac{1}{\mu}\right)\right]^{m/(m-1)} |Xl|^m \left(1+|Xl|^{m/(m-1)}\right) + [a(1+\mu)]^{\frac{2}{2-m}} + b^{m/(m-1)}\right\}\rho^2, \tag{3.15}$$

where we have used $|Xl| \leq M+1$.

The condition (H3) yields the following (note that $m-1 < m/2$):

$$III = \fint_{B_\rho(\xi_0)} [A_i^\alpha(\xi,u,Xl) - A_i^\alpha(\xi,u_0+Xl(\xi^1-\xi_0^1),Xl)]X_i\varphi^\alpha d\xi$$
$$\leq \fint_{B_\rho(\xi_0)} K(\cdot)(1+|Xl|)^{\frac{m}{2}}\mu(|v|)|X\varphi| d\xi, \tag{3.16}$$

where we have used the inequality $s\mu(t) \leq s\mu(s) + t$ for $s \in [0,1]$ and $t > 0$.

To obtain a suitable estimate for III, we need to split the domain $B_\rho(\xi_0)$ into four parts: $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|X\varphi| \leq 1\}$, $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|X\varphi| > 1\}$, $B_\rho(\xi_0) \cap \{|v/\rho| \leq 1\} \cap \{|X\varphi > 1\}$, and $B_\rho(\xi_0) \cap \{|v/\rho| \leq 1\} \cap \{|X\varphi| \leq 1\}$. We use Young's inequality, (2.9), and (2.10) repeatedly.

Case 1: For the set $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|X\varphi| > 1\}$,

$$\left[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m}) + \frac{|v|}{\rho^{(2-m)(m-1)/m}}\right]|X\varphi|$$
$$\leq 2\varepsilon|X\varphi|^m + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)} + \varepsilon^{-1}|v|^{m/(m-1)}\rho^{m-2}$$
$$\leq 2\varepsilon|X\varphi|^m + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)} + \varepsilon^{-1}\left|\frac{v}{\rho}\right||v|^{1/(m-1)}\rho^{m-1}$$
$$\leq 2\varepsilon|X\varphi|^m + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)} + \varepsilon^{-1}\left|\frac{v}{\rho}\right|^m$$
$$+ \varepsilon^{-1}|v|^{m/(m-1)^2}\rho^m$$
$$\leq 2\varepsilon C(m,M)|V(X\varphi)|^2 + \varepsilon^{-1}C(m,M)\left|V\left(\frac{v}{\rho}\right)\right|^2$$
$$+ \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)} + \varepsilon^{-1}|2M+Xl\rho|^{m/(m-1)^2}\rho^m.$$

Case 2: For $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|X\varphi| \leq 1\}$,

$$\left[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m}) + \frac{|v|}{\rho^{(2-m)(m-1)/m}}\right]|X\varphi|$$
$$\leq 2\varepsilon|X\varphi|^2 + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^2 + \varepsilon^{-1}\left|\frac{v}{\rho}\right||v|\rho^{(m-2)\frac{2(m-1)}{m}} \cdot \rho$$
$$\leq 2\varepsilon|X\varphi|^2 + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^2 + \varepsilon^{-1}\left|\frac{v}{\rho}\right||v|\rho^{m-2} \cdot \rho$$
$$\leq 2\varepsilon|X\varphi|^2 + \varepsilon^{-1}[K^2(\cdot)(1+|Xl|)^m\mu(\rho^{(2-m)(m-1)/m})]^2 + \varepsilon^{-1}\left|\frac{v}{\rho}\right|^m$$

$$
\begin{aligned}
&+\varepsilon^{-1}|v|^{m/(m-1)}\rho^{m}\\
\leq&2\varepsilon C(m,M)|V(X\varphi)|^{2}+\varepsilon^{-1}C(m,M)\left|V\left(\frac{v}{\rho}\right)\right|^{2}\\
&+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{2}+\varepsilon^{-1}|2M+Xl\rho|^{m/(m-1)}\rho^{m},
\end{aligned}
$$

where we have used the facts that $2(m-1)/m<1$.
Case 3: For $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|X\varphi|>1\}$, observing that $m/(m-1)>2$, one has

$$
\begin{aligned}
&\left[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})+\frac{|v|}{\rho^{(2-m)(m-1)/m}}\right]|X\varphi|\\
\leq&2\varepsilon|X\varphi|^{m}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)}+\varepsilon^{-1}|v|^{m/(m-1)}\rho^{m-2}\\
\leq&2\varepsilon|X\varphi|^{m}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)}+\varepsilon^{-1}\left|\frac{v}{\rho}\right||v|^{1/(m-1)}\rho^{m-1}\\
\leq&2\varepsilon C(m,M)|V(X\varphi)|^{2}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{m/(m-1)}\\
&+\varepsilon^{-1}|2M+Xl\rho|^{1/(m-1)}\rho^{m-1}.
\end{aligned}
$$

Case 4: For the case of $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|X\varphi|\leq 1\}$,

$$
\begin{aligned}
&\left[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})+\frac{|v|}{\rho^{(2-m)(m-1)/m}}\right]|X\varphi|\\
\leq&2\varepsilon|X\varphi|^{2}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{2}+\varepsilon^{-1}\left|\frac{v}{\rho}\right|^{2}\rho^{(m-2)\frac{2(m-1)}{m}}\cdot\rho^{2}\\
\leq&2\varepsilon|X\varphi|^{2}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{2}+\varepsilon^{-1}\rho^{m-2}\cdot\rho^{2}\\
\leq&2\varepsilon C(m,M)|V(X\varphi)|^{2}+\varepsilon^{-1}[K^{2}(\cdot)(1+|Xl|)^{m}\mu(\rho^{(2-m)(m-1)/m})]^{2}+\varepsilon^{-1}\rho^{m},
\end{aligned}
$$

where we have used the fact that $2(m-1)/m<1$.

Combining these estimations with (3.15), we get

$$
\begin{aligned}
III\leq&\varepsilon C(m,M)\fint_{B_\rho(\xi_0)}|V(Xv)|^{2}\mathrm{d}\xi+C(\varepsilon,m,M)\fint_{B_\rho(\xi_0)}\left|V\left(\frac{v}{\rho}\right)\right|^{2}\mathrm{d}\xi\\
&+\varepsilon^{-1}K^{2m/(m-1)}(\cdot)(1+|Xl|)^{m^{2}/(m-1)}\mu^{2}(\rho^{(2-m)(m-1)/m})\\
&+\varepsilon^{-1}(1+2M+Xl\rho)^{m/(m-1)^{2}}\rho^{m-1}.
\end{aligned}\tag{3.17}
$$

We apply (H3) to get (noting that $m/(m-1)<2$)

$$
\begin{aligned}
IV=&\fint_{B_\rho(\xi_0)}[A_i^\alpha(\xi,u_0+Xl(\xi^1-\xi_0^1),Xl)-A_i^\alpha(\xi_0,u_0,Xl)]X_i\varphi^\alpha\mathrm{d}\xi\\
\leq&\fint_{B_\rho(\xi_0)}K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)|X\varphi|\mathrm{d}\xi.
\end{aligned}\tag{3.18}
$$

To obtain a suitable estimate for IV, we need to split the domain $B_\rho(\xi_0)$ into four parts: $B_\rho(\xi_0)\cap\{|v/\rho|>1\}\cap\{|X\varphi|\leq 1\}$, $B_\rho(\xi_0)\cap\{|v/\rho|>1\}\cap\{|X\varphi|>1\}$, $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|X\varphi>1\}$, and $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|X\varphi|\leq 1\}$. We use Young's inequality, (2.9), and (2.10) repeatedly.

Case 1: $B_\rho(\xi_0)\cap\{|v/\rho|\le 1\}\cap\{|X\varphi|>1\}$ and $B_\rho(\xi_0)\cap\{|v/\rho|>1\}\cap\{|X\varphi|>1\}$, observing that $m/(m-1)>2$,

$$\begin{aligned}
&K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)|X\varphi|\mathrm{d}\xi\\
\le&\varepsilon|X\varphi|^m+C(\varepsilon)[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^{\frac{m}{m-1}}\\
\le&\varepsilon C(m,M)|V(X\varphi)|^2+C(\varepsilon)[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^{\frac{m}{m-1}}\\
\le&\varepsilon C(m,M)|V(Xv)|^2+\varepsilon C(m,M)\left|V\left(\frac{v}{\rho}\right)\right|^2+C(\varepsilon)[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^{\frac{m}{m-1}}.
\end{aligned}$$

Case 2: $B_\rho(\xi_0)\cap\{|v/\rho|\le 1\}\cap\{|X\varphi|\le 1\}$ and $B_\rho(\xi_0)\cap\{|v/\rho|>1\}\cap\{|X\varphi|\le 1\}$,

$$\begin{aligned}
&K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)|X\varphi|\mathrm{d}\xi\\
\le&\varepsilon|X\varphi|^2+\varepsilon^{-1}[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^2\\
\le&\varepsilon C(m,M)|V(X\varphi)|^2+\varepsilon^{-1}[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^2\\
\le&\varepsilon C(m,M)|V(Xv)|^2+\varepsilon C(m,M)\left|V\left(\frac{v}{\rho}\right)\right|^2+\varepsilon^{-1}[K(\cdot)(1+|Xl|)^{\frac{m+2}{2}}\mu(\rho)]^2.
\end{aligned}$$

Combining these estimations with (3.18), we get

$$\begin{aligned}
IV\le&\varepsilon C(m,M)⨍_{B_\rho(\xi_0)}|V(Xv)|^2\mathrm{d}\xi+\varepsilon C(m,M)⨍_{B_\rho(\xi_0)}\left|V\left(\frac{v}{\rho}\right)\right|^2\mathrm{d}\xi\\
&+C(\varepsilon)K(\cdot)^{\frac{m}{m-1}}(1+|Xl|)^{\frac{m(m+2)}{2(m-1)}}\mu^2(\rho),
\end{aligned}\tag{3.19}$$

where we used $\frac{m}{m-1}>2$, $\frac{m(m+2)}{2(m-1)}>m+2$.

By (H1), Lemma 2.1, and (3.9), it holds that

$$\begin{aligned}
V=&⨍_{B_\rho(\xi_0)}\int_0^1\frac{\partial A_i^\alpha(\xi,u,Xu-\theta(Xu-X\varphi-Xl))}{\partial p_j^\beta}d\theta X_j\phi^\beta X_i\varphi^\alpha\mathrm{d}\xi\\
\le& C⨍_{B_\rho(\xi_0)}\left[\int_0^1(1+|Xu+\theta(Xu-X\varphi-Xl)|^2)^{\frac{m-2}{2}}d\theta\right]|X\phi||X\varphi|\mathrm{d}\xi\\
\le& C⨍_{B_\rho(\xi_0)}\left[\int_0^1(1+|Xu+\theta[(Xu-X\phi)-Xu]|^2)^{\frac{m-2}{2}}d\theta\right]|X\phi||X\varphi|\mathrm{d}\xi\\
\le&\frac{8C}{m-1}⨍_{B_\rho(\xi_0)}(1+|Xu|^2+|Xu-X\phi|^2)^{\frac{m-2}{2}}|X\phi||X\varphi|\mathrm{d}\xi\\
\le&\frac{8C}{m-1}⨍_{B_\rho(\xi_0)}(1+|Xu|^2+|X\phi|^2)^{\frac{m-2}{2}}|X\phi||X\varphi|\mathrm{d}\xi\\
\le&\frac{8C}{m-1}⨍_{B_\rho(\xi_0)}(1+|X\phi|^2)^{\frac{m-2}{2}}|X\phi||X\varphi|\mathrm{d}\xi.
\end{aligned}\tag{3.20}$$

Noting that $-1/2<\frac{m-2}{2}<0$, and $\phi=(1-\eta)v=0$ due to $\eta=1$ on $B_{\rho/2}(\xi_0)$, we split the domain $B_\rho(\xi_0)$ into four parts: $B_\rho(\xi_0)\cap\{|X\phi|>1\}\cap\{X\varphi|>1\}$, $B_\rho(\xi_0)\cap\{|X\phi|\le 1\}\cap\{X\varphi|\le 1\}$, $B_\rho(\xi_0)\cap\{|X\phi|>$

$1\} \cap \{|X\varphi| \le 1\}$, and $B_\rho(\xi_0) \cap \{|X\phi| \le 1\} \cap \{X\varphi| > 1\}$. Thus by Young's inequality and the estimations in (2.9) and (2.10), there is

$$V \le \frac{C(m,M)}{m-1} \fint_{B_\rho(\xi_0)\setminus B_{\rho/2}(\xi_0)} |V(Xu - Xl)|^2 \mathrm{d}\xi + C(M,m,n) \fint_{B_\rho(\xi_0)} \left| V\left(\frac{v}{\rho}\right)\right|^2 \mathrm{d}\xi. \tag{3.21}$$

Substituting (3.12), (3.13), (3.15), (3.17), (3.19) and (3.21) into (3.11), we finally arrive at

$$\begin{aligned}
&\left[3^{m-2}\lambda C(M,m,n) + \frac{C(m,M)}{m-1}\right] \fint_{B_{\rho/2}(\xi_0)} |V(Xv)|^2 \mathrm{d}\xi \\
\le &\left[\frac{C(m,M)}{m-1} + (3\varepsilon + 2a(1+3M))C(m,M)\right] \fint_{B_\rho(\xi_0)} |V(Xv)|^2 \mathrm{d}\xi \\
&+ C(\varepsilon,a,n,m,M) \fint_{B_\rho(\xi_0)} \left| V\left(\frac{v}{\rho}\right)\right|^2 \mathrm{d}\xi + C(\varepsilon,a,n,m,M) \fint_{B_\rho(\xi_0)} \left|\frac{v}{\rho}\right|^2 \mathrm{d}\xi \\
&+ C(\varepsilon)[K(\cdot)(1+|Xl|)]^{2m/m-1}\mu^2(\rho^{(2-m)(m-1)/m}) \\
&+ (1+2M+|Xl|\rho)^{m/(m-1)^2}\rho^{m-1} + [2a(2+|Xl|)+b]^{2/(m-1)(2-m)}\rho,
\end{aligned}$$

where we have used $2m/(m-1) > m(m+2)/2(m-1) > m^2/(m-1)$ and the nondecreasing property of μ.

We take $\varepsilon = [3^{m-2}\lambda C(M,m,n) - 2a(1+3M)]/6$ with the assumption $\lambda > 2a(1+3M)/[3^{m-2}C(M,m,n)]$. Filling the gaps with $\theta = \frac{\frac{C(m,M)}{m-1} + [3\varepsilon + 2a(1+3M)C(M,m,n)]}{3^{m-2}\lambda C(M,m,n) + \frac{C(m,M)}{m-1}} < 1$ in Lemma 2.5 yields

$$\fint_{B_{\rho/2}(\xi_0)} |V(Xu - Xl)|^2 \mathrm{d}\xi \le C\left(\fint_{B_\rho(\xi_0)} \left| V\left(\frac{v}{\rho}\right)\right|^2 \mathrm{d}\xi + \fint_{B_\rho(\xi_0)} \left|\frac{v}{\rho}\right|^2 \mathrm{d}\xi + F\right) + \theta \fint_{B_\rho(\xi_0)} |V(Xu - Xl)|^2 \mathrm{d}\xi,$$

where

$$\begin{aligned}
F =&[K(\cdot)(1+|Xl|)]^{2m/(m-1)}\mu^2(\rho^{(2-m)(m-1)/m}) + (1+2M+|Xl|\rho)^{m/(m-1)^2}\rho^{m-1} \\
&+ [2a(2+|Xl|)+b]^{2/(m-1)(2-m)}\rho.
\end{aligned} \tag{3.22}$$

The proof is completed by noting that $[m(r-1)/r(m-1) - 1]Q = m/(m-1)$ and $\rho^{m/(m-1)} \le \rho^{2(2-m)(m-1)/m} \le \mu^2(\rho^{(2-m)(m-1)/m})$. □

3.2. Approximate $\mathcal{A}$-harmonicity by linearization

In this section, we provide a linearization strategy for non-linear sub-elliptic systems (1.1). Later on, this will be the starting point for the application of the $\mathcal{A}$-harmonic approximation lemma.

Lemma 3.3. *We claim that if $\rho \le \rho_1^{\frac{m}{(2-m)(m-1)}}(|u_0|, |Xl|)$ and $\varphi \in C_0^\infty(B_\rho(\xi_0), \mathbb{R}^N)$ with $\sup_{B_\rho(\xi_0)} |X\varphi| \le 1$, then there exist some constants $C_1 = C_1(m, M, C_p, K) > 1$ such that*

$$\fint_{B_\rho(\xi_0)} A^\alpha_{i,p^j_\beta}(\xi_0, u_0, Xl)(Xu - Xl)X\varphi^\alpha \mathrm{d}\xi$$

$$\leq C_1 \sup_{B_\rho(\xi_0)} |X\varphi| \Bigg[\omega\left(|Xl|, \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl)\right) \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl) + \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl) + \Phi(\xi_0,\rho,Xl) + \Phi^{\frac{1}{m}}(\xi_0,\rho,Xl) + \mu(\sqrt{\rho}) F(|u_0|,|Xl|) \Bigg], \tag{3.23}$$

where we assume that $F(s,t) = K^{4/(2-m)}(s+t)(2+t)^2 + a(1+t^m) + b$.

Proof. A straightforward computation yields

$$\begin{aligned}
&\fint_{B_\rho(\xi_0)} \left[\int_0^1 A^\alpha_{i,p^j_\beta}(\xi_0,u_0,\theta Xu + (1-\theta)Xl)(Xu - Xl)\mathrm{d}\theta \right] X\varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} \left[\int_0^1 \frac{d}{d\theta} A^\alpha_i(\xi_0,u_0,\theta Xu + (1-\theta)Xl)\mathrm{d}\theta \right] X\varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} [A^\alpha_i(\xi_0,u_0,Xu) - A^\alpha_i(\xi,u,Xu)] X\varphi^\alpha \mathrm{d}\xi + \fint_{B_\rho(\xi_0)} B^\alpha(\xi,u,Xu)\varphi^\alpha \mathrm{d}\xi \\
&+ \fint_{B_\rho(\xi_0)} X_i u X_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i} u X_i \varphi^\alpha \mathrm{d}\xi.
\end{aligned} \tag{3.24}$$

Then, we have

$$\begin{aligned}
&\fint_{B_\rho(\xi_0)} A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl)(Xu - Xl) X\varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} \left[\int_0^1 A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl)\mathrm{d}\theta (Xu - Xl) \right] X\varphi^\alpha \mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} X_i u X_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i} u X_i \varphi^\alpha \mathrm{d}\xi \\
&+ C \fint_{B_\rho(\xi_0)} (a|Xu|^m + b)\varphi^\alpha \mathrm{d}\xi \\
&+ \fint_{B_\rho(\xi_0)} \left\{ \int_0^1 |A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl) - A^\alpha_{i,p^j_\beta}(\xi_0,u_0,\theta Xu + (1-\theta)Xl)| |Xu - Xl| \mathrm{d}\theta \right\} \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
&+ \fint_{B_\rho(\xi_0)} |A^\alpha_i(\xi_0,u_0,Xu) - A^\alpha_i(\xi,u_0 + Xl(\xi^1 - \xi^1_0),Xu)| \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
&+ \fint_{B_\rho(\xi_0)} |A^\alpha_i(\xi,u_0 + Xl(\xi^1 - \xi^1_0),Xu) - A^\alpha_i(\xi,u,Xu)| \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
=& I' + II' + III' + IV' + V',
\end{aligned} \tag{3.25}$$

where we have used the fact that $\fint_{B_\rho(\xi_0)} A^\alpha_i(\xi_0,u_0,Xl)X\varphi^\alpha d\xi = 0$.

By the relationship of $T = X_i X_{n+i} - X_{n+i} X_i$, the term I' can be estimated as follows:

$$\begin{aligned}
I' =& \fint_{B_\rho(\xi_0)} X_i u X_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i} u X_i \varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} X_i(u - l) X_{n+i}\varphi^\alpha \mathrm{d}\xi + \fint_{B_\rho(\xi_0)} X_i l X_{n+i}\varphi^\alpha \mathrm{d}\xi
\end{aligned}$$

$$
\begin{aligned}
&-\fint_{B_\rho(\xi_0)} X_{n+i}(u-l)X_i\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i}lX_i\varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} X_i(u-l)X_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i}(u-l)X_i\varphi^\alpha \mathrm{d}\xi \\
&-\fint_{B_\rho(\xi_0)} X_{n+i}X_i l\varphi^\alpha \mathrm{d}\xi + \fint_{B_\rho(\xi_0)} X_iX_{n+i}l\varphi^\alpha \mathrm{d}\xi \\
=& \fint_{B_\rho(\xi_0)} X_i(u-l)X_{n+i}\varphi^\alpha \mathrm{d}\xi - \fint_{B_\rho(\xi_0)} X_{n+i}(u-l)X_i\varphi^\alpha \mathrm{d}\xi \\
\le& \sup_{B_\rho(\xi_0)} |X\varphi| \int_{B_\rho(\xi_0)} |Xu - Xl| \mathrm{d}\xi. \qquad (3.26)
\end{aligned}
$$

Let

$$
B_1 =: B_\rho(\xi_0) \cap \{|Xu - Xl| \le 1\}, \quad B_2 =: B_\rho(\xi_0) \cap \{|Xu - Xl| > 1\}.
$$

It follows that

$$
\begin{aligned}
\fint_{B_\rho(\xi_0)} |Xu - Xl| \mathrm{d}\xi &= \fint_{B_1} |Xu - Xl| \mathrm{d}\xi + \fint_{B_2} |Xu - Xl| \mathrm{d}\xi \\
&\le \left(\fint_{B_1} |Xu - Xl|^2 \mathrm{d}\xi \right)^{\frac{1}{2}} + \left(\fint_{B_2} |Xu - Xl|^m \mathrm{d}\xi \right)^{\frac{1}{m}} \\
&\le C \left[\left(\fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 \mathrm{d}\xi \right)^{\frac{1}{2}} + \left(\fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 \mathrm{d}\xi \right)^{\frac{1}{m}} \right] \\
&\le C \left(\Phi^{\frac{1}{2}}(\xi_0, \rho, Xl) + \Phi^{\frac{1}{m}}(\xi_0, \rho, Xl) \right). \qquad (3.27)
\end{aligned}
$$

We then obtain:

$$
I' \le C(\Phi^{\frac{1}{2}}(\xi_0, \rho, Xl) + \Phi^{\frac{1}{m}}(\xi_0, \rho, Xl)). \qquad (3.28)
$$

With the help of the fact that $\sup\limits_{B_\rho(\xi_0)} |\varphi| \le \rho \le 1$, we derive

$$
II' = \fint_{B_\rho(\xi_0)} (a|Xu|^m + b)|\varphi| \mathrm{d}\xi \le 2^{m-1}\rho \left[\fint_{B_\rho(\xi_0)} a|Xu - Xl|^m \mathrm{d}\xi + (a|Xl|^m + b) \right].
$$

For the case where $B_1 =: B_\rho(\xi_0) \cap \{|Xu - Xl| \le 1\}$, it follows, by Young's inequality and (2.9), that

$$
|Xu - Xl|^m \le |Xu - Xl|^{m \cdot \frac{2}{m}} + 1^{\frac{2}{2-m}} \le C(m, M)|V(Xu) - V(Xl)|^2 + 1,
$$

and thus

$$
II' \le C(m, M) \left[\fint_{B_1} a|V(Xu) - V(Xl)|^2 \mathrm{d}\xi + (a|Xl|^m + a + b)\mu(\sqrt{\rho}) \right].
$$

On the other hand, for $B_2 =: B_\rho(\xi_0) \cap \{|Xu - Xl| > 1\}$, it follows by Young's inequality and (2.10) that

$$II' \le C(m,M)\left[\fint_{B_2} a|V(Xu) - V(Xl)|^2 \mathrm{d}\xi + (a|Xl|^m + b)\mu(\sqrt{\rho}) \right].$$

Thus, by combining these estimates and noting the definition of $F(s,t)$, we infer that

$$II' \le C(a,m,M)[\Phi(\xi_0,\rho,Xl) + [a(|Xl|^m + 1) + b]\mu(\sqrt{\rho})]. \tag{3.29}$$

We can estimate the integrand of III' in different ways depending on whether $|Xu - Xl| \le 1$ or $|Xu - Xl| > 1$.

For the first case, $|Xu - Xl| \le 1$. Applying (1.7), Lemma 2.2 (1), Hölder's inequality, and Jensen's inequality leads to

$$\begin{aligned}
III' =& \fint_{B_\rho(\xi_0)} \left[\int_0^1 |A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl) - A^\alpha_{i,p^j_\beta}(\xi_0,u_0,\theta Xu + (1-\theta)Xl)||Xu - Xl|\mathrm{d}\theta \right] \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} \left\{ \int_0^1 \left[(1 + |Xl|^2 + |Xl + \theta(Xu - Xl)|^2)^{\frac{m-2}{2}} \omega(|Xl|, |\theta(Xu - Xl)|) \right] \mathrm{d}\theta \right\} |Xu - Xl| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} \omega(|Xl|, |Xu - Xl|)|Xu - Xl| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} \omega(|Xl|, |V(Xu - Xl)|)|V(Xu - Xl)| \mathrm{d}\xi \\
\le& C\omega \left(|Xl|, \left(\fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 \mathrm{d}\xi \right)^{\frac{1}{2}} \right) \left(\fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 \mathrm{d}\xi \right)^{\frac{1}{2}} \\
=& C\omega\left(|Xl|, \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl) \right) \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl).
\end{aligned} \tag{3.30}$$

For the second case, $|Xu - Xl| > 1$. By the assumption in (H1), Lemma 2.2(1), and $\frac{2Q}{Q+2} < m < 2$, one gets

$$\begin{aligned}
III' =& \fint_{B_\rho(\xi_0)} \left[\int_0^1 |A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl) - A^\alpha_{i,p^j_\beta}(\xi_0,u_0,\theta Xu + (1-\theta)Xl)||Xu - Xl|\mathrm{d}\theta \right] \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} \left\{ \int_0^1 \left[(1 + |Xl|^2)^{\frac{m-2}{2}} + \left(1 + |Xl + \theta(Xu - Xl)|^2\right)^{\frac{m-2}{2}} \right] \mathrm{d}\theta \right\} |Xu - Xl| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} |Xu - Xl| \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} |Xu - Xl|^m \mathrm{d}\xi \\
\le& C \fint_{B_\rho(\xi_0)} |V(Xu) - V(Xl)|^2 \mathrm{d}\xi \\
=& C\Phi(\xi_0,\rho,Xl).
\end{aligned} \tag{3.31}$$

Combining the last two estimates mentioned above implies that

$$III' = \fint_{B_\rho(\xi_0)} \left[\int_0^1 |A^\alpha_{i,p^j_\beta}(\xi_0,u_0,Xl) - A^\alpha_{i,p^j_\beta}(\xi_0,u_0,\theta Xu + (1-\theta)Xl)||Xu - Xl|\mathrm{d}\theta \right] \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi$$

$$\leq C\omega\left(|Xl|, \Phi^{\frac{1}{2}}(\xi_0,\rho,Xl)\right)\Phi^{\frac{1}{2}}(\xi_0,\rho,Xl) + C\Phi(\xi_0,\rho,Xl). \tag{3.32}$$

By employing (H3), Lemma 2.2, and Young's inequality, and noting the fact that $K(\cdot) > 1$, we deduce that

$$\begin{aligned}
IV' =& \fint_{B_\rho(\xi_0)} |A_i^\alpha(\xi_0,u_0,Xu) - A_i^\alpha(\xi, u_0 + Xl(\xi^1-\xi_0^1), Xu)| \sup_{B_\rho(\xi_0)} |X\varphi| \mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} K(\cdot)\mu(\rho)(1+|Xl|)(1+|Xu|)^{\frac{m}{2}}\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} K(\cdot)\mu(\rho)(1+|Xl|)\left[(1+|Xl|)^{\frac{m}{2}} + |Xu-Xl|^{\frac{m}{2}}\right]\mathrm{d}\xi \\
\leq& K(\cdot)\mu(\rho)(1+|Xl|)^{1+\frac{m}{2}} + \fint_{B_1+B_2} K(\cdot)\mu(\rho)(1+|Xl|)|Xu-Xl|^{\frac{m}{2}}\mathrm{d}\xi \\
\leq& K(\cdot)\mu(\rho)(1+|Xl|)^{1+\frac{m}{2}} + \fint_{B_1} K(\cdot)\mu(\rho)(1+|Xl|)|Xu-Xl|^{\frac{m}{2}}\mathrm{d}\xi \\
&+ \fint_{B_2} K(\cdot)\mu(\rho)(1+|Xl|)|Xu-Xl|^{\frac{m}{2}}\mathrm{d}\xi \\
\leq& K(\cdot)\mu(\rho)(1+|Xl|)^{1+\frac{m}{2}} + [K(\cdot)\mu(\rho)(1+|Xl|)]^2 \\
&+ [K(\cdot)\mu(\rho)(1+|Xl|)]^{\frac{4}{4-m}} + \fint_{B_2}|Xu-Xl|^m\mathrm{d}\xi + \fint_{B_1}|Xu-Xl|^2\mathrm{d}\xi \\
\leq& \Phi(\xi_0,\rho,Xl) + 3\left[K(\cdot)(1+|Xl|)\right]^2\mu(\rho),
\end{aligned} \tag{3.33}$$

where we have used the fact that $4/(4-m) < 2,\ 1+m/2 < 2$ and $\mu(\rho) \leq 1$ for $\rho \in [0,1]$.

Using the inequality $s\mu(t) \leq s\mu(s) + t$ for $s \in [0,1]$ and $t > 0$, we obtain:

$$\begin{aligned}
V' =& \fint_{B_\rho(\xi_0)} |A_i^\alpha(\xi, u_0 + Xl(\xi^1-\xi_0^1), Xu) - A_i^\alpha(\xi,u,Xu)|\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} K(\cdot)(1+|Xu|)^{\frac{m}{2}}\mu(|v|)\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} \frac{1}{\sqrt{\rho}}\left[K(\cdot)(1+|Xu|)^{\frac{m}{2}}\sqrt{\rho}\right]\mu(|v|)\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} \frac{1}{\sqrt{\rho}}\left[|v| + K(\cdot)(1+|Xu|)^{\frac{m}{2}}\sqrt{\rho}\mu\left(K(\cdot)(1+|Xu|)^{\frac{m}{2}}\sqrt{\rho}\right)\right]\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} \left[\left|\frac{v}{\rho}\right|\sqrt{\rho} + K^2(\cdot)(1+|Xu|)^m\mu(\sqrt{\rho})\right]\mathrm{d}\xi \\
\leq& \fint_{B_\rho(\xi_0)} \left[\left|\frac{v}{\rho}\right|\sqrt{\rho} + K^2(\cdot)(1+|Xl|)^m\mu(\sqrt{\rho}) + K^2(\cdot)(|Xu-Xl|)^m\mu(\sqrt{\rho})\right]\mathrm{d}\xi.
\end{aligned} \tag{3.34}$$

To further estimate the term V', we divide the ball $B_\rho(\xi_0)$ into four parts.

Case 1: $B_\rho(\xi_0) \cap \{|v/\rho| > 1\} \cap \{|Xu - Xl| \leq 1\}$. By Young's inequality, the estimates of (2.9) and (2.10), and Sobolev–Poincaré inequality (2.11), it follows that

$$\left|\frac{v}{\rho}\right|\sqrt{\rho} + K^2(\cdot)(1+|Xl|)^m\mu(\sqrt{\rho}) + K^2(\cdot)(|Xu-Xl|)^m\mu(\sqrt{\rho})$$

$$\leq C(m,M)(C_p+1)|V(Xu-Xl)|^2+K^{4/(2-m)}(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho}),$$

where we have used the fact that $\rho \leq \mu(\rho) \leq \mu(\sqrt{\rho})$.
Case 2: $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|Xu-Xl|\leq 1\}$. It yields

$$\begin{aligned}&\left|\frac{v}{\rho}\right|\sqrt{\rho}+K^2(\cdot)(1+|Xl|)^m\mu(\sqrt{\rho})+K^2(\cdot)(|Xu-Xl|)^m\mu(\sqrt{\rho})\\ \leq &C(m,M)(C_p+1)|V(Xu-Xl)|^2+K^{4/(2-m)}(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho}).\end{aligned}$$

Case 3: $B_\rho(\xi_0)\cap\{|v/\rho|\leq 1\}\cap\{|Xu-Xl|> 1\}$. It follows

$$\begin{aligned}&\left|\frac{v}{\rho}\right|\sqrt{\rho}+K^2(\cdot)(1+|Xl|)^m\mu(\sqrt{\rho})+K^2(\cdot)(|Xu-Xl|)^m\mu(\sqrt{\rho})\\ \leq &C(m,M)(C_p+K^2)|V(Xu-Xl)|^2+K^2(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho}).\end{aligned}$$

Case 4: $B_\rho(\xi_0)\cap\{|v/\rho|> 1\}\cap\{|Xu-Xl|> 1\}$. It leads to

$$\begin{aligned}&\left|\frac{v}{\rho}\right|\sqrt{\rho}+K^2(\cdot)(1+|Xl|)^m\mu(\sqrt{\rho})+K^2(\cdot)(|Xu-Xl|)^m\mu(\sqrt{\rho})\\ \leq &C(m,M)(C_p+K^2)|V(Xu-Xl)|^2+K^2(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho}).\end{aligned}$$

Combining these estimates above, we obtain:

$$\begin{aligned}V' \leq &\fint_{B_\rho(\xi_0)} K(\cdot)(1+|Xu|)^{\frac{m}{2}}\mu(|v|)\mathrm{d}\xi\\ \leq &C(m,M)\left(C_p+K^2(\cdot)\right)\fint_{B_\rho(\xi_0)}|V(Xu-Xl)|^2 d\xi+K^{4/(2-m)}(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho})\\ \leq &C(m,M)(C_p+K^2(\cdot))\Phi(\xi_0,p_0,Xl)+K^{4/(2-m)}(\cdot)(2+|Xl|)^m\mu(\sqrt{\rho}).\end{aligned} \tag{3.35}$$

Substituting (3.28), (3.31)–(3.33), and (3.35) into (3.25), we can immediately conclude that (3.23) holds. □

3.3. Excess improvement

In this part, we apply linearization tools and $\mathcal{A}$-harmonic approximation techniques to establish improved estimates for the excess functional Φ. For sake of simplicity, motivated by the form of the Caccioppoli-type inequalities, we set the following re-normalized excess functionals:

$$\Phi(\xi_0,\rho,l)=\fint_{B_\rho(\xi_0)}|\ V(Xu-Xl)\ |^2\ \mathrm{d}\xi,$$

and

$$\Psi(\xi_0,\rho,l)=\fint_{B_\rho(\xi_0)}\left|V\left(\frac{u-u_0-Xl(\xi^1-\xi_0^1)}{\rho}\right)\right|^2\mathrm{d}\xi+\fint_{B_\rho(\xi_0)}\left|\frac{u-u_0-Xl(\xi^1-\xi_0^1)}{\rho}\right|^2\mathrm{d}\xi.$$

Lemma 3.4. *Let $u \in HW^{1,m}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ satisfy the conditions of Theorem 1.1. Assume that the following smallness conditions are satisfied:*

$$\omega\left(|(Xu)_{\xi_0,\rho}|, \Phi^{\frac{1}{2}}(\xi_0, \rho, (Xu)_{\xi_0,\rho})\right) + \Phi^{\frac{1}{2}}(\xi_0, \rho, (Xu)_{\xi_0,\rho}) \leq \frac{\delta}{4}, \tag{3.36}$$

$$C_2 F^2(|u_{\xi_0,\rho}|, |(Xu)_{\xi_0,\rho}|)\mu(\sqrt{\rho}) \leq \delta^2, \tag{3.37}$$

with $C_2 = 8C_1^2 C^2(m, M)C_4$, together with the condition

$$\rho \leq \rho_1^{\frac{m}{(2-m)(m-1)}}(1 + |u_{\xi_0,\rho}|, 1 + |(Xu)_{\xi_0,\rho}|). \tag{3.38}$$

Then, we have the excess estimate for $\tau \in [\gamma, 1)$:

$$\Phi(\xi_0, \theta\rho, (Xu)_{\xi_0,\theta\rho}) \leq \theta^{2\tau}\Phi(\xi_0, \rho, (Xu)_{\xi_0,\rho}) + K^*(|u_{\xi_0,\rho}|, |(Xu)_{\xi_0,\rho}|)\mu^2(\rho^\sigma), \tag{3.39}$$

where $\sigma = \min\{(2 - m)(m - 1)/m, (m - 1)/2\}$, and $K^*(s, t) = C_7 H^{2/(m-1)^2(2-m)}(s, M + t)$.

Proof. For simplicity, we use the abbreviation $\Phi(\rho) = \Phi(\xi_0, \rho, (Xu)_{\xi_0,\rho})$ in what follows. For $\varepsilon > 0$ (to be determined later), we take $\delta \in (0, 1)$ and $\Upsilon \in [0, 1]$ to be the corresponding constant from the $\mathcal{A}$-harmonic approximation lemma and set

$$\omega = u - (u_{\xi_0,\rho} - \Upsilon h_{\xi_0,2\theta\rho}) - (Xu)_{\xi_0,\rho}(\xi^1 - \xi_0^1),\ \ \Upsilon = \tilde{C}\Gamma(\rho) \text{ with } \tilde{C} = max\{C_1, \sqrt{C_c}\}, \tag{3.40}$$

and

$$\Gamma(\rho) = \sqrt{\left(\frac{\delta}{4}\right)^{-2}\Phi(\rho) + \Psi(\rho) + \left(\frac{\delta}{4}\right)^{-2}\Phi^{\frac{2}{m}}(\rho) + 16\delta^{-2}\mu^2(\sqrt{\rho})F(|u_{\xi_0,\rho}|, |(Xu)_{\xi_0,\rho}|)}. \tag{3.41}$$

Noting the smallness assumptions in (3.36) and (3.37), we infer that

$$\omega(|(Xu)_{\xi_0,\rho}|, \Phi^{\frac{1}{2}}(\xi_0, \rho, (Xu)_{\xi_0,\rho})) + \Phi^{\frac{1}{2}}(\xi_0, \rho, (Xu)_{\xi_0,\rho}) \leq \frac{\delta}{4}, \tag{3.42}$$

$$\begin{aligned}
&\left|\fint_{B_\rho(\xi_0)}\left[A^\alpha_{i,p^j_\beta}(\xi_0, u_{\xi_0,\rho}, (Xu)_{\xi_0,\rho})X\omega\right]X_i\varphi^\alpha d\xi\right| \\
\leq&\Upsilon\frac{\omega(|(Xu)_{\xi_0,\rho}|, \Phi^{1/2}(\rho))\Phi^{1/2}(\rho) + \Phi(\rho) + \Phi^{1/2}(\rho) + \Phi^{1/m}(\rho) + \mu(\sqrt{\rho})F(|u_{\xi_0,\rho}|, |(Xu)_{\xi_0,\rho}|)}{C(m, M)\Gamma(\rho)}\sup_{B_\rho(\xi_0)}|X\varphi| \\
\leq&\Upsilon\left[\frac{\delta}{4}\left(\omega(|(Xu)_{\xi_0,\rho}|, \Phi^{1/2}(\rho)) + \Phi^{1/2}(\rho)\right) + \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4}\right]\sup_{B_\rho(\xi_0)}|X\varphi| \\
\leq&\Upsilon\left[\omega(|(Xu)_{\xi_0,\rho}|, \Phi^{1/2}(\rho)) + \Phi^{1/2}(\rho) + \frac{3\delta}{4}\right]\sup_{B_\rho(\xi_0)}|X\varphi| \\
\leq&\Upsilon\delta\sup_{B_\rho(\xi_0)}|X\varphi|.
\end{aligned} \tag{3.43}$$

Then, from the definition of Υ and the Caccioppoli-type inequality (3.6) with $l = l_{\xi_0,\rho}$

$$\fint_{B_\rho(\xi_0)} |V(X\omega)|^2 \mathrm{d}\xi \leq C_c\,(\Psi(\rho) + F) \leq \Upsilon^2 \leq 1. \tag{3.44}$$

We observe that (3.43) and (3.44) fulfill the conditions of the $\mathcal{A}$-harmonic approximation lemma, which ensures that we find an $\mathcal{A}$-harmonic function $h \in HW^{1,m}(B_\rho(\xi_0), \mathbb{R}^N)$ such that

$$\fint_{B_\rho(\xi_0)} |V(Xh)|^2 \mathrm{d}\xi \leq 1, \fint_{B_\rho(\xi_0)} \left|V\left(\frac{\omega - \Upsilon h}{\rho}\right)\right|^2 \mathrm{d}\xi \leq \Upsilon^2 \varepsilon. \tag{3.45}$$

With the help of Lemma 2.2,

$$\begin{aligned}
\Phi(\theta\rho) =& \fint_{B_\rho(\xi_0)} \left|V(Xu) - V\left((Xu)_{\xi_0,\theta\rho}\right)\right|^2 \mathrm{d}\xi \\
\leq& C(m, M) \fint_{B_\rho(\xi_0)} \left|V\left(Xu - (Xu)_{\xi_0,\theta\rho}\right)\right|^2 \mathrm{d}\xi \\
\leq& C(m, M) \fint_{B_\rho(\xi_0)} \left|V\left(Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}\right)\right|^2 \mathrm{d}\xi \\
&+ C(m, M) \left|V\left((Xu)_{\xi_0,\theta\rho} - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}\right)\right|^2.
\end{aligned} \tag{3.46}$$

Next, we proceed to estimate the right-hand side of (3.46), by decomposing $B_{\theta\rho}(\xi_0)$ into a set with

$$B_1 = B_{\theta\rho}(\xi_0) \cap \{|Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}| \leq 1\},$$

and

$$B_2 = B_{\theta\rho}(\xi_0) \cap \{|Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}| > 1\}.$$

Then, by Lemma 2.2 (1) and Hölder's inequality, we obtain:

$$\begin{aligned}
|(Xu)_{\xi_0,\theta\rho} - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}| =& \left|\fint_{B_{\theta\rho}(\xi_0)} |Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}| \mathrm{d}\xi\right| \\
\leq& \sqrt{2}\left[\fint_{B_{\theta\rho}(\xi_0)} |V(Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)})|^2 \mathrm{d}\xi\right]^{\frac{1}{2}} \\
&+ \sqrt[m]{2}\left[\fint_{B_{\theta\rho}(\xi_0)} |V(Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)})|^{\frac{2}{m}\cdot m} \mathrm{d}\xi\right]^{\frac{1}{m}} \\
\leq& \sqrt[m]{2}(E^{\frac{1}{2}} + E^{\frac{1}{m}}),
\end{aligned} \tag{3.47}$$

where we have used the term

$$E =: \fint_{B_{\theta\rho}(\xi_0)} \left|V\left(Xu - (Xu)_{\xi_0,\rho} - \Upsilon(Xh)_{(\xi_0,2\theta\rho)}\right)\right|^2 \mathrm{d}\xi. \tag{3.48}$$

Since $V(\zeta)$ is monotone increasing in ζ, it follows that, from (3.46),

$$\Phi(\theta\rho) \leq C\left(E + V^2(E^{1/2} + E^{1/m})\right) \leq C(E + E^{2/m}). \tag{3.49}$$

Now it remains for us to estimate E, noting that

$$\fint_{B_\rho(\xi_0)} |Xh| \mathrm{d}\xi \leq 2\sqrt{2} \fint_{B_\rho(\xi_0)} |V(Xh)|^2 \mathrm{d}\xi \leq 2\sqrt{2}. \tag{3.50}$$

Note that the smallness conditions in (3.36) and (3.37) imply that $C_4\Upsilon^2 \leq 1$ with $C_4 = \max\{8C_0, (2\theta)^{-Q}\}$, where we have assumed $4C_1^2C^2(m, M)C_4 \leq 1$, which is no restriction. By applying the priori estimate for constant coefficients sub-elliptic systems, we have the following:

$$\Upsilon|(Xh)_{(\xi_0,2\theta\rho)}| \leq \Upsilon \sup_{B_{\rho/2}(\xi_0)} |Xh| \leq \Upsilon\sqrt{C_0} \fint_{B_\rho(\xi_0)} |Xh| \mathrm{d}\xi \leq 2\sqrt{2}\Upsilon\sqrt{C_0} \leq 1. \tag{3.51}$$

The Caccioppoli-type inequality applied to $B_{\theta\rho}(\xi_0)$ with $u_0 = u_{\xi_0,\rho}$, $Xl = (Xu)_{\xi_0,\rho} + \Upsilon(Xh)_{(\xi_0,2\theta\rho)}$, and $\theta \in (0, 1/4]$ yields

$$\begin{aligned} E \leq & C_c\Bigg[\fint_{B_{2\theta\rho}(\xi_0)} \left|V\left(\frac{u - u_{\xi_0,\rho} - ((Xu)_{\xi_0,\rho} + \Upsilon(Xh)_{(\xi_0,2\theta\rho)})(\xi^1 - \xi_0^1)}{2\theta\rho}\right)\right|^2 \mathrm{d}\xi \\ & + \fint_{B_{2\theta\rho}(\xi_0)} \left|\frac{u - u_{\xi_0,\rho} - ((Xu)_{\xi_0,\rho} + \Upsilon(Xh)_{(\xi_0,2\theta\rho)})(\xi^1 - \xi_0^1)}{2\theta\rho}\right|^2 \mathrm{d}\xi + F\Bigg], \end{aligned} \tag{3.52}$$

where

$$\begin{aligned} F = & [K(|u_{\xi_0,\rho}| + |(Xu)_{\xi_0,\rho} + \Upsilon(Xh)_{(\xi_0,2\theta\rho)}|)(1 + |(Xu)_{\xi_0,\rho} + \gamma(Xh)_{(\xi_0,2\theta\rho)}|)]^{2m/(m-1)}\mu^2((2\theta\rho)^{(2-m)(m-1)/m}) \\ & + (1 + 2M + |(Xu)_{\xi_0,\rho} + \gamma(Xh)_{(\xi_0,2\theta\rho)}|)^{m/(m-1)^2}(2\theta\rho)^{m-1} \\ & + \mu^2((2\theta\rho)^{(2-m)(m-1)/m})[2a(2 + |Xl|) + b]^{2/(m-1)(2-m)}2\theta\rho. \end{aligned} \tag{3.53}$$

By Lemma 2.2, one gets

$$\begin{aligned} & \fint_{B_{2\theta\rho}(\xi_0)} \left|V\left(\frac{u - u_{\xi_0,\rho} - ((Xu)_{\xi_0,\rho} + \Upsilon(Xh)_{\xi_0,2\theta\rho})(\xi^1 - \xi_0^1)}{2\theta\rho}\right)\right|^2 \mathrm{d}\xi \\ \leq & \fint_{B_{2\theta\rho}(\xi_0)} \left|V\left(\frac{u - (u_{\xi_0,\rho} - \Upsilon h_{\xi_0,2\theta\rho}) - (Xu)_{\xi_0,\rho}(\xi^1 - \xi_0^1) - \Upsilon h(\xi)}{2\theta\rho} + \frac{\Upsilon h(\xi) - \Upsilon h_{\xi_0,2\theta\rho} - \Upsilon(Xh)_{\xi_0,2\theta\rho}(\xi^1 - \xi_0^1)}{2\theta\rho}\right)\right|^2 \mathrm{d}\xi \\ \leq & C\left[\fint_{B_{2\theta\rho}(\xi_0)} \left(\left|V\left(\frac{\omega - \Upsilon h(\xi)}{2\theta\rho}\right)\right|^2 + \left|V\left(\Upsilon\frac{h(\xi) - h_{\xi_0,2\theta\rho} - (Xh)_{\xi_0,2\theta\rho}(\xi^1 - \xi_0^1)}{2\theta\rho}\right)\right|^2\right) \mathrm{d}\xi\right]. \end{aligned} \tag{3.54}$$

To estimate the right-hand side, we employ (3.45) to infer that

$$\fint_{B_{2\theta\rho}(\xi_0)} \left|V\left(\frac{\omega - \Upsilon h(\xi)}{2\theta\rho}\right)\right|^2 \mathrm{d}\xi \leq C(2\theta)^{-Q-2} \fint_{B_\rho(\xi_0)} \left|V\left(\frac{\omega - \Upsilon h(\xi)}{\rho}\right)\right|^2 \mathrm{d}\xi \leq C(2\theta)^{-Q-2}\Upsilon^2\varepsilon.$$

Using Lemma 2.2 and the Sobolev–Poincaré-type inequality in Lemma 2.3 leads to

$$
\begin{aligned}
&\fint_{B_{2\theta\rho}(\xi_0)}\left|V\left(\Upsilon\frac{h(\xi)-h_{\xi_0,2\theta\rho}-(Xh)_{\xi_0,2\theta\rho}(\xi^1-\xi_0^1)}{2\theta\rho}\right)\right|^2\mathrm{d}\xi\\
\le& C_p^2\Upsilon^2\fint_{B_{2\theta\rho}(\xi_0)}\left|V\left(Xh(\xi)-(Xh)_{(\xi_0,2\theta\rho)}\right)\right|^2\mathrm{d}\xi\\
\le& C_p^4(2\theta\rho)^2\Upsilon^2\fint_{B_{2\theta\rho}(\xi_0)}|V(X^2h)|^2\mathrm{d}\xi\le C_p^4(2\theta\rho)^2\Upsilon^2\sup_{B_{\rho/2}(\xi_0)}|X^2h|^2\\
\le& CC_0C_p^4(2\theta)^2\Upsilon^2\fint_{B_\rho(\xi_0)}|V(Xh)|^2\mathrm{d}\xi\\
\le& C_5\theta^2\Upsilon^2,
\end{aligned}
\tag{3.55}
$$

where we assume that $C_5=4CC_0C_p^4$.

$$
\begin{aligned}
&\fint_{B_{2\theta\rho}(\xi_0)}\left|\frac{u-u_{\xi_0,\rho}-((Xu)_{\xi_0,\rho}+\Upsilon(Xh)_{(\xi_0,2\theta\rho)})(\xi^1-\xi_0^1)}{2\theta\rho}\right|^2\mathrm{d}\xi\\
\le&\fint_{B_{2\theta\rho}(\xi_0)}\left|\frac{u-(u_{\xi_0,\rho}-\Upsilon h_{\xi_0,2\theta\rho})-(Xu)_{\xi_0,\rho}(\xi^1-\xi_0^1)-\Upsilon h(\xi)}{2\theta\rho}\right.\\
&\left.+\frac{\Upsilon h(\xi)-\Upsilon h_{\xi_0,2\theta\rho}-\Upsilon(Xh)_{\xi_0,2\theta\rho}(\xi^1-\xi_0^1)}{2\theta\rho}\right|^2\mathrm{d}\xi\\
\le& C\left[\fint_{B_{2\theta\rho}(\xi_0)}\left|\frac{\omega-\Upsilon h(\xi)}{2\theta\rho}\right|^2+\left|\Upsilon\frac{h(\xi)-h_{\xi_0,2\theta\rho}-(Xh)_{\xi_0,2\theta\rho}(\xi^1-\xi_0^1)}{2\theta\rho}\right|^2\mathrm{d}\xi\right].
\end{aligned}
\tag{3.56}
$$

Now, we are in the position to estimate $\fint_{B_{2\theta\rho}(\xi_0)}\left|\frac{\omega-\Upsilon h(\xi)}{2\theta\rho}\right|^2\mathrm{d}\xi$. Since $\left|V\left(\frac{\omega-\Upsilon h}{\rho}\right)\right|^2$ is bounded almost everywhere (3.45), we denote its upper bound by M_1. Lemma 2.2 (1) implies that

$$
\left|\frac{\omega-\Upsilon h}{\rho}\right|\le\sqrt{2}\left|V\left(\frac{\omega-\Upsilon h}{\rho}\right)\right|\le\sqrt{2}M_1,\ \text{for}\left|\frac{\omega-\Upsilon h}{\rho}\right|\le 1,
$$

and

$$
\left|\frac{\omega-\Upsilon h}{\rho}\right|\le\left(\sqrt{2}\left|V\left(\frac{\omega-\Upsilon h}{\rho}\right)\right|\right)^{\frac{2}{m}}\le\left(\sqrt{2}M_1\right)^{\frac{2}{m}},\ \text{for}\left|\frac{\omega-\Upsilon h}{\rho}\right|>1.
$$

Hence, we have

$$
\left|\frac{\omega-\Upsilon h}{\rho}\right|\le\max\left\{\sqrt{2}M_1,\left(\sqrt{2}M_1\right)^{\frac{2}{m}}\right\}=M_2.
\tag{3.57}
$$

Furthermore, it leads to

$$
\fint_{B_{2\theta\rho}(\xi_0)}\left|\frac{\omega-\Upsilon h(\xi)}{2\theta\rho}\right|^2\mathrm{d}\xi\le C(2\theta)^{-Q-2}\fint_{B_\rho(\xi_0)}\left|\frac{\omega-\Upsilon h(\xi)}{\rho}\right|^2\mathrm{d}\xi
$$

$$
\begin{aligned}
&\leq C(2\theta)^{-Q-2}M_2^2 \fint_{B_\rho(\xi_0)} \left|\frac{\omega-\Upsilon h(\xi)}{M_2\rho}\right|^2 \mathrm{d}\xi \\
&\leq C2^{-Q-1}\theta^{-Q-2}M_2^2 \fint_{B_\rho(\xi_0)} \left|V\left(\frac{\omega-\Upsilon h(\xi)}{M_2\rho}\right)\right|^2 \mathrm{d}\xi \\
&\leq C(2\theta)^{-Q-2} \fint_{B_\rho(\xi_0)} \left|V\left(\frac{\omega-\Upsilon h(\xi)}{\rho}\right)\right|^2 \mathrm{d}\xi \\
&\leq C(2\theta)^{-Q-2}\Upsilon^2\varepsilon. \qquad (3.58)
\end{aligned}
$$

Using Lemmas 2.2–2.4 and (3.45) yields

$$
\begin{aligned}
\fint_{B_{2\theta\rho}(\xi_0)} \left|\Upsilon\frac{h(\xi)-h_{\xi_0,2\theta\rho}-(Xh)_{\xi_0,2\theta\rho}(\xi^1-\xi_0^1)}{2\theta\rho}\right|^2 \mathrm{d}\xi &\leq \Upsilon^2 \fint_{B_{2\theta\rho}(\xi_0)} \left|\frac{h(\xi)-h_{\xi_0,2\theta\rho}-(Xh)_{\xi_0,2\theta\rho}(\xi^1-\xi_0^1)}{2\theta\rho}\right|^2 \mathrm{d}\xi \\
&\leq C_p^2\Upsilon^2 \fint_{B_{2\theta\rho}(\xi_0)} \left|Xh(\xi)-(Xh)_{(\xi_0,2\theta\rho)}\right|^2 \mathrm{d}\xi \\
&\leq C_p^4(2\theta\rho)^2\Upsilon^2 \fint_{B_{2\theta\rho}(\xi_0)} |X^2h|^2\mathrm{d}\xi \\
&\leq C\theta^2\Upsilon^2. \qquad (3.59)
\end{aligned}
$$

Noting that $\Upsilon(Xh)_{\xi_0,2\theta\rho} \leq 2\sqrt{2}\sqrt{C_0}\Upsilon \leq 1$ and the definition of $H(\cdot)$, we obtain:

$$
\begin{aligned}
&[K(|u_{\xi_0,\rho}|+|(Xu)_{\xi_0,\rho}+\Upsilon(Xh)_{(\xi_0,2\theta\rho)}|)(1+|(Xu)_{\xi_0,\rho}+\gamma(Xh)_{(\xi_0,2\theta\rho)}|)]^{2m/(m-1)}\mu^2((2\theta\rho)^{(2-m)(m-1)/m}) \\
\leq&[K(|u_{\xi_0,\rho}|+|(Xu)_{\xi_0,\rho}+1|)(2+|(Xu)_{\xi_0,\rho})]^{2m/(m-1)}\mu^2(\rho^{(2-m)(m-1)/m}) \\
\leq&H^{m/(m-1)}(1+|u_{\xi_0,\rho}|,|(Xu)_{(\xi_0,\rho)}|)\mu^2(\rho^{(2-m)(m-1)/m}), \qquad (3.60)
\end{aligned}
$$

and

$$
\begin{aligned}
&(1+2M+|(Xu)_{\xi_0,\rho}+\gamma(Xh)_{(\xi_0,2\theta\rho)}|)^{m/(m-1)^2}(2\theta\rho)^{m-1}+[2a(2+|Xl|)+b]^{2/(m-1)(2-m)}2\theta\rho \\
\leq&\left[(2+2M+|(Xu)_{\xi_0,\rho}|)^{m/(m-1)^2}+[2a(2+|Xl|)+b]^{2/(m-1)(2-m)}\right]\mu^2(\rho^{(m-1)/2}), \qquad (3.61)
\end{aligned}
$$

where we have used the fact $\mu^2(\rho^{(m-1)/2}) \geq \mu^2(\sqrt{\rho})$ and the nondecreasing property of μ.

Combining all the above estimates with (3.52) and letting $\varepsilon = \theta^{Q+4}$, we arrive at

$$
E \leq C_6\left[\theta^2\Upsilon^2 + H^{2/(m-1)^2(2-m)}(|u_{\xi_0,\rho}|, M+|(Xu)_{\xi_0,\rho}|)\mu^2(\rho^\sigma)\right], \qquad (3.62)
$$

where $\sigma = \min\{(2-m)(m-1)/m, (m-1)/2\}$ and C_6 depends only on $Q,\ N,\ m,\ M,\ \lambda$, and C_p. For any given $\tau \in (\gamma, 1)$, choosing $\theta \in (0, \frac{1}{4})$ suitably such that $C_3C_6\theta^2 \leq \theta^{2\tau}$, we easily find (note the definition of γ)

$$
\begin{aligned}
\Phi(\theta\rho) \leq&\theta^{2\tau}\left[\Phi(\rho)+C_7H^{2/(m-1)^2(2-m)}(|u_{\xi_0,\rho}|, M+|(Xu)_{\xi_0,\rho}|)\mu^2(\rho^\sigma)\right] \\
:=&\theta^{2\tau}\left[\Phi(\rho)+K^*(|u_{\xi_0,\rho}|,|(Xu)_{\xi_0,\rho}|)\mu^2(\rho^\sigma)\right], \qquad (3.63)
\end{aligned}
$$

where the constant C_7 has the same dependencies as C_6 and$(2-m)(m-1)/m \leq 3-2\sqrt{2} < 1/2$, $K^*(s,t) = C_7H^{2/(m-1)^2(2-m)}(s, M+t)$. □

For $T > 0$, we find $\Phi_0(T) > 0$ (depending on Q, N, λ, L, τ and ω) such that

$$\omega^{\frac{1}{2}}(2T, 2\Phi_0^{\frac{1}{2}}(T)) + 2\Phi_0^{\frac{1}{2}}(T)) \leq \frac{1}{4}\delta, \tag{3.64}$$

and

$$16(2C_1^2)^{\frac{2}{m}}(1 + C_p^{\frac{2}{m}})^2\Phi_0(T) \leq \theta^{2Q}(1 - \theta^\tau)^2 T^2. \tag{3.65}$$

With $\Phi_0(T)$ from (3.64) and (3.65), we choose $\rho_0(T) \in (0, 1]$ (depending on $Q, N, \lambda, L, \tau, \omega, \eta$ and κ) such that

$$\rho_0(T) \leq \rho_1^{(2-m)(m-1)/m}(1 + 2T, 1 + 2T), \tag{3.66}$$

$$C_2 F^2(2T, 2T)\mu^2(\rho_0(T)) \leq \delta^2, \tag{3.67}$$

$$K_0(T)\mu(\rho_0(T)^2) \leq (\theta^{\sigma\gamma} - \theta^{2\tau})\Phi_0(T), \tag{3.68}$$

$$4(2C_1^2)^{\frac{2}{m}}(1 + C_p^{\frac{2}{m}})^2 K_0(T)H(\rho_0(T))^2 \leq \theta^{2Q}(1 - \theta^\gamma)^2(\theta^{\sigma\gamma} - \theta^{2\tau})T^2, \tag{3.69}$$

where $K_0(T) := K^*(2T, 2T)$.

The rest of the process to obtain Theorem 1.1 is very similar to that in [12]. We omit it here.

Author contributions

Beibei Chen: formal analysis, methodology, writing–original draft; Jialin Wang: formal analysis, methodology, writing–original draft, funding acquisition, supervision, writing–review and editing; Dongni Liao: formal analysis, methodology, funding acquisition, supervision, writing–review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors express their gratitude to the referees for their insightful comments and valuable suggestions. This research is supported by the National Natural Science Foundation of China (No. 12061010), Jiangxi Provincial Natural Science Foundation (No. 20242BAB26003), and the Science and Technology Planning Project of Jiangxi Province (No. GJJ2201204).

Conflict of interest

The authors declare no competing interests.

References

1. A. D. Austin, J. T. Tyson, A new proof of the C^∞ regularity of C^2 conformal mappings on the Heisenberg group, *Colloq. Math.*, **150** (2017), 217–228. http://dx.doi.org/10.4064/cm7193-3-2017
2. E. Acerbi, N. Fusco, Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$, *J. Math. Anal. Appl.*, **140** (1989), 115–135. http://dx.doi.org/10.1016/0022-247X(89)90098-X
3. V. Bögelein, F. Duzaar, N. A. Liao, C. Scheven, Gradient Hölder regularity for degenerate parabolic systems, *Nonlinear Anal.*, **225** (2022), 113119. http://dx.doi.org/10.1016/j.na.2022.113119
4. V. Bögelein, F. Duzaar, P. Marcellini, C. Scheven, Boundary regularity for elliptic systems with p, q-growth, *J. Math. Pure. Appl.*, **159** (2022), 250–293. http://dx.doi.org/10.1016/j.matpur.2021.12.004
5. M. Bramanti, M. C. Zhu, L^p and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift, *Analysis and PDE.*, **6** (2013), 1793–1855. http://dx.doi.org/10.2140/apde.2013.6.1793
6. L. Capogna, Regularity of quasi-linear equations in the Heisenberg group, *Commun. Pur. Appl. Math.*, **50** (1997), 867–889. http://dx.doi.org/10.1002/(SICI)1097-0312(199709)50:9¡867::AID-CPA3¿3.0.CO;2-3
7. L. Capogna, Regularity for quasilinear equation and 1-quasiconformal maps in Carnot groups, *Math. Ann.*, **313** (1999), 263–295. http://dx.doi.org/10.1007/s002080050261
8. L. Capogna, N. Garofalo, Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type, *J. Eur. Math. Soc.*, **5** (2003), 1–40. http://dx.doi.org/10.1007/s100970200043
9. G. Citti, S. Mukherjee, Regularity of quasi-linear equations with Hörmander vector fields of step two, *Adv. Math.*, **408** (2022), 108593. http://dx.doi.org/10.1016/j.aim.2022.108593
10. S. H. Chen, Z. Tan, Optimal partial regularity results for nonlinear elliptic systems in Carnot groups, *Discrete Cont. Dyn.*, **33** (2013), 3391–3405. http://dx.doi.org/10.3934/dcds.2013.33.3391
11. F. Duzaar, J. F. Grotowski, Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation, *Manuscripta Math.*, **103** (2000), 267–298. http://dx.doi.org/10.1007/s002290070007
12. F. Duzaar, A. Gastel, Nonlinear elliptic systems with Dini continuous coefficients, *Arch. Math.*, **78** (2002), 58–73. http://dx.doi.org/10.1007/s00013-002-8217-1
13. F. Duzaar, J. F. Grotowski, M. Kronz, Regularity of almost minimizers of quasi-convex variational integrals with subquadratic growth, *Annali di Matematica*, **184** (2005), 421–448. http://dx.doi.org/10.1007/s10231-004-0117-5
14. G. W. Du, J. Q. Han, P. C. Niu, Interior regularity for degenerate equations with drift on homogeneous groups, *RACSAM*, **113** (2019), 587–604. http://dx.doi.org/10.1007/s13398-018-0500-5

15. F. Duzaar, G. Mingione, The p-harmonic approximation and the regularity of p-harmonic maps, *Calc. Var.*, **20** (2004), 235–256. http://dx.doi.org/10.1007/s00526-003-0233-x

16. F. Duzaar, G. Mingione, Regularity for degenerate elliptic problems via p-harmonic approximation, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **21** (2004), 735–766. http://dx.doi.org/10.1016/j.anihpc.2003.09.003

17. A. Domokos, On the regularity of p-harmonic functions in the Heisenberg group, PhD Thesis, University of Pittsburgh, 2004.

18. F. Duzaar, K. Steffen, Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals, *J. Reine Angew. Math.*, **546** (2002), 73–138. http://dx.doi.org/10.1515/CRLL.2002.046

19. A. Föglein, Partial regularity results for sub-elliptic systems in the Heisenberg group, *Calc. Var.*, **32** (2008), 25–51. http://dx.doi.org/10.1007/S00526-007-0127-4

20. X. J. Feng, P. C. Niu, Interior regularity for degenerate elliptic equations with drift on homogeneous groups, *J. Lie Theory*, **23** (2013), 803–825.

21. Y. X. Hou, X. J. Feng, X. W. Cui, Global Hölder estimates for hypoelliptic operators with drift on homogeneous groups, *Miskolc Math. Notes*, **13** (2012), 337–347. http://dx.doi.org/10.18514/MMN.2012.458

22. F. W. Gehring, The L^p−integrability of the partial derivatives of quasiconformal mapping, *Acta Math.*, **130** (1973), 265–277. http://dx.doi.org/10.1007/BF02392268

23. M. Giaquinta, G. Modica, Modica, Partial regularity of minimisers of quasiconvex integrals, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **3** (1986), 185–208. http://dx.doi.org/10.1016/S0294-1449(16)30385-7

24. Y. X. Hou, P. C. Niu, Weighted Sobolev-Morrey estimates for hypoelliptic operators with drift on homogeneous groups, *J. Math. Anal. Appl.*, **428** (2015), 1319–1338. http://dx.doi.org/10.1016/j.jmaa.2015.03.080

25. E. Lanconelli, S. Polidoro, On a class of hypoelliptic evolution operators, *Rendiconti del Seminario Matematico*, **52** (1994), 29–63.

26. E. Lanconelli, A. Pascucci, P. Sergio, Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance, *STAMPA*, **2** (2002), 243–265.

27. J. J. Manfredi, G. Mingione, Regularity results for quasilinear elliptic equations in the Heisenberg group, *Math. Ann.*, **339** (2007), 485–544. http://dx.doi.org/10.1007/s00208-007-0121-3

28. S. Mukherjee, X. Zhong, $C^{1,\alpha}$-Regularity for variational problems in the Heisenberg group, *Anal. PDE*, **14** (2021), 567–594. http://dx.doi.org/10.2140/apde.2021.14.567

29. G. Mingione, A. Zatorska-Goldstein, X. Zhong, Gradient regularity for elliptic equations in the Heisenberg group, *Adv. Math.*, **222** (2009), 62–129. http://dx.doi.org/10.1016/j.aim.2009.03.016

30. E. Shores, Hypoellipticity for linear degenerate elliptic systems in Carnot groups and applications, PhD Thesis, University of Arkansas, 2005.

31. L. Simon, *Lectures on geometric measure theory*, Canberra: Australian National University Press, 1983.

32. J. L. Wang, D. N. Liao, Z. F. Yu, Hölder continuity for sub-elliptic systems under the sub-quadratic controllable growth in Carnot groups, *Rend. Sem. Mat. Univ. Padova*, **130** (2013), 169–202. http://dx.doi.org/10.4171/RSMUP/130-6
33. J. L. Wang, M. C. Zhu, S. J. Gao, D. N. Liao, Regularity for sub-elliptic systems with VMO-coefficients in the Heisenberg group: the sub-quadratic structure case, *Adv. Nonlinear Anal.*, **10** (2021), 420–449. http://dx.doi.org/10.1515/anona-2020-0145
34. S. Z. Zheng, Z. S. Feng, Regularity of subelliptic p-harmonic systems with subcritical growth in Carnot group, *J. Differ. Equations*, **258** (2015), 2471–2494. http://dx.doi.org/10.1016/J.JDE.2014.12.020
35. J. L. Zhang, P. C. Niu, $C^{1,\alpha}$-Regularity for quasilinear degenerate elliptic equation with a dirft term on the Heisenberg group, *B. Sci. Math.*, **175** (2022), 103097. http://dx.doi.org/10.1016/j.bulsci.2022.103097
36. J. L. Zhang, J. L. Wang, Regularity for a nonlinear discontinuous subelliptic system with drift on the Heisenberg group, *Adv. Math. Phys.*, **2022** (2022), 7853139. http://dx.doi.org/10.1155/2022/7853139
37. J. L. Zhang, J. L. Wang, Partial regularity for a nonlinear discontinuous sub-elliptic system with drift on the Heisenberg group: the superquadratic case, *Complex Var. Elliptic*, **69** (2024), 547–572. http://dx.doi.org/10.1080/17476933.2022.2152444