http://www.aimspress.com/journal/mine

Mathematics in Engineering, 3(1): 1–31.
DOI:10.3934/mine.2021008
Received: 23 January 2020
Accepted: 31 August 2020
Published: 22 September 2020

Research article

Regularity for a class of quasilinear degenerate parabolic equations in the Heisenberg group†

Luca Capogna[1], **Giovanna Citti**[2] **and Nicola Garofalo**[3,*]

[1] Department of Mathematical Sciences, Worcester Polytechnic Institute, MA 01609

[2] Dipartimento di Matematica, Piazza Porta S. Donato 5, Università di Bologna, 40126 Bologna, Italy

[3] Dipartimento di Ingegneria Civile e Ambientale (DICEA), Università di Padova, 35131 Padova, Italy

† **This contribution is part of the Special Issue:** Contemporary PDEs between theory and modeling—Dedicated to Sandro Salsa, on the occasion of his 70th birthday
Guest Editor: Gianmaria Verzini
Link: www.aimspress.com/mine/article/5753/special-articles

* **Correspondence:** Email: rembrandt54@gmail.com.

Abstract: We extend to the parabolic setting some of the ideas originated with Xiao Zhong's proof in [31] of the Hölder regularity of $p-$harmonic functions in the Heisenberg group $\mathbb{H}^n$. Given a number $p \geq 2$, in this paper we establish the C^∞ smoothness of weak solutions of a class of quasilinear PDE in $\mathbb{H}^n$ modeled on the equation

$$\partial_t u = \sum_{i=1}^{2n} X_i\Big((1 + |\nabla_0 u|^2)^{\frac{p-2}{2}} X_i u\Big).$$

Keywords: sub elliptic p-Laplacian; parabolic gradient estimates; Heisenberg group

Questa nota è dedicata a Sandro Salsa, con profondo affetto e ammirazione.

1. Introduction

In this paper we establish the C^∞ smoothness of solutions of a certain class of quasilinear parabolic equations in the Heisenberg group $\mathbb{H}^n$ (see [16, 18, 19, 29] for a review of the literature on subelliptic, and degenerate parabolic PDE in the Heisenberg group). In a cylinder $Q = \Omega \times (0, T)$, where $\Omega \subset \mathbb{H}^n$

is an open set and $T > 0$, we consider the equation

$$\partial_t u = \sum_{i=1}^{2n} X_i A_i(x, \nabla_0 u) \qquad \text{in} \quad Q = \Omega \times (0, T), \tag{1.1}$$

modeled on the regularized parabolic p-Laplacian

$$\partial_t u = \sum_{i=1}^{2n} X_i\Big((1 + |\nabla_0 u|^2)^{\frac{p-2}{2}} X_i u\Big), \tag{1.2}$$

where $p \geq 2$. The term *regularized* here refers to the fact that the non-linearity $(1 + |\nabla_0 u|^2)^{\frac{p-2}{2}}$ affects the ellipticity of the right hand side only when the gradient blows up, and not when it vanishes, thus presenting a weaker version of the singularity in the $p-$Laplacian. Here, we indicate with $x = (x_1, ..., x_{2n}, x_{2n+1})$ the variable point in $\mathbb{H}^n$. We alert the reader that, although it is customary to denote the variable x_{2n+1} in the center of the group with the letter t, we will be using z instead, since we have reserved the letter t for the time variable. Consequently, we will indicate with ∂_i partial differentiation with respect to the variable x_i, $i = 1, ..., 2n$, and use the notation $Z = \partial_z$ for the partial derivative $\partial_{x_{2n+1}}$. The notation $\nabla_0 u \cong (X_1 u, ..., X_{2n} u)$ represents the so-called *horizontal gradient* of the function u, where

$$X_i = \partial_i - \frac{x_{n+i}}{2}\partial_z, \qquad X_{n+i} = \partial_{n+i} + \frac{x_i}{2}\partial_z, \quad i = 1, ..., n.$$

As it is well-known, the $2n + 1$ vector fields $X_1, ..., X_{2n}, Z$ are connected by the following commutation relation: for every couple of index i, j, if $i \leq j$, then $i \leq n$ and $[X_i, X_j] = \delta_{i+nj} Z$, all other commutators being trivial.

We now introduce the relevant structural assumptions on the vector-valued function $(x, \xi) \to A(x, \xi) = (A_1(x, \xi), ..., A_{2n}(x, \xi))$: there exist $p \geq 2$, $\delta > 0$ and $0 < \lambda \leq \Lambda < \infty$ such that for a.e. $x \in \Omega, \xi \in \mathbb{R}^{2n}$ and for all $\eta \in \mathbb{R}^{2n}$, one has

$$\begin{cases} \lambda(\delta + |\xi|^2)^{\frac{p-2}{2}}|\eta|^2 \leq \partial_{\xi_j} A_i(x, \xi)\eta_i\eta_j \leq \Lambda(\delta + |\xi|^2)^{\frac{p-2}{2}}|\eta|^2, \\ |A_i(x, \xi)| + |\partial_{x_j} A_i(x, \xi)| \leq \Lambda(\delta + |\xi|^2)^{\frac{p-1}{2}}. \end{cases} \tag{1.3}$$

Given an open set $\Omega \subset \mathbb{H}^n$, we indicate with $W^{1,p}(\Omega)$ the Sobolev space associated with the p-energy $\mathscr{E}_{\Omega,p}(u) = \frac{1}{p}\int_\Omega |\nabla_0 u|^p$, i.e., the space of all functions $u \in L^p(\Omega)$ such that their distributional derivatives $X_i u$, $i = 1, ..., 2n$, are also in $L^p(\Omega)$. The corresponding norm is $||u||^p_{W^{1,p}(\Omega)} = ||u||_{L^p(\Omega)} + ||\nabla_0 u||_{L^p(\Omega)}$. We denote by $W^{1,p}_0(\Omega)$ the completion of $C^\infty_0(\Omega)$ with respect to such norm. A function $u \in L^p((0, T), W^{1,p}_0(\Omega))$ is a weak solution of (1.1) if

$$\int_0^T \int_\Omega u\phi_t \, dxdt - \int_0^T \int_\Omega \sum_{i=1}^{2n} A_i(x, \nabla_0 u) X_i \phi \, dxdt = 0, \tag{1.4}$$

for every $\phi \in C^\infty_0(Q)$. Our main result is the following.

Theorem 1.1. *Let A_i satisfy the structure conditions* (1.3) *for some $p \geq 2$ and $\delta > 0$. We also assume that* (1.1) *can be approximated as in* (1.6)-(1.8) *below. Let $u \in L^p((0, T), W^{1,p}_0(\Omega))$ be a weak solution* *of* (1.1) *in* $Q = \Omega \times (0,T)$. *For any open ball* $B \subset\subset \Omega$ *and* $T > t_2 \geq t_1 \geq 0$, *there exist constants* $C = C(n,p,\lambda,\Lambda,d(B,\partial\Omega),T-t_2,\delta) > 0$ *and* $\alpha = \alpha(n,p,\lambda,\Lambda,d(B,\partial\Omega),T-t_2,\delta) \in (0,1)$ *such that*

$$\|\nabla_0 u\|_{C^{\alpha}(B\times(t_1,t_2))} + \|Zu\|_{C^{\alpha}(B\times(t_1,t_2))} \leq C\Big(\int_0^T \int_\Omega (\delta + |\nabla_0 u|^2)^{\frac{p}{2}} dxdt\Big)^{\frac{1}{p}}. \tag{1.5}$$

Besides the structural hypothesis (1.3), Theorem 1.1 will be established under an additional technical approximating assumption. Namely, for $\varepsilon \geq 0$ we consider the left-invariant Riemannian metric g_ε in $\mathbb{H}^n$ in which the frame defined by $X_1^\varepsilon = X_1, ..., X_{2n}^\varepsilon = X_{2n}, X_{2n+1}^\varepsilon = \varepsilon Z$ is orthonormal, and denote by ∇_ε the gradient in such metric. We will adopt the unconventional notation $W^{1,p,\varepsilon}(\Omega)$ to indicate the Sobolev space associated with the p-energy $\mathscr{E}_{\Omega,p,\varepsilon}(u) = \frac{1}{p}\int_\Omega |\nabla_\varepsilon u|^p$. We assume that one can approximate A_i by a 1-parameter family of regularized approximants $A^\varepsilon(x,\xi) = (A_1^\varepsilon(x,\xi), ..., A_{2n+1}^\varepsilon(x,\xi))$ defined for a.e. $x \in \Omega$ and every $\xi \in \mathbb{R}^{2n+1}$, and such that for a.e. $x \in \Omega$, and for all $\xi = (\xi_1, ..., \xi_{2n}, \xi_{2n+1}) \in \mathbb{R}^{2n+1}$ one has uniformly on compact subsets of $\Omega \times (0,T)$,

$$(A_1^\varepsilon(x,\xi), ..., A_{2n+1}^\varepsilon(x,\xi)) \underset{\varepsilon\to 0^+}{\longrightarrow} (A_1(x,\xi_1,...,\xi_{2n}), ..., A_{2n}(x,\xi_1,...,\xi_{2n}), 0), \tag{1.6}$$

and furthermore

$$\begin{cases} \lambda(\delta + |\xi|^2)^{\frac{p-2}{2}}|\eta|^2 \leq \partial_{\xi_j} A_i^\varepsilon(x,\xi)\eta_i\eta_j \leq \Lambda(\delta + |\xi|^2)^{\frac{p-2}{2}}|\eta|^2, \\ |A_i^\varepsilon(x,\xi)| + |\partial_{x_j} A_i^\varepsilon(x,\xi)| \leq \Lambda(\delta + |\xi|^2)^{\frac{p-1}{2}}, \end{cases} \tag{1.7}$$

for all $\eta \in \mathbb{R}^{2n+1}$, and for some $0 < \lambda \leq \Lambda < \infty$ independent of ε. The proof of the $C^{1,\alpha}$ regularity in Theorem 1.1 is based on a priori estimates for solutions of the one-parameter family of regularized partial differential equations which approximate (1.1) as the parameter $\varepsilon \to 0$. The key will be in establishing estimates that do not degenerate as $\varepsilon \to 0$. Specifically, for any $\varepsilon > 0$ we will consider a weak solution u^ε to the equation

$$\partial_t u^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) \tag{1.8}$$

in a cylinder $Q_0 = B(x_0,R_0) \times (t_0,t_1)$, with $B(x_0,R_0) \subset \Omega$ and $(t_0,t_1) \subset (0,T)$, and with (parabolic) boundary data $u^\varepsilon = u$. Since (1.8) is strongly parabolic for every $\varepsilon > 0$, the solutions u^ε are smooth in every compact subset $K \subset Q_0$ and, in view of the comparison principle, and of the uniform Harnack inequality established in [2], converge uniformly on compact subsets to a function u_0. The bulk of the paper consists in establishing higher regularity estimates for u_ε that are uniform in $\varepsilon > 0$, to show that u_0 inherits such higher regularity and is a solution of (1.1), thus it coincides with u. Here is our main result in this direction.

Theorem 1.2. *In the hypothesis* (1.6), (1.7), *consider for each* $\varepsilon > 0$ *a weak solution* $u^\varepsilon \in L^p((0,T), W^{1,p,\varepsilon}(\Omega)) \cap C^2(Q)$ *of the approximating equation* (1.8) *in* Q. *For any open ball* $B \subset\subset \Omega$ *and* $T > t_2 \geq t_1 \geq 0$ *there exists a constant* $C = C(n,p,\lambda,\Lambda,d(B,\partial\Omega),T-t_2,\delta) > 0$, *such that*

$$\|\nabla_\varepsilon u^\varepsilon\|^p_{L^\infty(B\times(t_1,t_2))} + \int_{t_1}^{t_2}\int_B (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} \sum_{i,j=1}^{2n} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 dxdt \leq C\int_0^T\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} dxdt. \tag{1.9}$$

Moreover, for any open ball $B \subset\subset \Omega$ *and* $T > t_2 \geq t_1 \geq 0$, *there exist constants* $C > 0$ *and* $\alpha \in (0,1)$, *which depend on* $n, p, \lambda, \Lambda, d(B,\partial\Omega), T-t_2, \delta$, *such that*

$$\|\nabla_\varepsilon u^\varepsilon\|_{C^{\alpha}(B\times(t_1,t_2))} + \|Zu^\varepsilon\|_{C^{\alpha}(B\times(t_1,t_2))} \leq C\Big(\int_0^T \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} dxdt\Big)^{\frac{1}{p}}. \tag{1.10}$$

We emphasise that the constants in (1.9) and (1.10) are independent of ε.

It is worth mentioning here that the prototype for the class of equations (1.1) and for their parabolic approximation comes from considering the regularized $p-$Laplacian operator $L_p u = \mathrm{div}_{g_0,\mu_0}((\delta + |\nabla_0 u|^2_{g_0})^{\frac{p-2}{2}} \nabla_0 u)$ in a sub-Riemannian contact manifold (M, ω, g_0), where M is the underlying differentiable manifold, ω is the contact form and g_0 is a Riemannian metric on the contact distribution. The measure μ_0 is the corresponding Popp measure. The approximants are constructed through Darboux coordinates, considering the $p-$Laplacians associated to a family of Riemannian metrics g_ε that tame g_0 and such that the metric structure of the spaces (M, g_ε) converge in the Gromov-Hausdorff sense to the metric structure of (M, ω, g_0). For a more detailed description, see [8, Section 6.1]. As an immediate corollary of Theorem 1.1 one has the following.

Theorem 1.3. *Let (M, ω, g_0) be a contact, sub-Riemannian manifold and let $\Omega \subset M$ be an open set. For $p \geq 2$, consider $u \in L^p((0, T), W^{1,p}_0(\Omega))$ be a weak solution of*

$$\partial_t u = \mathrm{div}_{g_0,\mu_0}((\delta + |\nabla_0 u|^2_{g_0})^{\frac{p-2}{2}} \nabla_0 u),$$

in $Q = \Omega \times (0, T)$. For any open ball $B \subset\subset \Omega$ and $T > t_2 \geq t_1 \geq 0$, there exist constants $C = C(n, p, d(B, \partial\Omega), T - t_2, \delta) > 0$ and $\alpha = \alpha(n, p, d(B, \partial\Omega), T - t_2, \delta) \in (0, 1)$ such that

$$\|\nabla_0 u\|_{C^\alpha(B\times(t_1,t_2))} + \|Zu\|_{C^\alpha(B\times(t_1,t_2))} \leq C\Big(\int_0^T \int_\Omega (\delta + |\nabla_0 u|^2)^{\frac{p}{2}} dxdt\Big)^{\frac{1}{p}}. \tag{1.11}$$

The $C^{1,\alpha}$ estimates in (1.10) in Theorem 1.2 allow us to apply the Schauder theory developed in [5, 30], and finally deduce the following result.

Theorem 1.4. *Let $k \in \mathbb{N}$ and $\alpha \in (0, 1)$. If $A_i(x, \xi), \partial_{x_k} A_i(x, \xi), \partial_{\xi_j} A_i(x, \xi) \in C^{k,\alpha}_{loc}$ satisfy the structure conditions* (1.3) *for some $p \geq 2$ and $\delta > 0$, then any weak solution $u \in L^p((0, T), W^{1,p}_0(\Omega))$ is $C^{k+1,\alpha}$ on compact subsets of Q.*

The present paper is the first study of higher regularity of weak solutions for the non stationary p-Laplacian type in the sub-Riemannian setting, and it is based on the techniques introduced by Zhong in [31]. The stationary case has been developed so far essentially only in the Heisenberg group case thanks to the work of Domokos, [14], Manfredi, Mingione [21], Mingione, Zatorska-Goldstein and Zhong [22], Ricciotti [26, 27] and Zhong [31]. Regularity in more general contact sub-Riemannian manifolds, including the rototraslation group, has been recently established by the two of the authors and coauthors [8] and independently by Mukherjee [25] based on an extension of the techniques in [31]. Domokos and Manfredi [15] are rapidly making substantial progress in higher steps groups and in some special non-group structures, using the Riemannian approximation approach as in the work [8].

The plan of the paper is as follows. In Section 2 we collect some preparatory material that will be used in the main body of the paper. Section 3 is devoted to proving the first part of Theorem 1.2, which establishes the Lipschitz regularity of the approximating solutions u^ε. In Section 4 we prove the Hölder regularity of derivatives of u^ε in Theorem 1.2. Finally, in Section 5 we use the comparison principle and Theorem 1.2 to establish Theorem 1.1.

Some final comments are in order. The non-degeneracy hypothesis $\delta > 0$ in (1.3) (see also (1.7)) is not needed in the Euclidean setting and, in the stationary regime, it is not needed in the Heisenberg group either. We suspect the $C^{1,\alpha}$ regularity of weak solutions for (1.1) still holds without this hypothesis, but at the moment we are unable to prove it. In this note we use $\delta > 0$ notably in (3.17) and in Theorem 3.13.

In order to extend the parabolic regularity theory to the sub-Riemannian setting one has to find a way to implement, in this non-Euclidean framework, some of the techniques introduced by Di Benedetto [13] which rely on non-isotropic cylinders in space-time. The key idea is to work with cylinders whose dimensions are suitably rescaled to reflect the degeneracy exhibited by the partial differential equation. To give an example, if one sets $x \in \Omega$, $R, \mu > 0$, one can define the intrinsic cylinder

$$Q_R(\mu) := B(x, R) \times (-\mu^{2-p} R^2, 0), \text{ with } \sup_{Q_R(\mu)} |\nabla_0 u| \leq \mu.$$

In contrast with the usual parabolic cylinders of the linear theory, the shape of the $Q_R(\mu)$ cylinders is stretched in the time dimension by a factor of the order $|\nabla_0 u|^{2-p}$.

The use of such non-isotropic cylinders seems necessary in order to make-up for the different homogeneity of the time derivative and the space derivatives in the degenerate regime $\delta = 0$. In a future study we plan to return to the problem of extending Di Benedetto's Caccioppoli inequalities on non-isotropic cylinders to the Heisenberg group and beyond.

2. Preliminaries

In this section we collect a few definitions and preliminary results that will be used throughout the rest of the paper. As indicated in the introduction, for each $\varepsilon \in (0, 1)$ we define g_ε to be the Riemannian metric in $\mathbb{H}^n$ such that $X_1, ..., X_{2n}, \varepsilon Z$ is an orthonormal frame, and denote such frame as $X_1^\varepsilon, ..., X_{2n+1}^\varepsilon$. The corresponding gradient operator will be denoted by ∇_ε.

Definition 2.1. For $x_0 \in \Omega \subset \mathbb{H}^n$, we define a parabolic cylinder $Q_{\varepsilon,r}(x_0, t_0) \subset Q$ to be a set of the form $Q_{\varepsilon,r}(x_0, t_0) = B_\varepsilon(x_0, r) \times (t_0 - r^2, t_0)$. where $r > 0$, $B_\varepsilon(x_0, r) \subset \Omega$ denotes the g_ε-Riemannian ball of center x_0 and $t_0 \in (0, T)$. We call parabolic boundary of the cylinder $Q_{\varepsilon,r}(x_0, t_0) \subset Q$ the set $B_\varepsilon(x_0, r) \times \{t_0 - r^2\} \cup \partial B_\varepsilon(x_0, r) \times [t_0 - r^2, t_0)$.

First of all we recall the Hölder regularity, and local boundedness of weak solutions of (1.1) and (1.8) from [2].

Lemma 2.2. *Let $Q = \Omega \times (0, T) \subset \mathbb{H}^n \times \mathbb{R}^+$, and $\delta \geq 0$. For $\varepsilon \geq 0$ and $p \geq 2$, consider a weak solution $u^\varepsilon \in L^p((0, T), W^{1,p,\varepsilon}(\Omega)) \cap C^2(Q)$ of the approximating equation* (1.8) *in Q. For any open ball $B \subset\subset \Omega$ and $T > t_2 \geq t_1 \geq 0$ there exist constants $C = C(n, p, \lambda, \Lambda, d(B, \partial\Omega), T - t_2) > 0$, and $\alpha = \alpha(n, p, \lambda, \Lambda, d(B, \partial\Omega), T - t_2) \in (0, 1)$, such that*

$$||u^\varepsilon||_{C^\alpha(B\times(t_1,t_2))} \leq C\Big(\int_0^T \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} dxdt\Big)^{\frac{1}{p}}. \tag{2.1}$$

When $\varepsilon > 0$ and $\delta > 0$, classical regularity results (e.g., [20]) yield that weak solutions have bounded gradient, and hence (1.8) is strongly parabolic, thus leading to weak solutions being smooth. Clearly such smoothness may degenerate as $\varepsilon \to 0$, and the main point of this paper is to show that this does not happen.

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and let $Q = \Omega \times (0,T)$. For a function $u : Q \to \mathbb{R}$, and $1 \le p, q$ we define the Lebesgue spaces $L^{p,q}(Q) = L^q([0,T], L^p(\Omega))$, endowed with the norms

$$||u||_{L^{p,q}(Q)} = \Big(\int_0^T (\int_\Omega |u|^p dx)^{\frac{q}{p}} dt \Big)^{\frac{1}{q}}. \tag{2.2}$$

When $p = q$, we will refer to $L^{p,p}(Q)$ as $L^p(Q)$. One has the following useful reformulation of the Sobolev embedding theorem [16] in terms of $L^{p,q}$ spaces. In the next statement, we will indicate with $N = 2n + 2$ the homogenous dimension of $\mathbb{H}^n$ with respect to the non-isotropic group dilations, $\delta_\lambda(x) = (\lambda x_1, ..., \lambda^2 z)$, and we will denote by

$$N_1 = N + 2 = 2n + 4 \tag{2.3}$$

the corresponding parabolic dimension with respect to the dilations $(x,t) \to (\delta_\lambda x, \lambda^2 t)$.

Lemma 2.3. *Let v be a Lipschitz function in Q, and assume that for all $0 < t < T$, $v(\cdot, t)$ has compact support in $\Omega \times \{t\}$.*

(i) There exists $C = C(n) > 0$ such that for any $\varepsilon \in [0,1]$ one has

$$||v||_{L^{\frac{2N}{N-2},2}(Q)} \le C||\nabla_\varepsilon v||_{L^{2,2}(Q)}.$$

(ii) If $v \in L^{2,\infty}(Q)$, then $v \in L^{\frac{2N_1}{N_1-2},\frac{2N_1}{N_1-2}}(Q)$, and there exists $C > 0$, depending on n, such that for any $\varepsilon \in [0,1]$ one has

$$||v||^2_{L^{\frac{2N_1}{N_1-2},\frac{2N_1}{N_1-2}}(Q)} \le C(||v||^2_{L^{2,\infty}(Q)} + ||\nabla_\varepsilon v||^2_{L^{2,2}(Q)}).$$

We note that as ε decreases to zero, the background geometry shifts from Riemannian to sub-Riemannian. The stability with respect to ε of the constant C in the Lemma 2.3 is not trivial, see [7,10].

In the sequel we will use an interpolation inequality that will take the place of the Sobolev inequality in a Moser type iteration, just as, for example, in [11, Proposition 4.2]. Although the result does not use the equation at all, we state it in terms that will make it immediately applicable later on. Henceforth, to simplify the notation, we will routinely omit the indication of dx, $dxdt$, etc. in all integrals involved, unless there is risk of confusion.

Lemma 2.4. *Let u^ε be a weak solution of (1.8) in Q. If $\beta \ge 0$, and $\eta \in C^1([0,T], C_0^\infty(\Omega))$ vanishes on the parabolic boundary of Q, then there is a constant $C > 0$, depending only on $||u^\varepsilon||_{L^\infty(Q)}$, such that*

$$\int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+2)/2} |\eta|^{\beta+2} \le C(\beta + p + 1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta-2}{2}} \sum_{i,j=1}^{2n+1} |X_j^\varepsilon X_i^\varepsilon u^\varepsilon|^2 |\eta|^{\beta+2}$$
$$+ C\beta^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2} |\eta|^\beta (|\eta|^2 + |\nabla_\varepsilon \eta|^2).$$

Proof. Writing $(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+2)/2} = (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}(\delta + |\nabla_\varepsilon u^\varepsilon|^2)$, one has

$$\int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+2)/2} |\eta|^{\beta+2} = \delta \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2} |\eta|^{\beta+2}$$

$$+\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}X_i^\varepsilon u^\varepsilon X_i^\varepsilon u^\varepsilon|\eta|^{\beta+2}=\delta\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}|\eta|^{\beta+2}$$
$$-\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega X_i^\varepsilon\Big((\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}X_i^\varepsilon u^\varepsilon\Big)u^\varepsilon|\eta|^{\beta+2}-(\beta+2)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}X_i^\varepsilon u^\varepsilon u^\varepsilon|\eta|^{\beta+1}X_i^\varepsilon\eta$$
$$\le\delta\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}|\eta|^{\beta+2}+(\beta+p+1)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon||u^\varepsilon||\eta|^{\beta+2}$$
$$+C(\beta+2)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+1)/2}|\eta|^{\beta+1}|\nabla_\varepsilon\eta|.$$

To conclude the argument, it suffices to apply Young's inequality.

□

3. Caccioppoli type inequalities and Lipschitz regularity of u^ε

In this section we establish Lipschitz regularity for the derivatives of the solutions u^ε. The main results of this section are summarized in the following estimates, which are unform in $\varepsilon > 0$.

Theorem 3.1. *Let A_i^ε satisfy the structure conditions* (1.3) *for some $p \ge 2$ and $\delta > 0$. Consider an open set $\Omega \subset \mathbb{H}^n$ and $T > 0$, and let u^ε be a weak solution of* (1.8) *in $Q = \Omega\times(0,T)$. For any open ball $B \subset\subset \Omega$ and $T > t_2 \ge t_1 \ge 0$, there exists a constant $C > 0$, depending on $n, p, \lambda, \Lambda, d(B,\partial\Omega), T - t_2, \delta$, such that*

$$\|\nabla_\varepsilon u^\varepsilon\|^p_{L^\infty(B\times(t_1,t_2))}+\int_{t_1}^{t_2}\int_B(\delta+|\nabla_\varepsilon u|^2)^{\frac{p-2}{2}}\Big(\sum_{i,j=1}^{2n}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2+|\nabla_\varepsilon Zu^\varepsilon|\Big)\tag{3.1}$$
$$\le C\int_0^T\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}.$$

The proof of Theorem 3.1 will follow from combining the results in Theorem 3.11, Lemma 3.12, Proposition 3.13 and Proposition 4.1, that are all established later in the section. The Caccioppoli inequalities needed to prove Theorem 3.1 will take up most of the section, and they all apply to a solution u^ε of the approximating equation (1.8) in a cylinder $Q = \Omega\times(0,T)$. We begin with two lemmas in which we explicitly detail the PDE satisfied by the smooth approximants Zu^ε and $X_\ell^\varepsilon u^\varepsilon$.

Lemma 3.2. *Let u^ε be a solution of* (1.8) *in Q. If we set $v_\ell^\varepsilon = X_\ell^\varepsilon u^\varepsilon$, with $\ell = 1,...,2n+1$, and $s_\ell = (-1)^{[\ell/n]}$ for $\ell \le 2n$, $s_{2n+1} = 0$, then the function v_ℓ^ε solves the equation*

$$\partial_t v_\ell^\varepsilon=\sum_{i,j=1}^{2n+1}X_i^\varepsilon\Big(A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon\Big)\tag{3.2}$$
$$+\sum_{i=1}^{2n+1}X_i^\varepsilon\Big(A^\varepsilon_{i,x_\ell}(x,\nabla_\varepsilon u^\varepsilon)-\frac{s_\ell x_{\ell+s_\ell n}}{2}A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\Big)+s_\ell Z(A^\varepsilon_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)).$$

Proof. Differentiating (1.8) with respect to X_ℓ^ε, when $\ell \le n$, we find

$$\partial_t v_\ell^\varepsilon = \sum_{i=1}^{2n+1} X_\ell^\varepsilon X_i^\varepsilon A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon) = \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(X_\ell^\varepsilon A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\Big) + \sum_{i=1}^{2n+1}[X_\ell^\varepsilon, X_i^\varepsilon]A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)$$
$$= \sum_{i,j=1}^{2n+1} X_i^\varepsilon\Big(A_{i,\xi_j}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon\Big) + \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(A_{i,x_\ell}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon) - \frac{x_{\ell+n}}{2}A_{i,x_{2n+1}}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\Big) + Z(A_{\ell+n}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)).$$

Taking the derivative with respect to X_ℓ^ε when $n+1 \le \ell \le 2n$, we obtain

$$\partial_t v_\ell^\varepsilon = \sum_{i=1}^{2n+1} X_\ell^\varepsilon X_i^\varepsilon A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon) = \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(X_\ell^\varepsilon A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\Big) + \sum_{i=1}^{2n+1}[X_\ell^\varepsilon, X_i^\varepsilon]A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)$$
$$= \sum_{i,j=1}^{2n+1} X_i^\varepsilon\Big(A_{i,\xi_j}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_l^\varepsilon X_j^\varepsilon u^\varepsilon\Big) + \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(A_{i,x_\ell}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon) + \frac{x_{\ell-n}}{2}A_{i,x_{2n+1}}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\Big) - Z(A_{\ell-n}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)).$$

Since for $\ell = 2n+1$ the vector field X_ℓ^ε commutes with the others, taking the derivatives with respect to X_{2n+1}^ε we obtain the thesis.

□

Lemma 3.3. *Let u^ε be a solution of* (1.8) *in Q. Then, the function Zu^ε is a solution of the equation*

$$\partial_t Zu^\varepsilon = \sum_{i,j=1}^{2n+1} X_i^\varepsilon(A_{i,\xi_j}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon) + \sum_{i=1}^{2n+1} X_i^\varepsilon(A_{i,x_{2n+1}}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)).$$

Proof. The assertion immediately follows from Lemma 3.2, with $\ell = 2n+1$, since $X_{2n+1}^\varepsilon = \varepsilon Z$.

□

Lemma 3.4. *Let u^ε be a solution of* (1.8) *in Q. For any $\beta \ge 0$ and for all $\eta \in C^1([0,T], C_0^\infty(\Omega))$, one has*

$$\frac{1}{\beta+2}\int_\Omega |Zu^\varepsilon|^{\beta+2}\eta^2\Big|_{t_1}^{t_2} + \frac{\lambda(\beta+1)}{2}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon Zu^\varepsilon|^2|Zu^\varepsilon|^\beta|\eta|^2,$$
$$\le \frac{\Lambda}{(\beta+1)}\Big(\frac{16\Lambda}{\lambda}+2\Big)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon\eta|^2|Zu^\varepsilon|^{\beta+2} + \frac{2}{\beta+2}\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta\partial_t\eta$$
$$+\Lambda(\beta+1)\Big(\frac{16\Lambda}{\lambda}+2\Big)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^2|Zu^\varepsilon|^\beta.$$

Proof. We use $\phi = \eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon$ as a test function in the equation satisfied by Zu^ε, see Lemma 3.3, to obtain

$$\int_{t_1}^{t_2}\int_\Omega \partial_t Zu^\varepsilon\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon = \int_{t_1}^{t_2}\int_\Omega\sum_{i,j=1}^{2n+1}X_i^\varepsilon(A_{i,\xi_j}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon)\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon$$
$$+\int_{t_1}^{t_2}\int_\Omega\sum_{i,j=1}^{2n+1}X_i^\varepsilon(A_{i,x_{2n+1}}^\varepsilon(x,\nabla_\varepsilon u^\varepsilon))\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon.$$

The left-hand side of the latter equation can be expressed as follows:

$$\int_{t_1}^{t_2}\int_\Omega \partial_t Zu^\varepsilon \eta^2 |Zu^\varepsilon|^\beta Zu^\varepsilon = \frac{1}{\beta+2}\int_{t_1}^{t_2}\int_\Omega \partial_t |Zu^\varepsilon|^{\beta+2}\eta^2.$$

Considering the first term in the right-hand side, we obtain

$$\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} X_i^\varepsilon(A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon)\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon = -\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_i^\varepsilon(\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon)$$

$$= -2\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon \eta X_i^\varepsilon\eta|Zu^\varepsilon|^\beta Zu^\varepsilon - (\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon\eta^2|Zu^\varepsilon|^\beta X_i^\varepsilon Zu^\varepsilon.$$

As for the second term in the right-hand side, we have

$$\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} X_i^\varepsilon(A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon))\eta^2|Zu^\varepsilon|^\beta Zu^\varepsilon = -2\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta X_i^\varepsilon\eta|Zu^\varepsilon|^\beta Zu^\varepsilon$$

$$-(\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta^2|Zu^\varepsilon|^\beta X_i^\varepsilon Zu^\varepsilon.$$

Combining the latter three equations, we find

$$\frac{1}{\beta+2}\int_{t_1}^{t_2}\int_\Omega \partial_t|Zu^\varepsilon|^{\beta+2}\eta^2 + (\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1}\partial_{\xi_j}A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon\eta^2|Zu^\varepsilon|^\beta X_i^\varepsilon Zu^\varepsilon$$

$$= -2\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1}\partial_{\xi_j}A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_i^\varepsilon\eta\eta|Zu^\varepsilon|^\beta Zu^\varepsilon - 2\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta X_i^\varepsilon\eta|Zu^\varepsilon|^\beta Zu^\varepsilon$$

$$-(\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta^2|Zu^\varepsilon|^\beta X_i^\varepsilon Zu^\varepsilon.$$

The structure conditions (1.3) yield

$$\frac{1}{\beta+2}\int_\Omega |Zu^\varepsilon|^{\beta+2}\eta^2\Big|_{t_1}^{t_2} + \lambda(\beta+1)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon Zu^\varepsilon|^2|Zu^\varepsilon|^\beta|\eta|^2 \le$$

$$\frac{1}{\beta+2}\int_\Omega |Zu^\varepsilon|^{\beta+2}\eta^2\Big|_{t_1}^{t_2} + (\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1}\partial_{\xi_j}A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_i^\varepsilon Zu^\varepsilon\ \eta^2|Zu^\varepsilon|^\beta$$

$$= -2\int_{t_1}^{t_2}\int_\Omega \sum_{i,j=1}^{2n+1}\partial_{\xi_j}A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_i^\varepsilon\eta\eta|Zu^\varepsilon|^\beta Zu^\varepsilon + \frac{2}{\beta+2}\int_{t_1}^{t_2}\int_\Omega |Zu^\varepsilon|^{\beta+2}\eta\partial_t\eta$$

$$-2\int_{t_1}^{t_2}\int_\Omega \sum_{i=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta X_i^\varepsilon\eta|Zu^\varepsilon|^\beta Zu^\varepsilon - (\beta+1)\int_{t_1}^{t_2}\int_\Omega \sum_{i=1}^{2n+1} A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\eta^2|Zu^\varepsilon|^\beta X_i^\varepsilon Zu^\varepsilon$$

$$\le 2\Lambda\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon Zu|\eta|\nabla_\varepsilon\eta||Zu^\varepsilon|^{\beta+1} + \frac{2}{\beta+2}\int_{t_1}^{t_2}\int_\Omega |Zu^\varepsilon|^{\beta+2}\eta\partial_t\eta$$

$$+ 2\Lambda \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}} \eta |\nabla_\varepsilon \eta| |Zu^\varepsilon|^{\beta+1} + (\beta+1)\Lambda \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}} \eta^2 |Zu^\varepsilon|^\beta |\nabla_\varepsilon Z u^\varepsilon|,$$

thus concluding the proof.

□

Lemma 3.5. *Let u^ε be a weak solution of* (1.8) *in Q. There exists $C_0 = C_0(n, p, \lambda, \Lambda) > 0$. For any $t_2 \geq t_1 \geq 0$, $\beta \geq 0$ and all $\eta \in C_0^\infty(\Omega)$, we have*

$$\frac{1}{\beta+2} \int_\Omega \eta^2 [(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+2)/2}] \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2+\beta)/2} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \tag{3.3}$$
$$\leq C_0 \int_{t_1}^{t_2} \int_\Omega (\eta^2 + |\nabla_\varepsilon \eta|^2 + \eta |Z\eta|)(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}$$
$$+ C_0 (\beta+1)^4 \int_{t_1}^{t_2} \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2} |Zu^\varepsilon|^2.$$

Proof. In view of Lemma 3.2 we know that, if $u^\varepsilon \in C^\infty(Q)$ is a solution of $\partial_t u^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)$, then $v_\ell^\varepsilon = X_\ell^\varepsilon u^\varepsilon$ solves (3.2). If in the first term in the right-hand side of (3.2) we use the fact that $X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon = X_j^\varepsilon X_\ell^\varepsilon u^\varepsilon + [X_\ell^\varepsilon, X_j^\varepsilon] u^\varepsilon = X_j^\varepsilon v_\ell^\varepsilon + s_\ell Z v_\ell^\varepsilon$, we find

$$\partial_t v_\ell^\varepsilon = \sum_{i,j=1}^{2n+1} X_i^\varepsilon \Big(A_{i,\xi_j}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) X_j^\varepsilon v_\ell^\varepsilon \Big) + s_\ell \sum_{i=1}^{2n+1} X_i^\varepsilon \Big(A_{i,\xi_{\ell+s_\ell n}}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) Z u^\varepsilon \Big) + \tag{3.4}$$
$$+ \sum_{i=1}^{2n+1} X_i^\varepsilon \Big(A_{i,x_\ell}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) - \frac{s_\ell x_{\ell+s_\ell n}}{2} A_{i,x_{2n+1}}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) \Big) + s_\ell Z (A_{\ell+s_\ell n}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon)).$$

Fix $\eta \in C_0^\infty(\Omega)$ and let $\phi = \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} X_\ell^\varepsilon u^\varepsilon$. Taking such ϕ as the test-function in the weak form of (3.4), and integrating by parts the terms in divergence form, one has

$$\frac{1}{2} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}} \partial_t \Big[X_\ell^\varepsilon u^\varepsilon \Big]^2 \eta^2$$
$$+ \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega A_{i\xi_j}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) X_j^\varepsilon v_\ell^\varepsilon X_i^\varepsilon \Big(\eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} X_\ell^\varepsilon u^\varepsilon \Big)$$
$$= -s_\ell \sum_{i=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega A_{i\xi_{\ell+s_\ell n}}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) Z u^\varepsilon X_i^\varepsilon \Big(\eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} X_\ell^\varepsilon u^\varepsilon \Big)$$
$$+ \int_{t_1}^{t_2} \int_\Omega \sum_{i=1}^{2n+1} X_i^\varepsilon \Big(A_{i,x_\ell}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) - \frac{s_l x_{\ell+s_\ell n}}{2} A_{i,x_{2n+1}}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon) \Big) \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} X_\ell^\varepsilon u^\varepsilon$$
$$+ s_\ell \int_{t_1}^{t_2} \int_\Omega s_l Z (A_{\ell+s_\ell n}^\varepsilon (x, \nabla_\varepsilon u^\varepsilon)) \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} X_\ell^\varepsilon u^\varepsilon.$$

The latter equation implies that for every $\ell = 1, ..., 2n+1$ one has

$$\frac{1}{2(\beta+2)} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}} \partial_t \Big[X_\ell^\varepsilon u^\varepsilon \Big]^2 \eta^2$$

$$+\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X^\varepsilon_jX^\varepsilon_\ell u^\varepsilon X^\varepsilon_iX^\varepsilon_\ell u^\varepsilon\ \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}$$
$$+\sum_{i,j=1}^{2n+1}\frac{\beta}{2}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X^\varepsilon_jX^\varepsilon_\ell u^\varepsilon X^\varepsilon_\ell u^\varepsilon X^\varepsilon_i(|\nabla_\varepsilon u^\varepsilon|^2)\ \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta-2}{2}}$$
$$=-\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X^\varepsilon_jX^\varepsilon_\ell u^\varepsilon X^\varepsilon_\ell u^\varepsilon X^\varepsilon_i(\eta^2)(\delta+|\nabla_\varepsilon u|^2)^{\beta/2}$$
$$-s_\ell\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i\xi_{\ell+s_\ell n}}(x,\nabla_\varepsilon u^\varepsilon)ZuX^\varepsilon_i\Big(\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X^\varepsilon_\ell u^\varepsilon\Big)$$
$$-\int_{t_1}^{t_2}\int_\Omega\sum_{i=1}^{2n+1}\Big(A^\varepsilon_{i,x_\ell}(x,\nabla_\varepsilon u^\varepsilon)-\frac{s_\ell x_{\ell+s_\ell n}}{2}A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\Big)X^\varepsilon_i\Big(\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X^\varepsilon_\ell u^\varepsilon\Big)$$
$$+s_\ell\sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega Z(A^\varepsilon_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon))\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X^\varepsilon_\ell u^\varepsilon=I^1_\ell+I^2_\ell+I^3_\ell+I^4_\ell.$$

Summing over $\ell=1,...,2n+1$, by a simple application of the chain rule, and using the structural assumption (1.7), we see that the left-hand side can be bounded from below by

$$\frac{1}{\beta+2}\int_{t_1}^{t_2}\int_\Omega\partial_t\Big[(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}+1}\Big]\eta^2$$
$$+\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega\eta^2A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X^\varepsilon_jX^\varepsilon_\ell u^\varepsilon X^\varepsilon_iX^\varepsilon_\ell u^\varepsilon(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}$$
$$+\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\frac{\beta}{2}\int_{t_1}^{t_2}\int_\Omega\eta^2A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X^\varepsilon_jX^\varepsilon_\ell u^\varepsilon X^\varepsilon_\ell u^\varepsilon X^\varepsilon_i(|\nabla_\varepsilon u^\varepsilon|^2)(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta-2}{2}}$$
$$\geq\frac{1}{\beta+2}\int_{t_1}^{t_2}\int_\Omega\partial_t\Big[(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}+1}\Big]\eta^2+\lambda\int_{t_1}^{t_2}\int_\Omega\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2+\beta}{2}}\sum_{i,j=1}^{2n+1}|X^\varepsilon_iX^\varepsilon_ju^\varepsilon|^2$$
$$+\frac{\lambda\beta}{4}\int_{t_1}^{t_2}\int_\Omega\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta-4}{2}}|\nabla^\varepsilon(|\nabla_\varepsilon u^\varepsilon|^2)|^2.$$

Since the last term in the right-hand side of this estimate is nonnegative, we obtain from this bound

$$\frac{1}{\beta+2}\int_\Omega[(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}+1}\eta^2]\Big|_{t_1}^{t_2}+\lambda\int_{t_1}^{t_2}\int_\Omega\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2+\beta}{2}}\sum_{i,j=1}^{2n+1}|X^\varepsilon_iX^\varepsilon_ju^\varepsilon|^2 \tag{3.5}$$
$$\leq\sum_{\ell=1}^{2n+1}\Big(I^1_\ell+I^2_\ell+I^3_\ell+I^4_\ell\Big).$$

Next, we estimate each of the terms in the right-hand side separately. Recalling that from (1.7) one has $|A_{i\xi_j}(x,\eta)|=|\partial_{\xi_j}A^\varepsilon_i(x,\eta)|\leq C(\delta+|\eta|^2)^{\frac{p-2}{2}}$, one has that for any $\alpha>0$ there exists $C_\alpha>0$ depending only on α, p, n and the structure constants, such that

$$\sum_{\ell=1}^{2n+1} I_\ell^1 = -\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon X_\ell^\varepsilon u^\varepsilon X_\ell^\varepsilon u^\varepsilon X_i^\varepsilon(\eta^2)(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} \tag{3.6}$$
$$\le 2\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega |\eta|(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2}|X_j^\varepsilon X_i^\varepsilon u^\varepsilon||\nabla_\varepsilon u^\varepsilon||\nabla_\varepsilon\eta|(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}}$$
$$\le \alpha\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2}|X_j^\varepsilon X_i^\varepsilon u^\varepsilon|^2 + C_\alpha\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}|\nabla_\varepsilon\eta|^2.$$

Analogously, we find

$$\sum_{\ell=1}^{2n+1} I_\ell^2 \le \alpha\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \tag{3.7}$$
$$+ C\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}|\nabla_\varepsilon\eta|^2 + C_\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2}|Zu^\varepsilon|^2.$$

In a similar fashion, we obtain

$$\sum_{\ell=1}^{2n} I_\ell^3 \le \alpha\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega (\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2\eta^2 \tag{3.8}$$
$$+ C_\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}(|\nabla_\varepsilon\eta|^2+|\eta|^2).$$

Finally, integrating by parts twice, and using the structural assumptions, one has

$$\sum_{\ell=1}^{2n+1} I_\ell^4 = -\sum_{\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega Z(A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon))\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X_\ell^\varepsilon u^\varepsilon \tag{3.9}$$
$$= 2\sum_{\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta Z\eta(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X_\ell^\varepsilon u^\varepsilon$$
$$+\beta\sum_{\ell=1}^{2n+1}\sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta-2}{2}}X_j u^\varepsilon X_j Zu^\varepsilon X_\ell^\varepsilon u^\varepsilon$$
$$+\sum_{\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X_\ell^\varepsilon Zu^\varepsilon$$
$$= 2\sum_{\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta Z\eta(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}X_\ell^\varepsilon u^\varepsilon$$
$$-\beta\sum_{\ell=1}^{2n+1}\sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega X_j\Big(A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta-2}{2}}X_j u^\varepsilon X_\ell^\varepsilon u^\varepsilon\Big)Zu^\varepsilon$$
$$-\sum_{\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega X_\ell^\varepsilon\Big(A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon)\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2}\Big)Zu^\varepsilon$$

$$
\begin{aligned}
&\leq \alpha \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \eta^2 \\
&+ C(\beta+1) \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2} \Big(\eta^2 + |\nabla_\varepsilon \eta|^2 + |\eta Z\eta|\Big) \\
&+ C_\alpha(\beta+1)^4 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta-2)/2} |Zu^\varepsilon|^2.
\end{aligned}
$$

Combining (3.6)–(3.9) with (3.5), we reach the desired conclusion (3.3).

□

In the case $\beta = 0$ we obtain the following stronger estimate, which we will need in the sequel. We denote by $\|\cdot\|$ the L^∞ norm of a function on the parabolic cylinder Q.

Lemma 3.6. *Let u^ε be a weak solution of* (1.8) *in Q, let $t_2 \geq t_1 \geq 0$, and $\eta \in C^1([0,T], C_0^\infty(\Omega))$ be such that $0 \leq \eta \leq 1$, and for which $\|\partial_t \eta\| \leq C\|\nabla_\varepsilon \eta\|^2$, where $C > 0$ is a universal constant. For every $\alpha > 0$ there exists $C_\alpha > 0$ such that*

$$
\begin{aligned}
&\frac{1}{2}\int_\Omega ((\delta + |\nabla_\varepsilon u^\varepsilon|^2)\eta^2)\Big|_{t_1}^{t_2} + \lambda \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \eta^2 \\
&\leq \alpha \int_{t_1}^{t_2}\int_\Omega |Zu^\varepsilon|^2 \eta^3 + C_\alpha \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2} \Big(\eta^2 + |\nabla_\varepsilon \eta|^2 + |\eta Z\eta|\Big).
\end{aligned}
$$

Proof. In view of Lemma 3.2 we notice that, if $u^\varepsilon \in C^\infty(Q)$ is a solution of $\partial_t u^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)$, then $v_\ell^\varepsilon = X_\ell^\varepsilon u^\varepsilon$ solves

$$
\begin{aligned}
\partial_t v_\ell^\varepsilon = &\sum_{i,j=1}^{2n+1} X_i^\varepsilon \Big(X_\ell^\varepsilon (A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)) \Big) + \sum_{i=1}^{2n+1} X_i^\varepsilon \Big(A_{i,x_\ell}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) - \frac{s_\ell x_{\ell+s_\ell n}}{2} A_{i,x_{2n+1}}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) \Big) \qquad (3.10) \\
&+ s_\ell Z(A_{\ell+s_\ell n}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)).
\end{aligned}
$$

With η as in the statement of the lemma, we take $\phi = \eta^2 X_\ell^\varepsilon u^\varepsilon$ as a test function in the weak form of (3.10). Integrating by parts the terms in divergence form, one has

$$
\begin{aligned}
&\frac{1}{2}\int_{t_1}^{t_2}\int_\Omega \eta^2 \partial_t (X_\ell^\varepsilon u^\varepsilon)^2 + \sum_{i=1}^{2n+1} \int_{t_1}^{t_2}\int_\Omega X_\ell(A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)) X_i^\varepsilon \Big(\eta^2 X_\ell^\varepsilon u^\varepsilon\Big) \\
&= \int_{t_1}^{t_2}\int_\Omega \eta^2 \sum_{i=1}^{2n+1} X_i^\varepsilon \Big(A_{i,x_\ell}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) - \frac{s_\ell x_{\ell+s_\ell n}}{2} A_{i,x_{2n+1}}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) \Big) X_\ell^\varepsilon u^\varepsilon \\
&+ s_\ell \int_{t_1}^{t_2}\int_\Omega \eta^2 Z(A_{\ell+s_\ell n}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon)) X_\ell^\varepsilon u^\varepsilon.
\end{aligned}
$$

This gives

$$
\frac{1}{2}\int_{t_1}^{t_2}\int_\Omega \eta^2 \partial_t (X_\ell^\varepsilon u^\varepsilon)^2 + \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2}\int_\Omega \eta^2 A_{i,\xi_j}^\varepsilon(x, \nabla_\varepsilon u^\varepsilon) X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon X_\ell^\varepsilon X_i^\varepsilon u^\varepsilon
$$

$$= -\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2 X_\ell(A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon))Zu^\varepsilon - \sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega X_\ell(A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon))\eta X_i^\varepsilon \eta X_\ell^\varepsilon u^\varepsilon$$
$$-\int_{t_1}^{t_2}\int_\Omega \sum_{i=1}^{2n+1}\Big(A^\varepsilon_{i,x_\ell}(x,\nabla_\varepsilon u^\varepsilon) - \frac{s_\ell x_{\ell+s_\ell n}}{2}A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\Big)X_i^\varepsilon\Big(\eta^2 X_\ell^\varepsilon u^\varepsilon\Big)$$
$$+ s_\ell \sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2 Z(A^\varepsilon_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon))X_\ell^\varepsilon u^\varepsilon = I_\ell^1 + I_\ell^2 + I_\ell^3 + I_\ell^4.$$

Summing over $\ell = 1,...,2n+1$, in view of the structural hypothesis (1.7), after an integration by parts in the first term in the left-hand side we obtain the following bound

$$\frac{1}{2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)\eta^2\Big|_{t_1}^{t_2} + \lambda\int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2\eta^2$$
$$\le I_\ell^1 + I_\ell^2 + I_\ell^3 + I_\ell^4 + \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)\eta\partial_t\eta.$$

Next, we estimate each of the terms in the right-hand side separately. Recalling that $|A^\varepsilon_{i\xi_j}(x,\eta)| \le C(\delta + |\eta|^2)^{\frac{p-2}{2}}$, we find that for any $\alpha_1, \alpha_2 > 0$ there exist $C_{\alpha_1}, C_{\alpha_2} > 0$, depending only on α_1, α_2, p, n and the structure constants, such that

$$\sum_{\ell=1}^{2n+1} I_\ell^1 = \sum_{\ell=1}^{2n}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2 X_\ell(A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon))Zu^\varepsilon$$
$$= -2\sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\eta X_\ell\eta Zu^\varepsilon - \sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2 A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon Zu^\varepsilon$$
$$= -2\sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)\eta X_\ell\eta Zu^\varepsilon$$
$$+\sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)2\eta Z\eta X_\ell^\varepsilon u^\varepsilon + \sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2 A^\varepsilon_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_\ell^\varepsilon u^\varepsilon$$
$$\le \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-1)/2}\eta|\nabla_\varepsilon\eta||Zu^\varepsilon| + \sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2}|2\eta Z\eta|$$
$$+\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-1)/2}|\nabla_\varepsilon Zu^\varepsilon|$$
$$\le \alpha_1\int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2}\eta^2|Zu^\varepsilon|^2 + C_{\alpha_1}\int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2}|\nabla_\varepsilon\eta|^2$$
$$+\sum_{\ell=1}^{2n+1}\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2}|2\eta Z\eta|$$
$$+\frac{\alpha_2}{\|\nabla_\varepsilon\eta\|^2}\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^4(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2}|\nabla_\varepsilon Zu^\varepsilon|^2 + C_{\alpha_2}\|\nabla_\varepsilon\eta\|^2\int_{t_1}^{t_2}\int_{supp(\eta)}(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2}.$$

Now, we apply Lemma 3.4 to find, for any $\alpha > 0$,

$$
\begin{aligned}
&\frac{\alpha}{\|\nabla_\varepsilon \eta\|^2} \sum_{\ell=1}^{2n+1} \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2} |\nabla_\varepsilon Z u^\varepsilon|^2 \eta^4 \\
&\leq \alpha C \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |Zu^\varepsilon|^2 \eta^2 + \frac{\alpha \|\partial_t \eta\|}{\|\nabla_\varepsilon \eta\|^2} \int_{t_1}^{t_2} \int_\Omega |Zu^\varepsilon|^2 \eta^3 \\
&+ \alpha \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} \eta^4 .
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
&\sum_{\ell=1}^{2n+1} I_\ell^2 + \sum_{\ell=1}^{2n+1} I_\ell^3 \leq \alpha \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \eta^2 \\
&+ C \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2} |\nabla_\varepsilon \eta|^2 .
\end{aligned}
$$

Using the structure conditions, one has

$$
\begin{aligned}
\sum_{\ell=1}^{2n+1} I_\ell^4 &\leq \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-1)/2} |\nabla_\varepsilon Z u^\varepsilon| \eta^2 \\
&\leq \alpha \sum_{i,j=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2} |\nabla_\varepsilon Z u^\varepsilon|^2 \eta^2 \\
&+ C_\alpha \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2} \Big(\eta^2 + |\nabla_\varepsilon \eta|^2 + |\eta Z \eta| \Big),
\end{aligned}
$$

thus concluding the proof.

□

Next, we need to establish mixed type Caccioppoli inequalities, where the left-hand side includes terms with both horizontal derivatives and derivatives along the second layer of the stratified Lie algebra of $\mathbb{H}^n$.

Lemma 3.7. *Set $T > t_2 > t_1 > 0$. Let u^ε be a weak solution of* (1.8) *in $Q = \Omega \times (0, T)$. Let $\beta \geq 2$ and let $\eta \in C^1((0, T), C_0^\infty(\Omega))$, with $0 \leq \eta \leq 1$. For all $\alpha \leq 1$ there exist constants C_Λ, $C_\alpha = C(\alpha, \lambda, \Lambda) > 0$ such that*

$$
\begin{aligned}
&\int_{t_1}^{t_2} \int_\Omega \eta^{\beta+2} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |Zu^\varepsilon|^\beta \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \qquad (3.11) \\
&+ \int_\Omega \eta^{\beta+2} |Zu^\varepsilon|^\beta |\nabla_\varepsilon u^\varepsilon|^2 \Big|_{t_1}^{t_2} \\
&+ (\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Z u^\varepsilon|^2 |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2
\end{aligned}
$$

$$\leq C_\alpha(\beta+1)^2(1+|\nabla_\varepsilon\eta\|^2_{L^\infty})\int_{t_1}^{t_2}\int_\Omega(\eta^\beta+\eta^{\beta+4})(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+\frac{2\alpha}{(1+\|\nabla_\varepsilon\eta\|^2)(\beta+2)}\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+3}|\partial_t\eta|dx+\frac{\alpha}{(\beta+2)^2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+4}\Big|_{t=t_1}$$
$$+C_\Lambda(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+2}{2}}|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}$$
$$+\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^\beta|\nabla_\varepsilon u^\varepsilon|^2\partial_t(\eta^{\beta+2}).$$

Proof. Let $\eta\in C_0^\infty(\Omega\times(0,T))$ be a nonnegative cutoff function. Fix $\beta\geq 2$ and $\ell\in\{1,...,2n\}$. Note that

$$\partial_t(|X_\ell^\varepsilon u^\varepsilon|^2|Zu^\varepsilon|^\beta)=2X_\ell^\varepsilon u^\varepsilon\partial_tX_\ell^\varepsilon u^\varepsilon|Zu^\varepsilon|^\beta+\beta|X_\ell^\varepsilon u^\varepsilon|^2|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\partial_tZu^\varepsilon,$$

which suggests to use $2X_\ell^\varepsilon u^\varepsilon|Zu^\varepsilon|^\beta$ as a test function in the Eq (3.2) satisfied by $X_\ell^\varepsilon u^\varepsilon$ and to choose $\beta|X_\ell^\varepsilon u^\varepsilon|^2|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon$ as a test function in the Eq (3.3) satisfied by Zu^ε. Eq (3.2) becomes in weak form

$$\int_{t_1}^{t_2}\int_\Omega\partial_tX_\ell^\varepsilon u^\varepsilon\phi=-\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega\Big(A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon\Big)X_i^\varepsilon\phi+s_\ell Z(A^\varepsilon_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon))\phi$$
$$-\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega\Big(A^\varepsilon_{i,x_\ell}(x,\nabla_\varepsilon u^\varepsilon)-\frac{s_\ell x_{\ell+s_\ell n}}{2}A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\Big)X_i^\varepsilon\phi.$$

Consequently, if we substitute the test function $\phi=2\eta^{\beta+2}|Zu|^\beta X_\ell^\varepsilon u$, we obtain

$$\begin{aligned}
&2\int_{t_1}^{t_2}\int_\Omega\partial_tX_\ell^\varepsilon u^\varepsilon\eta^{\beta+2}|Zu^\varepsilon|^\beta X_\ell^\varepsilon u^\varepsilon \qquad (3.12)\\
&+2\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon\eta^{\beta+2}|Zu^\varepsilon|^\beta X_\ell^\varepsilon X_i^\varepsilon u^\varepsilon\\
&=-\sum_{i,j=1}^{2n+1}2\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon X_i^\varepsilon\Big(\eta^{\beta+2}|Zu^\varepsilon|^\beta\Big)X_\ell^\varepsilon u^\varepsilon\\
&-2\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon\eta^{\beta+2}|Zu^\varepsilon|^\beta[X_i^\varepsilon,X_\ell^\varepsilon]u^\varepsilon\\
&-2s_l\int_{t_1}^{t_2}\int_\Omega Z(A_{\ell+s_\ell n}(x,\nabla_\varepsilon u^\varepsilon))\eta^{\beta+2}|Zu^\varepsilon|^\beta X_\ell^\varepsilon u^\varepsilon\\
&-2\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega\Big(A^\varepsilon_{i,x_\ell}(x,\nabla_\varepsilon u^\varepsilon)-\frac{s_\ell x_{\ell+s_\ell n}}{2}A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)\Big)X_i^\varepsilon\Big(\eta^{\beta+2}|Zu^\varepsilon|^\beta X_\ell^\varepsilon u^\varepsilon\Big)\\
&=I_\ell^1+I_\ell^2+I_\ell^3+I_\ell^4.
\end{aligned}$$

We will show that these terms satisfy the following estimate

$$\sum_{k=1}^4\sum_{\ell=1}^{2n+1}|I_\ell^k|\leq\alpha\int_{t_1}^{t_2}\int_\Omega\eta^{\beta+2}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \qquad (3.13)$$

$$+ C_\alpha(\beta+1)^2(1+\|\nabla_\varepsilon\eta\|^2_{L^\infty}\int_{t_1}^{t_2}\int_\Omega(\eta^\beta+\eta^{\beta+4})(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+\frac{2\alpha}{(1+\|\nabla_\varepsilon\eta\|^2)(\beta+2)}\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+3}|\partial_t\eta|+\frac{\alpha}{(\beta+2)^2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+4}\Big|_{t=t_1}$$
$$+\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+4}|Zu^\varepsilon|^{\beta-2}|\nabla_\varepsilon Zu^\varepsilon|^2|\nabla_\varepsilon u^\varepsilon|^2.$$

We first note that

$$\sum_{\ell=1}^{2n+1}|I_\ell^1|\le 2\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega|A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon X_i\Big(\eta^{\beta+2}|Zu^\varepsilon|^\beta\Big)X_\ell^\varepsilon u^\varepsilon|$$
$$\le 2n\Lambda(\beta+2)\sum_{\ell=1}^{2n+1}\sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}|X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon|\eta^{\beta+1}|\nabla_\varepsilon\eta||\nabla_\varepsilon u^\varepsilon||Zu^\varepsilon|^\beta$$
$$+2n\beta\sum_{\ell=1}^{2n+1}\sum_{j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}|X_\ell^\varepsilon X_j^\varepsilon u^\varepsilon|\eta^{\beta+2}|Zu^\varepsilon|^{\beta-1}|\nabla_\varepsilon Zu^\varepsilon|$$
$$\le\alpha\int_{t_1}^{t_2}\int_\Omega\eta^{\beta+2}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+C_\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega\eta^\beta|\nabla_\varepsilon\eta|^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^\beta$$
$$+C_\alpha(\beta+1)^2(1+\|\nabla_\varepsilon\eta\|^2)\int_{t_1}^{t_2}\int_\Omega\eta^\beta(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+\frac{\alpha}{1+\|\nabla_\varepsilon\eta\|^2}\int_{t_1}^{t_2}\int_\Omega\eta^{\beta+4}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta|\nabla_\varepsilon Zu^\varepsilon|^2.$$

The last term can be estimated, as follows, using Lemma 3.4:

$$\alpha\int_{t_1}^{t_2}\int_\Omega\eta^{\beta+4}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta|\nabla_\varepsilon Zu^\varepsilon|^2 \tag{3.14}$$
$$\le\alpha C_{\Lambda,\lambda}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon\eta|^2\eta^{\beta+2}|Zu^\varepsilon|^{\beta+2}$$
$$+\frac{2\alpha}{\beta+2}\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+3}\partial_t\eta$$
$$+\frac{\alpha}{(\beta+1)^2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+4}\Big|_{t=t_1}+\alpha C_{\Lambda,\lambda}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^{\beta+4}|Zu^\varepsilon|^\beta$$
$$\le\alpha C_{\Lambda,\lambda}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon\eta|^2\eta^{\beta+2}|Zu^\varepsilon|^\beta\sum_{ij}|X_iX_ju|^2$$
$$+\frac{2\alpha}{\beta+2}\int_{t_1}^{t_2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+3}\partial_t\eta+\frac{\alpha}{(\beta+1)^2}\int_\Omega|Zu^\varepsilon|^{\beta+2}\eta^{\beta+4}\Big|_{t=t_1}$$

$$+\alpha C_{\Lambda,\lambda}\int_{t_1}^{t_2}\int_\Omega (\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^{\beta+4}|Zu^\varepsilon|^\beta.$$

From here estimate (3.13) holds. Integrating by parts we have

$$\sum_{\ell=1}^{2n+1}|I_\ell^2| = -2\sum_{\ell=1}^{2n+1}\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \Big|A_i^\varepsilon(x,\nabla_\varepsilon u^\varepsilon)X_l^\varepsilon\Big(\eta^{\beta+2}|Zu^\varepsilon|^\beta[X_i^\varepsilon,X_\ell^\varepsilon]u^\varepsilon\Big)\Big|$$
$$\le 2(\beta+2)\Lambda\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}\eta^{\beta+1}|\nabla_\varepsilon\eta||Zu^\varepsilon|^{\beta+1}+$$
$$+2(\beta+1)\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}\eta^{\beta+2}|Zu^\varepsilon|^\beta|\nabla_\varepsilon Zu^\varepsilon|$$
$$\le C_\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^\beta|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+\alpha\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+2}|Zu^\varepsilon|^\beta\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+\alpha\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+4}|Zu^\varepsilon|^\beta|\nabla_\varepsilon Zu^\varepsilon|^2$$
$$+C_\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^\beta|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2.$$

From here, using inequality (3.14), we deduce that I_ℓ^2 satisfies inequality (3.13). The estimate of I_ℓ^3 can be made as follows:

$$|I_\ell^3|\le\alpha\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+4}|Zu^\varepsilon|^\beta|\nabla_\varepsilon Zu^\varepsilon|^2$$
$$+C_\alpha\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}\eta^\beta|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$

From here and (3.14) the inequality (3.13) follows. The estimate of I_ℓ^4 is analogous:

$$|I_\ell^4|\le 2(\beta+1)\Lambda\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}\eta^{\beta+1}|\nabla_\varepsilon\eta||Zu^\varepsilon|^\beta|\nabla_\varepsilon u^\varepsilon|$$
$$+2(\beta+1)\Lambda\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}\eta^{\beta+2}|Zu^\varepsilon|^{\beta-1}|\nabla_\varepsilon Zu^\varepsilon||\nabla_\varepsilon u^\varepsilon|$$
$$+\Lambda\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}}\eta^{\beta+2}|Zu^\varepsilon|^\beta|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|$$
$$\le\alpha(\beta+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+4}|Zu^\varepsilon|^{\beta-2}|\nabla_\varepsilon Zu^\varepsilon|^2|\nabla_\varepsilon u^\varepsilon|^2$$
$$+\alpha\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\eta^{\beta+2}|Zu^\varepsilon|^\beta|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$

$$+ C_\alpha(\beta+1)(1+\|\nabla_\varepsilon\eta\|^2_{L^\infty})\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^\beta(\eta^\beta+\eta^{\beta+2}).$$

We now recall the following pde from Lemma 3.3

$$\partial_t Zu^\varepsilon = \sum_{i,j=1}^{2n+1} X_i^\varepsilon(A^\varepsilon_{i,\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon) + \sum_{i=1}^{2n+1} X_i^\varepsilon(A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)).$$

Substituting in this equation the test function $\phi = \beta|Zu|^{\beta-2}Zu\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2$, one obtains

$$\begin{aligned}
&\beta\int_{t_1}^{t_2}\int_\Omega \partial_t Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2 \qquad (3.15)\\
&+\beta(\beta-1)\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon X_i^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2\\
&= -\beta\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon X_i^\varepsilon\Big(\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2\Big)\\
&-\beta\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)X_i^\varepsilon\Big(|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2\Big)\\
&= -\beta(\beta+2)\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon X_i^\varepsilon\eta\eta^{\beta+1}|\nabla_\varepsilon u^\varepsilon|^2\\
&-2\beta\sum_{\ell,i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\eta^{\beta+2}X_\ell^\varepsilon u^\varepsilon X_i^\varepsilon X_\ell^\varepsilon u^\varepsilon\\
&-\beta(\beta-1)\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)|Zu^\varepsilon|^{\beta-2}X_i^\varepsilon Zu^\varepsilon\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2\\
&-\beta(\beta+1)\sum_{i=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\eta^{\beta+1}X_i^\varepsilon\eta|\nabla_\varepsilon u^\varepsilon|^2\\
&-\beta\sum_{i,\ell=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A^\varepsilon_{i,x_{2n+1}}(x,\nabla_\varepsilon u^\varepsilon)|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon\eta^{\beta+2}X_\ell^\varepsilon u^\varepsilon X_i^\varepsilon X_\ell^\varepsilon u^\varepsilon\\
&= I^5+\cdots+I^9.
\end{aligned}$$

We observe that the ellipticity condition yields

$$\begin{aligned}
&\beta(\beta-1)\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon ZuX_i^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2\\
&\geq(\beta+1)^2C_\lambda\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon Zu^\varepsilon|^2|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2.
\end{aligned}$$

Let us now consider I^5:

$$I^5 = -\beta(\beta+2)\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega A_{i\xi_j}(x,\nabla_\varepsilon u^\varepsilon)X_j^\varepsilon Zu^\varepsilon|Zu^\varepsilon|^{\beta-2}Zu^\varepsilon X_i^\varepsilon\eta\eta^{\beta+1}|\nabla_\varepsilon u^\varepsilon|^2$$

$$\leq 2(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Z u^\varepsilon| |Zu^\varepsilon|^{\beta-1} |\nabla_\varepsilon \eta| \eta^{\beta+1} |\nabla_\varepsilon u^\varepsilon|^2$$
$$\leq \alpha(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Z u^\varepsilon|^2 |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2$$
$$+ C_\alpha(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} |Zu^\varepsilon|^{\beta} |\nabla_\varepsilon \eta| \eta^{\beta}.$$

The estimate of I^6 is identical to that I^1_ℓ and we thus omit it. Let us consider I^7. One has

$$I^7 \leq (\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}} |Zu^\varepsilon|^{\beta-2} |\nabla_\varepsilon Z u^\varepsilon| \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2$$
$$\leq \alpha(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |Zu^\varepsilon|^{\beta-2} |\nabla_\varepsilon Z u^\varepsilon|^2 \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2$$
$$+ C_\alpha(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2.$$

Similar consideration holds for I^8

$$I^8 \leq (\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}} |Zu^\varepsilon|^{\beta-1} \eta^{\beta+1} |\nabla_\varepsilon \eta| |\nabla_\varepsilon u^\varepsilon|^2$$
$$\leq C_\Lambda(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} |Zu^\varepsilon|^{\beta} \eta^{\beta} |\nabla_\varepsilon \eta|^2$$
$$+ C_\Lambda(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+2}{2}} |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2}.$$

Finally, we estimate I^9.

$$I^9 \leq C_\Lambda(\beta+1) \sum_{\ell,i=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-1}{2}} |Zu^\varepsilon|^{\beta-1} \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon| |X_i^\varepsilon X_\ell^\varepsilon u^\varepsilon|$$
$$\leq C_\Lambda(\beta+1) \sum_{\ell,i=1}^{2n+1} \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2} |X_i^\varepsilon X_\ell^\varepsilon u^\varepsilon|^2.$$

It follows that

$$\sum_{k=1}^{4} \sum_{\ell=1}^{2n+1} I^k_\ell + \sum_{k=5}^{9} I^k \leq \alpha \int_{t_1}^{t_2} \int_\Omega \eta^{\beta+2} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |Zu^\varepsilon|^{\beta} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \tag{3.16}$$
$$+ C_\alpha(\beta+1)^2 (1 + |\nabla_\varepsilon \eta||^2_{L^\infty}) \int_{t_1}^{t_2} \int_\Omega (\eta^\beta + \eta^{\beta+4}) (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} |Zu^\varepsilon|^{\beta-2} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+ \frac{2\alpha}{(1 + \|\nabla_\varepsilon \eta\|^2)(\beta+2)} \int_{t_1}^{t_2} \int_\Omega |Zu^\varepsilon|^{\beta+2} \eta^{\beta+3} |\partial_t \eta| dx + \frac{\alpha}{(\beta+2)^2} \int_\Omega |Zu^\varepsilon|^{\beta+2} \eta^{\beta+4} \Big|_{t=t_1}$$
$$+ \alpha(\beta+1)^2 \int_{t_1}^{t_2} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Z u^\varepsilon|^2 |Zu^\varepsilon|^{\beta-2} (\eta^{\beta+2} + \eta^{\beta+4}) |\nabla_\varepsilon u^\varepsilon|^2$$

$$+ C_\Lambda(\beta+1)^2 \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+2}{2}} |Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}.$$

Summing up Eqs (3.12) and (3.15), we obtain

$$\sum_{i,j=1}^{2n+1}\int_{t_1}^{t_2}\int_\Omega \eta^{\beta+2}(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 + \int_\Omega (\eta^{\beta+2}|Zu^\varepsilon|^\beta |\nabla_\varepsilon u^\varepsilon|^2)\Big|_{t_1}^{t_2}$$
$$+ \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Zu^\varepsilon|^2 |Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2$$
$$= \int_{t_1}^{t_2}\int_\Omega |Zu^\varepsilon|^\beta |\nabla_\varepsilon u^\varepsilon|^2 \partial_t(\eta^{\beta+2}) + \sum_{k=1}^{4}\sum_{\ell=1}^{2n+1} I_\ell^k + \sum_{k=5}^{9} I^k.$$

Applying (3.16), the proof is completed.

□

At this point we make use of the non-degeneracy condition $\delta > 0$, and recalling that Z is obtained as a commutator of the horizontal vector fields and that $\eta \le 1$, we estimate

$$\int_{t_1}^{t_2}\int_\Omega |Zu^\varepsilon|^2\eta^3 dxdt \le C_\delta \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 \eta^2 dxdt. \tag{3.17}$$

Lemma 3.6 and (3.17) yield the following

Corollary 3.8. *Let u^ε be a weak solution of* (1.8) *in Q. For any $t_2 \ge t_1 \ge 0$, and all $\eta \in C_0^\infty(\Omega)$, such that $\eta \le 1$, $\|\partial_t\eta\| \le C\|\nabla_\varepsilon\eta\|^2$. For every fixed value of δ there exists C_δ depending on δ, p, n and on the structure constants, such that*

$$\frac{1}{2}\int_\Omega ((\delta + |\nabla_\varepsilon u^\varepsilon|^2)\eta^2)\Big|_{t_1}^{t_2} + \lambda \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2\eta^2$$
$$\le C_\delta \int_{t_1}^{t_2}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{p/2}\Big(\eta^2 + |\nabla_\varepsilon\eta|^2 + |\eta Z\eta|\Big).$$

Corollary 3.9. *Let u^ε be a solution of* (1.8) *in $\Omega\times(0,T)$ and $B_\varepsilon(x_0,r)\times(t_0-r^2,t_0)$ a parabolic cylinder. Let $\eta \in C^\infty(B_\varepsilon(x_0,r)\times(t_0-r^2,t_0))$ be a non-negative test function $\eta \le 1$, which vanishes on the parabolic boundary and such that there exists a constant $C_{\lambda,\Lambda} > 1$ for which $\|\partial_t\eta\|_{L^\infty} \le C_{\lambda,\Lambda}(1 + \|\nabla_\varepsilon\eta\|_{L^\infty}^2)$. Set $t_1 = t_0 - r^2$. There exists a constant $C_{\delta,\lambda,\Lambda}$, also depending on δ, such that for all $\beta \ge 2$ one has*

$$\int_{t_0-r^2}^{t_0}\int_\Omega \eta^{\beta+2}(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^\beta \sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 + \max_{t\in(t_0-r^2,t_0]}\int_\Omega \eta^{\beta+2}|Zu^\varepsilon|^\beta|\nabla_\varepsilon u^\varepsilon|^2 \tag{3.18}$$
$$+ (\beta+1)^2\int_{t_0-r^2}^{t_0}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|\nabla_\varepsilon Zu^\varepsilon|^2|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}|\nabla_\varepsilon u^\varepsilon|^2$$
$$\le C_{\lambda,\Lambda}(\beta+1)^2(1 + |\nabla_\varepsilon\eta\|_{L^\infty}^2)\int_{t_0-r^2}^{t_0}\int_\Omega \eta^\beta(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}|Zu^\varepsilon|^{\beta-2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$+ C_{\lambda,\Lambda}(\beta+1)^2\int_{t_0-r^2}^{t_0}\int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+2}{2}}|Zu^\varepsilon|^{\beta-2}\eta^{\beta+2}.$$

Proof. The statement follows at once by standard parabolic pde arguments, after choosing α appropriately small in (3.11) and applying (3.17), once one notes that $|Zu^\varepsilon| \leq \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|$.

□

Corollary 3.10. *In the hypotheses of the previous corollary we have*

$$\sum_{i,j=1}^{2n+1} \int_{t_0-r^2}^{t_0} \int_\Omega \eta^{\beta+2} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |Zu^\varepsilon|^\beta |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 + \max_{t\in(t_0-r^2,t_0]} \int_\Omega \eta^{\beta+2} |Zu^\varepsilon|^\beta |\nabla_\varepsilon u^\varepsilon|^2$$
$$+ (\beta+1)^2 \int_{t_0-r^2}^{t_0} \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}} |\nabla_\varepsilon Z u^\varepsilon|^2 |Zu^\varepsilon|^{\beta-2} \eta^{\beta+2} |\nabla_\varepsilon u^\varepsilon|^2$$
$$\leq C^{\beta/2} (\beta+1)^\beta (\|\nabla_\varepsilon \eta\|^2_{L^\infty} + 1))^{\beta/2} \sum_{i,j=1}^{2n+1} \int_{t_0-r^2}^{t_0} \int_\Omega \eta^\beta \Big(\delta + |\nabla_\varepsilon u^\varepsilon|^2\Big)^{\frac{p-2+\beta}{2}} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2,$$

where $c = c(n, p, L) > 0$.

Proof. In order to handle the first term in the right-hand side of the sought for conclusion, it suffices to observe that

$$C(\beta+1)^2 (\|\nabla_\varepsilon \eta\|^2_{L^\infty} + 1) \eta^\beta \Big(\delta + |\nabla_\varepsilon u^\varepsilon|^2\Big)^{p/2} |Zu|^{\beta-2} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 =$$
$$= \eta^{\beta-2} \Big(\delta + |\nabla_\varepsilon u|^2\Big)^{(p-2)(\beta-2)/2\beta} |Zu|^{\beta-2} (|\sum_{i,j=1}^{2n+1} X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2)^{(\beta-2)/\beta}$$
$$+ \eta^2 \Big(\delta + |\nabla_\varepsilon u^\varepsilon|^2\Big)^{(p+\beta-2)/\beta} (\sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2)^{2/\beta} C(\beta+1)^2 (\|\nabla_\varepsilon \eta\|^2_{L^\infty} + 1).$$

The conclusion then follows from Hölder's inequality. We can handle the second term in the same way

$$C(\beta+1)^2 (\|\nabla_\varepsilon \eta\|^2_{L^\infty}) \eta^\beta \Big(\delta + |\nabla_\varepsilon u^\varepsilon|^2\Big)^{(p+2)/2} |Zu|^{\beta-4} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$
$$= \eta^{\beta-2} \Big(\delta + |\nabla_\varepsilon u|^2\Big)^{\frac{(p-2)(\beta-4)}{2\beta}} |Zu|^{\beta-4} (|\sum_{i,j=1}^{2n+1} X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2)^{(\beta-4)/\beta}$$
$$\times \eta^2 \Big(\delta + |\nabla_\varepsilon u^\varepsilon|^2\Big)^{2(p+\beta-2)/\beta} (\sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2)^{4/\beta} C(\beta+1)^2 (\|\nabla_\varepsilon \eta\|^2_{L^\infty} + 1).$$

□

The key step in the proof of the Lipschitz regularity of solutions is the following Caccioppoli type inequality which is a parabolic analogue of [31, Theorem 3.1].

Theorem 3.11. *Let u^ε be a solution of* (1.8) *in $\Omega \times (0, T)$ and $B_\varepsilon(x_0, r) \times (t_0 - r^2, t_0)$ a parabolic cylinder. Let $\eta \in C^\infty(B_\varepsilon(x_0, r) \times (t_0 - r^2, t_0])$ be a non-negative test function $\eta \leq 1$, which vanishes on* *the parabolic boundary such that there exists a constant* $C_{\lambda,\Lambda} > 1$ *for which* $\|\partial_t\eta\|_{L^\infty} \le C_{\lambda,\Lambda}(1+\|\nabla_\varepsilon\eta\|^2_{L^\infty})$. *Set* $t_1 = t_0 - r^2, t_2 = t_0$. *There exists a constant* $C > 0$ *depending on* δ p, *and* Λ *such that for all* $\beta \ge 2$ *one has*

$$\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p-2+\beta)/2}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 dxdt + \frac{1}{\beta+2}\max_{t\in(t_0-r^2,t_0]}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta}{2}+1}\eta^2$$
$$\le C(\beta+1)^5(\|\nabla_\varepsilon\eta\|^2_{L^\infty}+\|\eta Z\eta\|_{L^\infty}+1)\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}.$$

Proof. In view of Lemma 3.5, the conclusion will follow once we provide an appropriate estimate of the term

$$\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p-2+\beta)/2}|Zu^\varepsilon|^2.$$

The first step is to apply Hölder's inequality to obtain

$$\int_{t_1}^{t_2}\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p-2+\beta)/2}|Zu^\varepsilon|^2$$
$$\le\Big(\int_{t_1}^{t_2}\int\eta^{\beta+2}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^{\beta+2}dxdt\Big)^{\frac{2}{\beta+2}}\Big(\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta}{2}}\Big)^{\frac{\beta}{\beta+2}}$$

(since $|Zu^\varepsilon| \le \sum_{i,j=1}^n |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|$)

$$\le\Big(\int_{t_1}^{t_2}\int\eta^{\beta+2}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}|Zu^\varepsilon|^{\beta}\sum_{i,j=1}^n|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2\Big)^{\frac{2}{\beta+2}}\Big(\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta}{2}}\Big)^{\frac{\beta}{\beta+2}}$$

(the first integral in the right-hand side can be bounded by applying Corollary 3.10, resulting in the estimate)

$$\le C^{\frac{\beta}{\beta+2}}(\beta+1)^{\frac{2\beta}{\beta+2}}(\|\nabla_\varepsilon\eta\|^2_{L^\infty}+1)^{\frac{\beta}{\beta+2}}\Big(\int_{t_1}^{t_2}\int_\Omega\eta^\beta\Big(\delta+|\nabla_\varepsilon u^\varepsilon|^2\Big)^{\frac{p-2+\beta}{2}}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u|^2\Big)^{\frac{2}{\beta+2}}$$
$$\times\Big(\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta}{2}}\Big)^{\frac{\beta}{\beta+2}}$$

(by Young' s inequality, recalling C_0 from the statement of Lemma 3.5)

$$\le C\frac{\beta}{\beta+2}\Big(\frac{4C_0(\beta+1)^4}{(\beta+2)}\Big)^{\frac{2}{\beta}}(\beta+1)^2(\|\nabla_\varepsilon\eta\|^2_{L^\infty}+1)\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta}{2}}$$
$$+\frac{1}{2C_0(\beta+1)^4}\int_{t_1}^{t_2}\int_\Omega\eta^\beta\Big(\delta+|\nabla_\varepsilon u^\varepsilon|^2\Big)^{\frac{p-2+\beta}{2}}\sum_{i,j=1}^{2n+1}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2.$$

Now we note that

$$\frac{\beta}{\beta+2}\Big(\frac{4C_0(\beta+1)^4}{(\beta+2)}\Big)^{\frac{2}{\beta}}(\beta+1)^2 \le C_{\lambda,\Lambda}(\beta+1)^5.$$

Substituting the previous estimate in Lemma 3.5, we conclude

$$\int_{t_1}^{t_2}\int_\Omega\eta^2(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2+\beta}{2}}|X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2$$

$$\leq C_{\lambda,\Lambda}(\beta+1)^5(\|\nabla_\varepsilon\eta\|^2_{L^\infty}+\|\eta Z\eta\|_{L^\infty}+1)\int_{t_1}^{t_2}\int_{spt(\eta)}(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta}{2}}.$$

This completes the proof of the theorem.

□

In the next result, from Lemma 2.4 and Theorem 3.11 we will establish local integrability of $\nabla_\varepsilon u^\varepsilon$ in L^q for every $q \geq p$.

Lemma 3.12. *Let u^ε be a solution of (1.8) in Q. For any open ball $B \subset\subset \Omega$ and $T > t_2 \geq t_1 \geq 0$, consider a test function $\eta \in C^\infty([0,T]\times B)$, vanishing on the parabolic boundary, such that $\eta \leq 1$, $\|\partial_t\eta\| \leq C\|\nabla_\varepsilon\eta\|^2$. For every $\beta \geq 0$, there exists a constant $C = C(n,p,\lambda,\Lambda,d(B,\partial\Omega),T-t_2,\delta) > 0$, such that*

$$\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+2)/2}|\eta|^{\beta+2} \leq C^\beta(\beta+1)^\beta\int_{t_1}^{t_2}\int_B(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{p/2}.$$

Proof. We begin by examining the case $\beta = 0$. Applying Lemma 2.4 and Corollary 3.8 one can find positive constants C_1, C_2, C_3, depending on $n, p, \lambda, \Lambda, d(B,\partial\Omega), T-t_2, \delta$, such that

$$\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+2)/2}|\eta|^2 \leq C_1(p+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p-2}{2}}\sum_{i,j}|X_j^\varepsilon X_i^\varepsilon u^\varepsilon|^2|\eta|^2$$
$$+C_2\beta^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{p/2}(|\eta|^2+|\nabla_\varepsilon\eta|^2) \leq C_3\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{p/2}\big(\eta^2+|\nabla_\varepsilon\eta|^2+|\eta Z\eta|\big),$$

concluding the proof in the case $\beta = 0$. Next, we consider the range $\beta \geq 2$. The interpolation inequality Lemma 2.4 and Theorem 3.11 imply the existence of positive constants $C_4, ..., C_7$, depending on $n, p, \lambda, \Lambda, d(B,\partial\Omega), T-t_2$, and δ, such that

$$\begin{aligned}
&\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p+2)/2}|\eta|^{\beta+2} \qquad (3.19)\\
&\leq C_4(\beta+p+1)^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p+\beta-2}{2}}\sum_{i,j}|X_j^\varepsilon X_i^\varepsilon u^\varepsilon|^2|\eta|^{\beta+2}\\
&+C_5\beta^2\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+p)/2}|\eta|^\beta(|\eta|^2+|\nabla_\varepsilon\eta|^2)\\
&\leq C_6(\beta+p+1)^7\int_{t_1}^{t_2}\int_\Omega(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}\big(\eta^2+|\nabla_\varepsilon\eta|^2+|\eta Z\eta|\big)\\
&\leq C_7(\beta+1)^7(\|\nabla_\varepsilon\eta\|^2_{L^\infty}+\|\eta Z\eta\|_{L^\infty}+1)\int_B(\delta+|\nabla_\varepsilon u^\varepsilon|^2)^{(p+\beta)/2}.
\end{aligned}$$

Iterating the latter $[\beta]/2$ times, the conclusion follows.

□

In the next result we establish Lipschitz bounds that are uniform in ε. The argument consists in implementing Moser iterations, and rests on the observation that the quantity $\delta+|\nabla_\varepsilon u^\varepsilon|^2$ is bounded from below by $\delta > 0$, and that for every $\beta \geq 0$ it is bounded in $L^{p+\beta}$ in a parabolic cylinder, uniformly in ε.

In the iteration itself, we will consider metric balls B_ε defined through the Carnot-Caratheodory metric associated to the Riemannian structure g_ε defined by the orthonormal frame $X_1^\varepsilon, ..., X_{2n+1}^\varepsilon$. We recall here that g_ε converges to the sub-Riemannian structure of the Heisenberg group in the Gromov-Hausdorff sense [19], and in particular $B_\varepsilon \to B_0$ in terms of Hausdorff distance. These considerations should make it clear that the estimates in the following theorem are stable as $\varepsilon \to 0$.

Theorem 3.13. *Let u^ε be a solution of (1.8) in $\Omega \times (0,T)$ and $Q_0^\varepsilon = B_\varepsilon(x_0, r) \times (t_0 - r^2, t_0)$ a parabolic cylinder contained in $\Omega \times (0,T)$. For given $\sigma \in (0,1)$, there exists a constant $C = C(p, \sigma, \beta_0, \lambda, \Lambda, \delta) > 0$ such that*

$$\sup_{B(x_0,\sigma r)\times(t_0-(\sigma r)^2,t_0)} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} \leq C \fint_{t_0-r^2}^{t_0} \fint_{B(x_0,r)} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}}. \tag{3.20}$$

Proof. We recall the main steps. Let us consider a family of cylinders $Q_i^\varepsilon = B_\varepsilon(x_0, r_i) \times (t_0 - r_i^2, t_0) \subset\subset Q_0^\varepsilon$ and with $r_i < r_{i-1}$. Applying (ii) in Lemma 2.3 to the function $w_\beta = (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta+2}{4}}$, one obtains

$$\begin{aligned}
&\left(\int_{t_0-r_i^2}^{t_0} \int_{B_\varepsilon(x_0,r_i)} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{(\beta+2)N_1}{2(N_1-2)}}\right)^{\frac{N_1-2}{N_1}} = \|w_\beta\|^2_{\frac{2N_1}{N_1-2},\frac{2N_1}{N_1-2},Q_i^\varepsilon} \\
&\leq \|w_\beta\|^2_{2,\infty,Q_i^\varepsilon} + \|\nabla_\varepsilon w_\beta\|^2_{2,2,Q_i^\varepsilon} \\
&\leq \int_{t_0-r_i^2}^{t_0} \int_{B_\varepsilon(x_0,r_i)} \eta^2 (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\beta/2} \sum_{i,j=1}^{2n+1} |X_i^\varepsilon X_j^\varepsilon u^\varepsilon|^2 + \frac{1}{\beta+2} \max_{t\in(t_0-r^2,t_0]} \int_{B_\varepsilon(x_0,r_i)} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{\beta+2}{2}} \eta^2.
\end{aligned}$$

Next, we set $g = (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(p-2)/2}$. Using Theorem 3.11, along with the fact that $(\delta + |\nabla_\varepsilon u^\varepsilon|) \geq \delta > 0$, we obtain

$$\left(\int_{t_0-r_i^2}^{t_0} \int_{B_\varepsilon(x_0,r_i)} (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{(\beta+2)N_1}{2(N_1-2)}}\right)^{\frac{N_1-2}{N_1}} \leq \frac{C(\beta+p)^6}{(r_i - r_{i-1})^2} \int_{t_0-r_i^2}^{t_0} \int_{B_\varepsilon(x_0,r_i)} g(\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{(\beta+2)/2}.$$

Setting $q = \frac{(\beta+2)N_1}{N_1-1}$ and $k = \frac{N_1-1}{N_1-2}$ in the latter inequality, we deduce

$$\begin{aligned}
&\left(\fint_{t_0-r_i^2}^{t_0} \fint_{B(x_0,r_i)} (\sqrt{\delta + |\nabla_\varepsilon u^\varepsilon|^2})^{qk}\right)^{\frac{1}{qk}} \\
&\leq C^{\frac{1}{\beta+2}} (\beta+p)^{\frac{6}{\beta+2}} \left(\frac{r_{i-1}^{2+N}}{r_i^N (r_i - r_{i-1})^2}\right)^{\frac{1}{2+\beta}} \left(\fint_{t_0-r_{i-1}^2}^{t_0} \fint_{B(x_0,r_{i-1})} g(\sqrt{\delta + |\nabla_\varepsilon u^\varepsilon|^2})^{\beta+2}\right)^{\frac{1}{\beta+2}} \\
&\leq C^{\frac{1}{\beta+2}} (\beta+p)^{\frac{6}{\beta+2}} \left(\frac{r_{i-1}^{2+N}}{r_i^N (r_i - r_{i-1})^2}\right)^{\frac{1}{2+\beta}} \left(\fint_{t_0-r_{i-1}^2}^{t_0} \fint_{B(x_0,r_{i-1})} (\sqrt{\delta + |\nabla_\varepsilon u^\varepsilon|^2})^{q}\right)^{\frac{1}{q}}.
\end{aligned}$$

The classical Moser iteration scheme in see [24] now applies, leading to the sought for conclusion.

□

4. Hölder regularity of derivatives of u^ε

This section focuses on the proof of the second part of Theorem 1.2. Namely, we want to prove that for each $\delta, \varepsilon > 0$ a weak solution

$$u^\varepsilon \in L^p((0,T), W^{1,p,\varepsilon}(\Omega)) \cap C^2(Q)$$

of the approximating PDE (1.8) in $Q = \Omega \times (0,T)$ satisfies the Hölder estimates

$$\|\nabla_\varepsilon u^\varepsilon\|_{C^\alpha(B\times(t_1,t_2))} + \|Zu^\varepsilon\|_{C^\alpha(B\times(t_1,t_2))} \le C\Big(\int_0^T \int_\Omega (\delta + |\nabla_\varepsilon u^\varepsilon|^2)^{\frac{p}{2}} dxdt\Big)^{\frac{1}{p}},$$

for any open ball $B \subset\subset \Omega$ and $T > t_2 \ge t_1 \ge 0$, and for some constants $C = C(n, p, \lambda, \Lambda, d(B, \partial\Omega), T - t_2, \delta) > 0$ and $\alpha = \alpha(n, p, \lambda, \Lambda, d(B, \partial\Omega), T - t_2, \delta) \in (0,1)$ independent of ε. It is clear that the above estimate represents the ε-version of (1.11).

We begin by studying the regularity of the derivatives of u^ε. In view of Lemma (3.2) and (3.3), for each $\varepsilon > 0$, $\ell = 1, \cdots, 2n+1$ all derivatives $X^\varepsilon_\ell u^\varepsilon$ and Zu^ε of u^ε satisfy the PDE

$$\partial_t w^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(\sum_{j=1}^{2n+1} a^\varepsilon_{ij}(x,t) X^\varepsilon_j w^\varepsilon + a^\varepsilon_i(x,t)\Big) + a^\varepsilon(x,t), \tag{4.1}$$

where

$$a^\varepsilon_{ij}(x,t) = A^\varepsilon_{i,\xi_j}(x, \nabla_\varepsilon u^\varepsilon),$$

$$a^\varepsilon_i(x,t) = A^\varepsilon_{i,x_\ell}(x, \nabla_\varepsilon u^\varepsilon) - \frac{s_\ell x_{\ell+s_\ell n}}{2} A^\varepsilon_{i,x_{2n+1}}(x, \nabla_\varepsilon u^\varepsilon),$$

$$a^\varepsilon(x,t) = s_\ell Z(A^\varepsilon_{\ell+s_\ell n}(x, \nabla_\varepsilon u^\varepsilon)).$$

By Lemma 3.3, Zu^ε satisfies the same equation, for $s_\ell = 0$. For every $K \subset\subset Q$, by Theorem 3.13, $|\nabla_\varepsilon u^\varepsilon|$ is bounded in K uniformly in ε. Hence $a_{i,j}$ and a_i are locally bounded in K, with ellipticity constants uniform in $\varepsilon > 0$, but dependent on δ. Precisely, there exists a constant C_0, and constants $\lambda_\delta = \delta$ and $\Lambda_\delta = \Lambda(\delta^2 + C_1^2)$ such that for every $\eta \in \mathbb{R}^{2n+1}$ and for a.e. $(x,t) \in K$, $\varepsilon > 0$, $i, j = 1, \cdots, 2n+1$

$$\|a^\varepsilon_{ij}(x,t)\|_{L^\infty(K)} + \|a^\varepsilon_i(x,t)\|_{L^\infty(K)} \le C_0 \quad \lambda_\delta|\eta|^2 \le \sum_{i,j=1}^{2n+1} a^\varepsilon_{ij}(x,t)\eta_i\eta_j \le \Lambda_\delta|\eta|^2. \tag{4.2}$$

Since $a^\varepsilon = 0$ in the equation satisfied by Zu^ε, we will then start with studying the regularity of derivatives of the solution along the center of the group.

Proposition 4.1. *Let u^ε be a solution of* (1.8) *in $\Omega \times (0,T)$ and $Q_r = B(x_0, r) \times (t_0 - r^2, t_0)$ a parabolic cylinder contained in $\Omega \times (0,T)$. There exists constants $C = C(p, \sigma, \lambda, \Lambda, \delta) > 0$ and $\alpha = \alpha(p, \sigma, \lambda, \Lambda, \delta) \in (0,1)$ such that*

$$\|Zu^\varepsilon\|_{C^\alpha(Q_{r/2})} + \|\nabla_\varepsilon Zu^\varepsilon\|_{L^2(Q_{r/2})} \le C\Big(\|u^\varepsilon\|_{L^p(Q_r)} + \|\nabla_\varepsilon u^\varepsilon\|_{L^p(Q_r)}\Big).$$

Proof. First of all, we observe that since $\delta > 0$ is fixed, Lemma 3.12 and Theorem 3.11 imply that for all $i, j = 1, ..., 2n$, one has $|X_iX_ju^\varepsilon|$ is bounded in L^2 uniformly in $\varepsilon > 0$. It follows that $Zu^\varepsilon \in L^2_{loc}(Q)$ uniformily in $\varepsilon > 0$. Since Zu^ε is a solution of (4.1), with $a^\varepsilon = 0$, the Caccioppoli inequality implies that $\nabla_\varepsilon Zu^\varepsilon$ is in $L^2_{loc}(Q)$, uniformly in $\varepsilon > 0$. The stable Harnack inequality established in [9] and [2] (see also [1, 17, 28] for the Riemannian case) yields interior Hölder estimates for w^ε in Q_r, which are stable as $\varepsilon \to 0$. □

Actually, we will prove a stronger result, in parabolic Morrey spaces. $M^{q,\alpha}(Q)$ denotes the space of all functions $f \in L^q(Q)$ such that

$$\|f\|_{M^{q,\lambda}(Q)} = \sup_{r\in S} \frac{1}{r^{\alpha-1}}\Big(r^{-N_1}\int_{\min(t_0-r^2,0)}^{t_0}\int_{B\cap\Omega}|f|^q dxdt\Big)^{1/q} < \infty, \tag{4.3}$$

where S is the set of positive radius r such that $B = B(x_0, r) \subset \Omega$, and $r^2 < t_0 < T$.

We also recall that the parabolic Campanato spaces $\mathscr{L}^{q,\alpha}(Q)$ is the collection of all $f \in L^q(Q)$ such that

$$\|f\|_{\mathscr{L}^{q,\alpha}(Q)} = \sup_{r\in S} \frac{1}{r^{\alpha}}\Big(r^{-N_1}\int_{\min(t_0-r^2,0)}^{t_0}\int_{B\cap\Omega}|f - f_{(x_0,t_0),r}|^q dxdt\Big)^{1/q} < +\infty. \tag{4.4}$$

Here, we have set

$$f_{(x_0,t_0)} = r^{-N_1}\int_{\min(t_0-r^2,0)}^{t_0}\int_{B\cap\Omega} f(x,t)dxdt.$$

Remark 4.2. Let $\alpha \in (0, 1)$ denote the Hölder exponent of Zu^ε (which is uniform in $\varepsilon > 0$). By observing that $w^\varepsilon - w^\varepsilon(x_0, t_0)$ is also a solution of (4.1), then a standard Caccioppoli type argument yields

$$\int_{t_0-r^2}^{t_0}\int_B |\nabla_\varepsilon Zu|^2 dxdt \le C\frac{1}{r^2}\int_{t_0-(2r)^2}^{t_0}\int_{2B}|Zu^\varepsilon - Zu^\varepsilon(x_0,t_0)|^2 dxdt \le Cr^{2\alpha-2}r^{N_1}, \tag{4.5}$$

where $N_1 = 2n + 4$ is the parabolic dimension, defined in (2.3).

This shows, in particular, that for every compact K cointained in Q there is a constant $C > 0$ independent of ε such that $\|\nabla_\varepsilon Zu\|_{M^{2,\alpha}(K)} \le C$, so that the coefficient a^ε in Eq (4.1) satisfies

$$\|a^\varepsilon\|_{L^2(K)} + \|a^\varepsilon\|_{M^{2,\alpha}(K)} \le C_0. \tag{4.6}$$

A standard argument, see for instance [12], shows that the Campanato space is isomorphic to the space of Hölder continuous functions. In particular, we rely on the following instance of this general result.

Lemma 4.3. *Let $K \subset\subset Q$. There exists $M, r_0 > 0$ such that for any $(x_0, t_0) \in K$ and $0 < r < r_0$, if $f \in \mathscr{L}^{q,\alpha}(B(x_0, r) \times (t_0 - r^2, t_0))$ then $f \in C^\alpha_\varepsilon(B(x_0, r/M) \times (t_0 - r^2/M^2, t_0))$.*

We need to invoke a standard result from the theory of Morrey-Campanato which adapts immediately to the Heisenberg group setting, see [6, 23].

Lemma 4.4. *For each $\varepsilon \ge 0$, let w^ε be a weak solution in a cylinder $Q = \Omega \times (0, T)$ to the Eq (4.1) with smooth coefficients. Assume that for every compact $K \subset\subset Q$ there are constants $C_0, \Lambda_\delta, \lambda_\delta, > 0$, $\alpha \in (0, 1)$ such that (4.2) and (4.6) are satisfied. Also assume that*

$$\|w^\varepsilon\|_{L^\infty(K)} + \|\nabla_\varepsilon w^\varepsilon\|_{L^2(K)} \le C_0. \tag{4.7}$$

Then for every $K \subset\subset \Omega$, there exists a constant $C > 0$ depending on $C_0, \Lambda_\delta, \lambda_\delta, \alpha$ such that $\|\nabla_\varepsilon w\|_{M^{2,\alpha}(K)} \leq C$.

Proof. Choose $r > 0$ such that the cylinder $Q_r \subset K$, and denote by z^ε the unique solution of the linear PDE, (where we omit the term a):

$$\partial_t z^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon\Big(\sum_{j=1}^{2n+1} a_{ij}^\varepsilon(x,t) X_j^\varepsilon z^\varepsilon\Big), \quad z^\varepsilon = w^\varepsilon \text{ on the parabolic boundary of } Q_r.$$

From the maximum principle $\|z^\varepsilon\|_{L^\infty(Q_r)} \leq \|w^\varepsilon\|_{L^\infty(Q_r)} \leq C_0$, by assumption. Arguing as in Remark 4.2, we see that $\|\nabla_\varepsilon z^\varepsilon\|_{M^{2,\lambda}(Q_r)} \leq C$. Choosing the test function $\varphi = w^\varepsilon - z^\varepsilon$ in the weak formulation of (4.1) we obtain

$$\int_{t_0-r^2}^{t_0}\int_B |\nabla_\varepsilon(w^\varepsilon - z^e)|^2 dxdt \leq \int_{t_0-r^2}^{t_0}\int_B a_i X_i^\varepsilon(w^\varepsilon - z^\varepsilon) dxdt + \int_{t_0-r^2}^{t_0}\int_B a(w^\varepsilon - z^\varepsilon) dxdt$$

From the hypothesis (4.7), (4.6), and using Young inequality, it immediately follows that

$$\int_{t_0-r^2}^{t_0}\int_B |\nabla_\varepsilon(w^\varepsilon - z^e)|^2 dxdt \leq Cr^{N_1} + Cr^{N_1/2}\Big(\int_{t_0-r^2}^{t_0}\int_B a^2 dxdt\Big)^{\frac{1}{2}} \leq Cr^{N_1+\alpha-1}. \tag{4.8}$$

The thesis follows from the fact that

$$\|\nabla_\varepsilon w^\varepsilon\|_{M^{2,\alpha}(Q_r)} \leq \|\nabla_\varepsilon(z^\varepsilon - w^\varepsilon)\|_{M^{2,\alpha}(Q_r)} + \|\nabla_\varepsilon z^\varepsilon\|_{M^{2,\alpha}(Q_r)} \leq 2C$$

and the right hand side is bounded independently of ε. □

Remark 4.5. If u^ε be a solution of (1.8) in $\Omega \times (0,T)$, the derivative $\partial_t u^\varepsilon$ satisfies the same equation as Zu^ε. As a result, arguing as in Remark 4.2 we deduce that for every compact K contained in Q there is a constant $C > 0$ independent of ε such that

$$\|\nabla_\varepsilon \partial_t u^\varepsilon\|_{M^{2,\alpha}(K)} \leq C$$

Proof of Theorem 1.2. For every $K \subset\subset Q$, by Theorem 3.13, there exists a constant C_0 independent of ε such that $|\nabla_\varepsilon u^\varepsilon| \leq C_0$ and Theorem 3.11 imply that for all $i, j = 1, ..., 2n$, one has $\|X_i X_j u^\varepsilon\|_{L^2} \leq C_0$. Hence the function $w_\ell = X_\ell^\varepsilon u^\varepsilon$ for every $\ell = 1, ..., 2n$ satisfies (4.7). Furthermore we already noted that it is a solution of equation (4.1) with smooth coefficients satisfying (4.2) and (4.6) uniformly in $\varepsilon > 0$. One can apply Lemma 4.4 and Remark 4.5 to conclude that $\|\partial_t X_l^\varepsilon u^\varepsilon\|_{M^{2,\alpha}} + \|\nabla_\varepsilon X_l^\varepsilon u^\varepsilon\|_{M^{2,\alpha}} \leq C$, for a suitable constant C. In view of the Poincaré inequality, and recalling that its costant is independent of ε (see [7, 9]), one then has that $\nabla_\varepsilon u^\varepsilon$ belongs to the Campanato spaces $\mathscr{L}^{2,\alpha}$. Finally, by virtue of Lemma 4.3 it follows that $\nabla_\varepsilon u^\varepsilon$ is Hölder continuous, with norm independent of ε, thus concluding the proof.

5. Proof of Theorem 1.1

We will need a simple form of the comparison principle, see [3] and [4].

Lemma 5.1. *Let u, w be weak solutions of* (1.1) *in a cylinder $B \times (t_1, t_2)$. If on the parabolic boundary $B \times \{t_1\} \cup \partial B \times (t_1, t_2)$ one has that $u \geq w$, then $u \geq w$ in $B \times (t_1, t_2)$.*

We now show how Theorem 1.1 follows from the comparison principle and from Theorem 1.2.

Proof of Theorem 1.1. Recall from Lemma 2.1 that u is Hölder continuous in any compact subdomain of Q, in particular in the closure of $B \times (t_1, t_2)$. For each $\varepsilon > 0$ consider u^ε, the unique smooth solution of the quasilinear parabolic problem

$$\begin{cases} \partial_t u^\varepsilon = \sum_{i=1}^{2n+1} X_i^\varepsilon A_i^\varepsilon(x, \nabla_\varepsilon u^\varepsilon), & \text{in } B \times (t_1, t_2) \\ u^\varepsilon = u & \text{in } B \times \{t_1\} \cup \partial B \times (t_1, t_2), \end{cases} \tag{5.1}$$

where $A_i^\varepsilon(x, \xi)$ satisfies the structure conditions (1.7). By virtue of Theorem 3.1 and of the Hölder regularity from Theorem 1.2, one has that for every $K \subset\subset Q$, and $q \geq 1$, there exist $M = M(p, q, \lambda, \Lambda, n, \delta) > 0$ and $\alpha = \alpha(p, q, \lambda, \Lambda, n, \delta) \in (0, 1)$, such that for every $\varepsilon > 0$, $(x_0, t_0) \in K$ and $B(x_0, r) \times (t_0 - r^2, t_0) \subset Q$,

$$\begin{aligned} &||\nabla_\varepsilon |\nabla_\varepsilon u^\varepsilon|^q||_{L^2(B(x_0,r)\times(t_0-r^2,t_0))} \leq M, \\ &||Z|\nabla_\varepsilon u^\varepsilon|^q||_{L^2(B(x_0,r)\times(t_0-r^2,t_0))} \leq M \\ &||\nabla_\varepsilon u^\varepsilon||_{C^\alpha_\varepsilon(B(x_0,r)\times(t_0-r^2,t_0))} + ||Zu^\varepsilon||_{C^\alpha_\varepsilon(B(x_0,r)\times(t_0-r^2,t_0))} \leq M. \end{aligned}$$

By the theorem of Ascoli-Arzelà, one can find $u_0 \in C^{1,\alpha}_{loc}(Q)$ and a sequence $\varepsilon_k \to 0$ such that

$$u^{\varepsilon_k} \to u_0 \text{ and } \nabla_{\varepsilon_k} u^{\varepsilon_k} \to \nabla_0 u_0 \text{ uniformly on compact subsets of Q.}$$

The latter implies that u_0 is a weak solution of (1.1), in $B(x_0, r) \times (t_0 - r^2, t_0)$, which agrees with the function u on the parabolic boundary of $B(x_0, r) \times (t_0 - r^2, t_0)$. By the comparison principle, the solution to this boundary values problem is unique, and hence we conclude that $u \in C^{1,\alpha}_{loc}(B(x_0, r) \times (t_0 - r^2, t_0))$. □

Acknowledgments

The first author was partially funded by NSF awards DMS 1101478, and by a Simons collaboration grant for mathematicians 585688. The second author was partially funded by Horizon 2020 Project ref. 777822: GHAIA, and by PRIN 2015 Variational and perturbative aspects of nonlinear differential problems. The third author was supported in part by a Progetto SID (Investimento Strategico di Dipartimento) "Non-local operators in geometry and in free boundary problems, and their connection with the applied sciences", University of Padova, 2017.

We thank Vira A. Markasheva, who collaborated with us on an earlier version of this project.

Conflict of interest

The authors declare no conflict of interest.

References

1. Aronson DG, Serrin J (1967) Local behavior of solutions of quasilinear parabolic equations. *Arch Ration Mech Anal* 25: 81–122.
2. Avelin B, Capogna L, Citti G, et al. (2014) Harnack estimates for degenerate parabolic equations modeled on the subelliptic p-Laplacian. *Adv Math* 257: 25–65.
3. Bieske T (2005) Comparison principle for parabolic equations in the Heisenberg group. *Electron J Differ Eq* 95: 1–11.
4. Bieske T (2006) Equivalence of weak and viscosity solutions to the p-Laplace equation in the Heisenberg group. *Ann Acad Sci Fenn Math* 31: 363–379.
5. Bramanti M, Brandolini L (2007) Schauder estimates for parabolic nondivergence operators of Hörmander type. *J Differ Equations* 234: 177–245.
6. Capogna L (1999) Regularity for quasilinear equations and 1-quasiconformal mappings in Carnot groups. *Math Ann* 313: 263–295.
7. Capogna L, Citti G (2016) Regularity for subelliptic PDE through uniform estimates in multiscale geometries. *B Math Sci* 6: 173–230.
8. Capogna L, Citti G, Ottazzi A, et al. (2019) Conformality and Q-harmonicity in sub-Riemannian manifolds. *J Math Pure Appl* 122: 67–124.
9. Capogna L, Citti G, Rea G (2013) A subelliptic analogue of Aronson-Serrin's Harnack inequality. *Math Ann* 357: 1175–1198.
10. Citti G, Manfredini M (2006) Uniform estimates of the fundamental solution for a family of hypoelliptic operators. *Potential Anal* 25: 147–164.
11. Citti G, Lanconelli E, Montanari A (2002) Smoothness of Lipschitz-continuous graphs with nonvanishing Levi curvature. *Acta Math* 188: 87–128.
12. Da Prato G (1964) Spazi $L^{p,\theta}(\Omega, \delta)$ e loro proprietà. *Annali di Matematica* 69: 383–392.
13. Di Benedetto E (1993) *Degenerate Parabolic Equations*, Springer, Universitext.
14. Domokos A (2004) Differentiability of solutions for the non-degenerate p-Laplacian in the Heisenberg group. *J Differ Equations* 204: 439–470.
15. Domokos A, Manfredi JJ (2020) $C^{1,\alpha}$-subelliptic regularity on $SU(3)$ and compact, semi-simple Lie groups. *Anal Math Phys* 10: 4.
16. Folland GB, Stein EM (1982) *Hardy Spaces on Homogeneous Groups*, Princeton: Princeton University Press.
17. Grigor'yan AA (1992) The heat equation on noncompact Riemannian manifolds. *Sb Math* 72: 47–77.
18. Hörmander L (1967) Hypoelliptic second order differential equations. *Acta Math* 119: 147–171.
19. Koranyi A (1983) Geometric aspects of analysis on the Heisenberg group, In: *Topics in Modern Harmonic Analysis, Vol. I, II (Turin/Milan, 1982)*, Francesco Severi, Rome: Ist. Naz. Alta Mat., 209-258.

20. Ladyzhenskaya O, Solonnikov VA, Ural'tseva NN (1967) *Linear and Quasilinear Parabolic Equations*, Moskow: Nauka.

21. Manfredi JJ, Mingione G (2007) Regularity results for quasilinear elliptic equations in the Heisenberg group. *Math Ann* 339: 485–544.

22. Mingione G, Zatorska-Goldstein A, Zhong X (2009) Gradient regularity for elliptic equations in the Heisenberg group. *Adv Math* 222: 62–129.

23. Morrey Jr CB (1959) Second order elliptic equations in several variables and Hölder continuity. *Math Z* 72: 146–164.

24. Moser J (1964) A Harnack inequality for parabolic differential equations. *Commun Pure Appl Math* 17: 101–134.

25. Mukherjee S (2019) On local Lipschitz regularity for quasilinear equations in the Heisenberg group. *arXiv:1804.00751*.

26. Ricciotti D (2015) *p-Laplace Equation in the Heisenberg Group*, Cham: Springer.

27. Ricciotti D (2018) On the $C^{1,\alpha}$ regularity of p-harmonic functions in the Heisenberg group. *P Am Math Soc* 146: 2937–2952.

28. Saloff-Coste L (1992) A note on Poincaré, Sobolev, and Harnack inequalities. *Int Math Res Notices* 1992: 27–38.

29. Stein EM (1993) *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton: Princeton University Press.

30. Xu CJ (1992) Regularity for quasilinear second-order subelliptic equations. *Commun Pure Appl Math* 45: 77–96.

31. Zhong X (2009) Regularity for variational problems in the Heisenberg group. *arXiv:1711.03284*.

AIMS Press