GRADIENT ESTIMATES FOR AN ORTHOTROPIC NONLINEAR DIFFUSION EQUATION IN THE HEISENBERG GROUP

MICHELE CIRCELLI

Abstract. We prove local Lipschitz regularity for weak solutions to a parabolic orthotropic p-Laplacian-type equation in the Heisenberg group $\mathbb{H}^n$, for the range $2 \leq p \leq 4$.

Contents

1. Introduction

In this paper, we establish the local Lipschitz regularity of weak solutions to a quasilinear, degenerate parabolic equation in the Heisenberg group $\mathbb{H}^n$. Specifically, we extend the results of [13] to the non-stationary setting, adapting the methodologies introduced in [8].

Given a domain Ω in the Heisenberg group $\mathbb{H}^n$ and $T > 0$, we consider the equation

$$\partial_t u = \operatorname{div}_H(Df(\nabla_H u)), \quad \text{in} \quad Q := \Omega \times (0, T), \tag{1.1}$$

where $f : \mathbb{R}^{2n} \to \mathbb{R}$ is defined as

$$f(z) = \frac{1}{p} \sum_{i=1}^{n} \left(z_i^2 + z_{i+n}^2 \right)^{\frac{p}{2}}, \tag{1.2}$$

with $p \geq 1$, and $Df = (D_1 f, \ldots, D_{2n} f)$ denotes its Euclidean $2n$-dimensional gradient. Here, $\nabla_H u = (X_1 u, \ldots, X_{2n} u)$ represents the horizontal gradient of a weak solution $u \in L^p((0,T), HW^{1,p}(\Omega))$, where $HW^{1,p}(\Omega)$ denotes the Sobolev space associated with ∇_H.

If we define

$$\lambda_i(z) = \left(z_i^2 + z_{i+n}^2\right)^{\frac{p-2}{2}}, \quad \forall i \in \{1, \dots, n\},$$
$$\lambda_i(z) = \left(z_{i-n}^2 + z_i^2\right)^{\frac{p-2}{2}}, \quad \forall i \in \{n+1, \dots, 2n\},$$

then equation (1.1) can be rewritten as

$$\partial_t u = \sum_{i=1}^{2n} X_i(\lambda_i(\nabla_H u) X_i u), \quad \text{in} \quad Q = \Omega \times (0, T). \tag{1.3}$$

We prove the local Lipschitz regularity for solutions in the range $2 \le p \le 4$ and, as byproducts, the local L^q integrability both for the non-horizontal and the time derivatives of solutions.

Our main result is the following.

Theorem (Main Theorem)**.** *Let $2 \le p \le 4$. If $u \in L^p((0,T), HW^{1,p}(\Omega))$ is a weak solution to* (1.3) *in $Q = \Omega \times (0, T)$, then*

$$\nabla_H u \in L^\infty_{loc}(Q, \mathbb{R}^{2n}),$$

and, for any $B(x_0, r) \times (t_0 - \mu\, 4r^2, t_0) \subset Q$, there exists a constant $c = c(n, p, L) > 0$ such that

$$\sup_{B(x_0,r)\times(t_0-\mu\, r,t_0)} |\nabla_H u| \le c\mu^{\frac{1}{2}} \max\left\{ \left(\int_{t_0-\mu\, 4r^2}^{t_0} \int_{B(x_0,r)} |\nabla_H u|^p \, dxdt \right)^{\frac{1}{p}}, \mu^{\frac{p}{2(2-p)}} \right\}. \tag{1.4}$$

Moreover $\partial_t u, Zu \in L^q_{loc}(\Omega)$, for any $1 \le q < \infty$.

The present paper, along with [7] and [8], represents the first instances in the literature that study higher regularity for weak solutions of non-stationary degenerate quasilinear equations in the sub-Riemannian setting. In these references, the authors investigate the regularity of solutions to equation (1.1), modeled on the p-Laplace equation, i.e. when

$$f(z) = \frac{1}{p}\left(\delta + |\nabla_H u|^2\right)^{\frac{p}{2}}.$$

In [7], the authors prove the smoothness of solutions in the non-degenerate case, namely $\delta > 0$, while in [8], they establish the Lipschitz regularity for solutions in the degenerate case ($\delta = 0$), with $2 \le p \le 4$. Both studies are based on techniques introduced by Zhong in [21], where the optimal regularity for solutions in the stationary case is addressed.

However, none of these results apply to our equation (1.1). Indeed, all of them rely on the fact that the loss of ellipticity of the operator $\mathrm{div}_H Df$ occurs only at a single point $z = 0$, whereas equation (1.3) is much more degenerate, as it degenerates in the unbounded set

$$\bigcup_{i=1}^{n} \{\lambda_i(z) = 0\} = \bigcup_{i=1}^{n} \left\{z_i^2 + z_{i+n}^2 = 0\right\} \subset \mathbb{R}^{2n},$$

which is the union of $2n - 2$ dimensional submanifolds. For this reason, we do not expect Hölder regularity of the gradient in the spatial variable, but only boundedness of the gradient. Such result, in the Euclidean setting, was proven in [4]: Theorem 1 shows that this is also the case in the setting of the Heisenberg group for the range $2 \le p \le 4$. We believe that the boundedness of the gradient is true for the entire range $1 < p < \infty$, but it requires a different technique. For a comprehensive overview of degenerate parabolic equations in the Euclidean setting, modeled on the p-Laplacian, we refer to [17].

The study of the stationary case of equation (1.3), motivated by its relation to the optimal transport problem with congestion in the Heisenberg group (see [12] and [14]), was recently addressed in [13], where the authors proved the local Lipschitz regularity for solutions for the range $p \geq 2$, adapting to the orthotropic case the techniques introduced by Zhong in [21].

Regarding the stationary case of equation (1.3) in the Euclidean setting, it is worth mentioning [5], where the Lipschitz regularity for solutions is addressed for $q \geq 2$, even in a more degenerate case, and [16] for an alternative proof based on viscosity methods. In the particular case of the plane, Bousquet and Brasco [1] proved that weak solutions are C^1 for $1 < p < \infty$. Moreover, derivatives of solutions have a logarithmic modulus of continuity: see [20] for the case $1 < q < 2$, and [19] for the case $q \geq 2$. See also [2] and [3] for the same result in the anisotropic case.

The proof of Theorem 1 relies on adapting the results from [13] to the parabolic setting and using a Poincaré-type inequality for smooth functions originally established in [15]. As a first step, we approximate equation (1.3) with a uniformly elliptic one through the Riemannian approximation of the Heisenberg group. The smoothness of solutions for this approximating equation directly follows from standard regularity results for parabolic equations.

The main ingredient in the proof of the local Lipschitzity for solutions is the Caccioppoli-type estimate for the first derivatives of approximating solutions in Theorem 6.1, which is uniform in the approximating parameters. A key result in proving this Caccioppoli-type inequality is the aforementioned Poincaré-type inequality for smooth functions (Lemma 5.3). This inequality allows us to establish a uniform local bound for the non-horizontal derivative of approximating solutions and is the only point in the paper where we impose the limited range $2 \leq p \leq 4$.

From Theorem 6.1, a local uniform bound for the gradient of the approximating solutions follows through a Moser-type iteration, leading to the Lipschitz regularity for solutions as we pass to the limit. Additionally, the local L^q integrability for the non-horizontal derivative of solutions, for any $1 < q < \infty$, easily follows by letting the approximating parameters tend to zero. Moreover, using a standard argument, where the Heisenberg setting plays no role, we also establish the local L^q integrability of the time derivative of solutions, for any $1 < q < \infty$.

The plan of the paper is as follows. In Section 2, we introduce the Heisenberg group and its Riemannian approximation, establishing the notations used throughout the paper. In Section 3, we approximate our equation with a uniformly elliptic one, using the Riemannian approximation of the Heisenberg group. Section 4 is dedicated to proving Caccioppoli-type inequalities for derivatives of approximating solutions, which involve the derivative of solutions in the non-horizontal direction. In Section 5, we establish a uniform (in the approximating parameters) integrability estimate for the vertical derivative of the approximating solutions, using a Poincaré-type inequality. In Section 6, we use the results from the previous section to prove the main Caccioppoli-type inequality for the first derivatives of approximating solutions (Theorem 1.3). This inequality is uniform in the approximating parameters and does not involve the derivative of approximating solutions in the non-horizontal direction. In Section 7, we iterate the previous Caccioppoli-type inequality using the Moser iteration scheme to obtain a local uniform bound on the L^∞ norm of the approximating solutions. Section 8 focuses on a uniform integrability estimate for the time derivative of approximating solutions. Finally, in Section 9, we prove the main theorem by letting the approximating parameters tend to zero.

2. Preliminaries

In this section we fix our notation: we introduce the Heisenebrg group $\mathbb{H}^n$ and we collect some preliminary results that will be used throughout the rest of the paper.

2.1. **The Heisenberg group.** Let $n \geq 1$, we identify the Heisenberg group $\mathbb{H}^n$ with the Euclidean space $\mathbb{R}^{2n+1}$, equipped with the group multiplication

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

for any two points $x = (x_1, \dots, x_{2n}, z), y = (y_1, \dots, y_{2n}, s) \in \mathbb{H}^n$.

The left invariant vector fields corresponding to the canonical basis of the Lie algebra

$$X_i = \partial_{x_i} - \frac{x_{n+i}}{2} \partial_t, \quad X_{n+i} = \partial_{x_{n+i}} + \frac{x_i}{2} \partial_t, \quad 1 \leq i \leq n,$$

are also called horizontal vector fields and we denote by $\nabla_H u = \sum_{i=1}^{2n} X_i u X_i \cong (X_1 u, \dots, X_{2n} u)$ the horizontal gradient of any smooth function $u : \mathbb{H}^n \to \mathbb{R}$. Given a smooth horizontal vector field $\phi = \sum_{i=1}^{2n} \phi_i X_i$, its horizontal divergence is $\mathrm{div}_H \phi = \sum_{i=1}^{2n} X_i \phi_i$.

The only non-trivial commutator

$$Z = \partial_z = [X_i, X_{n+i}] = X_i X_{n+i} - X_{n+i} X_i, \quad 1 \leq i \leq n,$$

is also called vertical vector field. We denote by $N := 2n + 2$ the homogeneous dimension of $\mathbb{H}^n$.

Let us introduce the Carnot-Carathéodory distance d. An absolutely continuous curve $\gamma : [0,1] \to \mathbb{R}^{2n+1}$ is said to be horizontal if its tangent vector $\dot{\gamma}(t) \in \mathrm{span}\,(X_1(\gamma(t)), \dots, X_{2n}(\gamma(t)))$, at almost every $t \in [0,1]$. Due to the stratification of the space the Hörmander condition is satisfied, and the Rashevsky-Chow's theorem guarantees that any couple of points can be joined with an horizontal curve [11]. It is then possible to give the following definition of distance. Given $x, y \in \mathbb{H}^n$, the Carnot-Carathéodory distance between them is defined as

$$d(x,y) := \inf \left\{ \int_0^1 |\dot{\gamma}(t)|\, dt \ : \ \gamma \text{ is horizontal, } \gamma(0) = x, \ \gamma(1) = y \right\},$$

where $|\cdot|$ denotes the norm associated with the left-invariant sub-Riemannian metric g, defined by $g\,(X_i, X_j) = \delta_{i,j}$, for $i, j = 1, \dots, 2n$.

All of the balls are defined with respect to the Carnot-Carathéodory distance:

$$B(x,r) = \{ y \in \mathbb{H}^n : d(x,y) < r \},\ x \in \mathbb{H}^n, r > 0.$$

The Haar measure in $\mathbb{H}^n$ is the Lebesgue measure of $\mathbb{R}^{2n+1}$.

If $1 \leq p < \infty$ and $\Omega \subset \mathbb{H}^n$, the horizontal Sobolev space $HW^{1,p}(\Omega)$ is the Sobolev space associated with the p-energy $\mathcal{E}_{\Omega,p}(u) = \frac{1}{p} \int_\Omega |\nabla_H u|^p\, dx$, i.e. it consists of functions $u \in L^p(\Omega)$ such that the horizontal distribution gradient $\nabla_H u$ is in $L^p(\Omega, \mathbb{R}^{2n})$.

The space $HW^{1,p}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{HW^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\nabla_H u\|_{L^p(\Omega, \mathbb{R}^{2n})}.$$

The space $HW_0^{1,p}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in $HW^{1,p}(\Omega)$ with this norm.

Definition 2.1. Let $(x_0, t_0) \in \Omega \times (0, T)$. For $r, \mu > 0$, the parabolic cylinder $Q_{\mu,r}(x_0, t_0) \subset Q$ of center (x_0, t_0) is the set

$$Q_{\mu,r}(x_0, t_0) = B(x_0, r) \times (t_0 - \mu\, r^2, t_0).$$

We call parabolic boundary of the cylinder $Q_{\mu,r}(x_0, t_0) \subset Q$ the set

$$\partial Q_{\mu,r}(x_0, t_0) := B(x_0, r) \times \{t_0 - \mu\, r^2\} \cup \partial B(x_0, r) \times [t_0 - \mu\, r^2, t_0).$$

2.2. **Riemannian approximation of the Heisenberg group.** The left-invariant sub-Riemannian structure of $\mathbb{H}^n$ arises as the pointed Hausdorff-Gromov limit of Riemannian manifolds, in which the non-horizontal direction is increasingly penalized.

Let $\varepsilon > 0$, we denote by g_ε the left-invariant Riemannian metric for which the frame defined by

$$X_1^\varepsilon = X_1, ..., X_{2n}^\varepsilon = X_{2n}, X_{2n+1}^\varepsilon = \varepsilon Z$$

is orthonormal.

It has been proved by Gromov in [18] that the left invariant Riemannian manifolds $(\mathbb{H}^n, g_\varepsilon)$ converge to the left invariant sub-Riemannian manifold $(\mathbb{H}^n, g)$, as $\varepsilon \to 0^+$, in the pointed Hausdorff-Gromov sense, i.e. the g_ε-Riemannian balls B_ε satisfy $B_\varepsilon \to B$, as $\varepsilon \to 0^+$, in the Hausdorff-Gromov sense.

If $u : \mathbb{H}^n \to \mathbb{R}$ is any smooth function, the gradient associated with the Riemannian metric g_ε is

$$\nabla_\varepsilon u := \sum_{i=1}^{2n+1} X_i^\varepsilon u X_i^\varepsilon = \sum_{i=1}^{2n} X_i u X_i + \varepsilon^2 Z u Z \cong (X_1 u, \dots, X_{2n} u, \varepsilon Z u) \,.$$

For a smooth vector field $\phi = \sum_{i=1}^{2n+1} \phi_i X_i^\varepsilon$ will also denote

$$\mathrm{div}_\varepsilon \phi = \sum_{i=1}^{2n+1} X_i^\varepsilon \phi_i.$$

Formally we have

$$\nabla_\varepsilon u \to \nabla_H u, \quad \mathrm{div}_\varepsilon \phi \to \mathrm{div}_H \phi, \qquad \text{as } \varepsilon \to 0^+.$$

We note explicitly that, again formally, we have

$$|\nabla_\varepsilon u|_\varepsilon := |\nabla_\varepsilon u|_{g_\varepsilon} = \sum_{i=1}^{2n} (X_i u)^2 + \varepsilon^2 (Zu)^2 \to |\nabla_H u|, \quad \text{as } \varepsilon \to 0^+.$$

Given $\Omega \subset \mathbb{H}^n$, we will adopt the unconventional notation $W^{1,p,\varepsilon}(\Omega)$ to indicate the Sobolev space associated with the p-energy $\mathcal{E}_{\Omega,p,\varepsilon}(u) = \frac{1}{p}\int_\Omega |\nabla_\varepsilon u|^p \, dx$, i.e. the space of functions $u \in L^p(\Omega)$ such that the distribution gradient $\nabla_\varepsilon u$ is in $L^p(\Omega, \mathbb{R}^{2n+1})$.

The space $W^{1,p,\varepsilon}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{W^{1,p,\varepsilon}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\nabla_\varepsilon u\|_{L^p(\Omega, \mathbb{R}^{2n+1})}.$$

Definition 2.2. Let $(x_0, t_0) \in \Omega \times (0, T)$. For $r, \mu > 0$, the Riemannian parabolic cylinder $Q^\varepsilon_{\mu,r}(x_0, t_0) \subset Q$ of center (x_0, t_0) is the set

$$Q^\varepsilon_{\mu,r}(x_0, t_0) = B_\varepsilon(x_0, r) \times (t_0 - \mu r^2, t_0).$$

We call parabolic boundary of the cylinder $Q^\varepsilon_{\mu,r}(x_0, t_0) \subset Q$ the set

$$\partial Q^\varepsilon_{\mu,r}(x_0, t_0) := B_\varepsilon(x_0, r) \times \{t_0 - \mu r^2\} \cup \partial B_\varepsilon(x_0, r) \times [t_0 - \mu r^2, t_0).$$

The following Sobolev embedding theorem is important for the Moser iteration.

Lemma 2.3. *Let $v \in C^\infty(Q)$ such that, for any $0 < t < T$, the function $v(\cdot, t)$ has compact support in $\Omega \times \{t\}$. Then, there exists $c = c(n) > 0$ such that, for any $\varepsilon \in [0, 1]$, one has*

$$\|v\|_{L^{\frac{2N}{N-2},2}(Q)} \le c \|\nabla_\varepsilon v\|_{L^{2,2}(Q, \mathbb{R}^{2n+1})}.$$

We note that, as ε decreases to zero, the background geometry shifts from Riemannian to sub-Riemannian but the constant C in the Lemma 2.3 is stable with respect to ε, see [6].

3. The approximating equation

The proof of Theorem 1 is based on uniform a-priori estimates for solutions of a regularized partial differential equation that approximate (1.3). The approximation procedure we introduce below is a regularization scheme widely used in literature to prove a-priori estimates for weak solutions of PDEs in the Heisenebrg group and it strongly relies on the Riemannian approximation of the Heisenebrg group, see for instance [7, 8, 9, 10] and [13].

Let $\Omega \subset \mathbb{H}^n$ be a domain and $T > 0$ and let consider the equation

$$\partial_t u = \text{div}_H(Df(\nabla_H u)), \qquad \text{in} \quad Q = \Omega \times (0,T), \tag{3.1}$$

where $f : \mathbb{R}^{2n} \to \mathbb{R}$ is the function

$$f(z) = \frac{1}{p}\sum_{i=1}^{n}\left(z_i^2 + z_{i+n}^2\right)^{\frac{p}{2}},$$

$p \geq 1$. If we denote by

$$\lambda_i(z) = \left(z_i^2 + z_{i+n}^2\right)^{\frac{p-2}{2}}, \quad \forall i \in \{1,\dots,n\},$$
$$\lambda_i(z) = \left(z_{i-n}^2 + z_i^2\right)^{\frac{p-2}{2}}, \quad \forall i \in \{n+1,\dots,2n\},$$

then

$$Df(z) = \Big(\lambda_1(z)z_1, \dots, \lambda_{2n}(z)z_{2n}\Big)$$

and therefore equation (1.1) reads as

$$\partial_t u = \sum_{i=1}^{2n} X_i(\lambda_i(\nabla_H u)X_i u), \qquad \text{in} \quad Q = \Omega \times (0,T),$$

and it satisfies the following structure condition:

$$\sum_{i=1}^{n}\lambda_i(z)\left(\xi_i^2 + \xi_{i+n}^2\right) \leq \langle D^2 f(z)\xi, \xi\rangle \leq (p-1)\sum_{i=1}^{n}\lambda_i(z)\left(\xi_i^2 + \xi_{i+n}^2\right).$$

We say that a function $u \in L^p\left((0,T), HW^{1,p}(\Omega)\right)$ is a weak solution of (3.1) if

$$\int_0^T\int_\Omega u\,\partial_t\phi\;dxdt = \int_0^T\int_\Omega\sum_{i=1}^{2n} D_i f(\nabla_H u)X_i\phi\;dxdt, \tag{3.2}$$

for every $\phi \in C_0^\infty(Q)$.

As in [13], we denote by $f_\delta : \mathbb{R}^{2n+1} \to \mathbb{R}$ the function

$$f_\delta(z) = \frac{1}{p}\sum_{i=1}^{n}\left(\delta + z_i^2 + z_{i+n}^2 + z_{2n+1}^2\right)^{\frac{p}{2}},$$

with $\delta > 0$, and we consider the Riemannian parabolic equation

$$\partial_t u = \text{div}_\varepsilon(Df_\delta(\nabla_\varepsilon u)), \tag{3.3}$$

where Df_δ denotes the Euclidean gradient of f_δ in $\mathbb{R}^{2n+1}$ and div_ε and ∇_ε are the Riemannian divergence and gradient, respectively, see Subsection 2.2.

If we denote by

$$
\begin{aligned}
&\lambda_{i,\delta}(z) := (\delta + z_i^2 + z_{i+n}^2 + z_{2n+1}^2)^{(p-2)/2}, \quad \forall i \in \{1, \dots, n\},\\
(3.4)\qquad &\lambda_{i,\delta}(z) := (\delta + z_{i-n}^2 + z_i^2 + z_{2n+1}^2)^{(p-2)/2}, \quad \forall i \in \{n+1, \dots, 2n\},\\
&\lambda_{2n+1,\delta}(z) := \sum_{i=1}^{n} \lambda_{i,\delta}(z),
\end{aligned}
$$

then the gradient of f_δ can be written as follows

$$
Df_\delta(z) = \Big(\lambda_{1,\delta}(z)z_1, \cdots, \lambda_{2n,\delta}(z)z_{2n}, \lambda_{2n+1,\delta}(z)z_{2n+1}\Big). \tag{3.5}
$$

Hence, the explicit expression of the regularized equation becomes

$$
\partial_t u = \sum_{i=1}^{2n+1} X_i^\varepsilon(\lambda_{i,\delta}(\nabla_\varepsilon u) X_i^\varepsilon u),
$$

and the structure condition

$$
\sum_{i=1}^{n} \lambda_{i,\delta}(z) \left(\xi_i^2 + \xi_{i+n}^2 + \xi_{2n+1}^2\right) \le \langle D^2 f_\delta(z)\xi, \xi\rangle \le L \sum_{i=1}^{n} \lambda_{i,\delta}(z) \left(\xi_i^2 + \xi_{i+n}^2 + \xi_{2n+1}^2\right), \tag{3.6}
$$

for any $\xi \in \mathbb{R}^{2n+1}$, where $L = L(n,p) > 1$ is a constant.

Definition 3.1 (ε-weak solution)**.** We say that a function $u^\varepsilon \in L^p\left((0,T), W^{1,p,\varepsilon}(\Omega)\right)$ is a weak solution to the equation (3.3) if

$$
\int_0^T \int_\Omega u\, \partial_t \phi \; dxdt = \int_0^T \int_\Omega \sum_{i=1}^{2n+1} D_i f_\delta(\nabla_\varepsilon u) X_i^\varepsilon \phi \; dxdt, \tag{3.7}
$$

for any $\phi \in C_0^\infty(Q)$.

Since (3.3) is strongly parabolic for every $\delta, \varepsilon > 0$, the solutions $u_{\delta,\varepsilon}$ are smooth in every compact subset $K \subset Q_0$ and they will converge uniformly on compact subsets to the solution u, see Section 9.

In the next lemma collect the partial differential equations solved by the first derivatives of approximating solutions $X_\ell^\varepsilon u_{\delta,\varepsilon}$, with $\ell \in \{1, \dots, 2n+1\}$.

Lemma 3.2. *Let u be a weak solution of* (3.3) *in Q and let us denote by $v_\ell = X_\ell^\varepsilon u_{\delta,\varepsilon}$, with $\ell = 1, \dots, 2n+1$.*
If $\ell \in \{1, \dots, 2n\}$, then the function v_ℓ solves the equation

$$
\partial_t v_\ell = \sum_{i,j=1}^{2n+1} X_i^\varepsilon \left(D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u\right) + s_\ell Z(D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)), \tag{3.8}
$$

where $s_\ell = (-1)^{\left[\frac{\ell}{n+1}\right]}$ and $[\cdot]$ denotes the floor function;
If $\ell = 2n+1$, then v_{2n+1} solves

$$
\partial_t v_{2n+1} = \sum_{i,j=1}^{2n+1} X_i^\varepsilon(D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u). \tag{3.9}
$$

For the proof see [7, Lemma]

4. A Caccioppoli-type inequality for the first derivatives of approximating solutions

The aim of the next three sections is to get higher regularity estimates for weak solutions $u_{\delta,\varepsilon}$ to (3.3) that are stable in ε and δ. Through these sections, with an abuse of notation, we will drop the indexes ε, δ and we will denote by u a weak solution to (3.3). Moreover, we denote by c a positive constant, that may vary from line to line. Except explicitly being specified, it depends only on the dimension n and on the constants p and L in the structure condition (3.6). However, it does not depend on the approximating parameters ε and δ, thus it does not degenerate as $\varepsilon, \delta \to 0$.

We start with a uniform Caccioppoli-type estimates for the first derivatives of approximating solutions u, which depends also on the vertical derivatives of such solutions Zu. The term containing the derivative zu will be removed in Theorem 6.1.

Lemma 4.1. *Let u be a weak solution of (3.3) in Q. There exists $c = c(n,p,L) > 0$ such that, for any $\beta \geq 0$ and any non-negative function $\eta \in C^1([0,T], C_0^\infty(\Omega))$, vanishing on the parabolic boundary of Q, one has*

$$\begin{aligned}
&\frac{1}{\beta+2} \sup_{0<t<T} \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}}\ dx \\
&+ \int_0^T \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right)\ dxdt \\
&\leq c(\beta+1) \int_0^T \int_\Omega \left(|\nabla_\varepsilon \eta|^2 + \eta |Z\eta|\right) (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)\ dxdt \\
&+ \frac{2}{\beta+2} \int_0^T \int_\Omega \eta |\partial_t \eta| (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}}\ dx \\
&+ c(\beta+1)^2 \int_0^T \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) |Zu|^2\ dxdt.
\end{aligned} \tag{4.1}$$

Proof. We fix $\eta \in C^1([0,T], C_0^\infty(\Omega))$ and we want to use $\phi = \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} X_\ell^\varepsilon u$, with $\ell \in \{1, \dots, 2n+1\}$ as test function in the weak formulation of the equations solved by the first derivatives, i.e. the weak formulation of (3.8) and (3.9).

If $\ell \in \{1, \dots, 2n\}$, we use ϕ as test test function in (3.8) and, through an integration by parts in the right hand side, we obtain

$$\begin{aligned}
&\frac{1}{2} \int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \partial_t \left[(X_\ell^\varepsilon u)^2\right]\ dxdt \\
&+ \sum_{i,j=1}^{2n+1} \int_0^T \int_\Omega D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_i^\varepsilon \left(\eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} X_\ell^\varepsilon u\right)\ dxdt \\
&= -s_\ell \int_0^T \int_\Omega D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u) Z \left(\eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} X_\ell^\varepsilon u\right)\ dxdt.
\end{aligned}$$

We compute $X_i^\varepsilon \left(\eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} X_\ell^\varepsilon u\right)$ in the latter equation and, using the rule

$$\sum_{i=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_i^\varepsilon X_\ell^\varepsilon u = \sum_{i=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_i^\varepsilon - s_\ell D_{\ell+s_\ell n,j}^2 f_\delta(\nabla_\varepsilon u) Zu,$$

we obtain that for every $\ell = 1, \dots, 2n$ it holds

$$
\begin{aligned}
&\frac{1}{2}\int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \partial_t \left[(X_\ell^\varepsilon u)^2\right] \, dxdt \\
&+ \int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_\ell^\varepsilon X_i^\varepsilon u \, dxdt \\
&+ \int_0^T \int_\Omega \eta^2 X_\ell^\varepsilon u \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_i^\varepsilon \left(\left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \right) \, dxdt \\
&= -2 \int_0^T \int_\Omega \eta X_\ell^\varepsilon u (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_i^\varepsilon \eta \, dxdt \\
&+ s_\ell \int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \sum_{j=1}^{2n+1} D^2_{\ell+s_\ell n, j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u Z u \, dxdt \\
&- s_\ell \int_0^T \int_\Omega D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u) Z \left(\eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} X_\ell^\varepsilon u \right) \, dxdt = I_\ell^1 + I_\ell^2 + I_\ell^3.
\end{aligned}
\tag{4.2}
$$

If $\ell = 2n+1$, we use $\phi = \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} X_\ell^\varepsilon u$ as test function in the weak formulation of (3.9). Again integrating by parts and computing the derivative we obtain

$$
\begin{aligned}
&\frac{1}{2}\int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \partial_t \left[(X_{2n+1}^\varepsilon u)^2\right] \, dxdt \\
&+ \int_0^T \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_{2n+1}^\varepsilon X_j^\varepsilon u X_{2n+1}^\varepsilon X_i^\varepsilon u \, dxdt \\
&+ \int_0^T \int_\Omega \eta^2 X_{2n+1}^\varepsilon u \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_{2n+1}^\varepsilon X_j^\varepsilon u X_i^\varepsilon \left(\left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta}{2}} \right) \, dxdt \\
&= -2 \int_0^T \int_\Omega \eta X_{2n+1}^\varepsilon u (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_{2n+1}^\varepsilon X_j^\varepsilon u X_i^\varepsilon \eta \, dxdt =: I_{2n+1}^1.
\end{aligned}
\tag{4.3}
$$

Since the left hand side is the same for all values of $\ell = 1, \dots, 2n+1$, we will handle together the last integral in it: computing the derivative and using again the rule

$$
\sum_{k=1}^{2n+1} X_i^\varepsilon X_k^\varepsilon u X_k^\varepsilon u = \sum_{k=1}^{2n+1} X_k^\varepsilon X_i^\varepsilon u X_k^\varepsilon u + s_i Z u X_{i+s_i n}^\varepsilon u,
\tag{4.4}
$$

we obtain

$$\int_0^T\int_\Omega \eta^2 X_\ell^\varepsilon u \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_i^\varepsilon \left((\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \right)\, dxdt$$

$$(4.5) \qquad = \beta \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta-2}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_\ell^\varepsilon u \sum_{k=1}^{2n+1} X_k^\varepsilon X_i^\varepsilon u X_k^\varepsilon u dx$$

$$+ s_i \beta \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta-2}{2}} X_\ell^\varepsilon u Z u \sum_{i,j=1}^{2n} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X^\varepsilon_{i+s_i n} u dx$$

where we recall that $s_i = (-1)^{[\frac{i}{n+1}]}$, $i = 1, \dots, 2n$.

We denote by I_4^ℓ the last integral in the right hand side of (4.5) and we set $I^2_{2n+1} = I^3_{2n+1} = 0$. Using (4.5) in (4.2) and (4.3) and summing up over $\ell = 1, \dots, 2n+1$ we obtain:

$$\frac{1}{\beta+2}\int_0^T\int_\Omega \eta^2 \partial_t \left[(\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}+1} \right]\, dxdt$$

$$+ \int_0^T\int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i,j,\ell=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_\ell^\varepsilon X_i^\varepsilon u \; dxdt$$

$$+ \beta \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta-2}{2}} \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) \sum_{\ell=1}^{2n+1} X_\ell^\varepsilon X_j^\varepsilon u X_\ell^\varepsilon u \sum_{k=1}^{2n+1} X_k^\varepsilon X_i^\varepsilon u X_k^\varepsilon u dx$$

$$= \sum_{\ell=1}^{2n+1} \left(I_1^\ell + I_2^\ell + I_3^\ell - I_4^\ell \right)$$

Using the structure condition (3.6) and the fact that the third term in the left-hand side is always positive, it follows that

$$\frac{1}{\beta+2}\int_0^T\int_\Omega \eta^2 \partial_t \left[(\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}+1} \right]\, dxdt$$

$$(4.6) \qquad + \int_0^T\int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^{n} \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2 \right)\, dxdt$$

$$\le \sum_{\ell=1}^{2n+1} \left(I_1^\ell + I_2^\ell + I_3^\ell + |I_4|^\ell \right).$$

The structure condition (3.6) and the Young inequality imply

$$\sum_{\ell=1}^{2n+1} I_\ell^1 \le 2\int_0^T\int_\Omega \eta (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+1}{2}} \sum_{\ell=1}^{2n+1} \left| \sum_{i,j=1}^{2n+1} D^2_{i,j} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon X_j^\varepsilon u X_i^\varepsilon \eta \right|\, dxdt$$

$$(4.7) \qquad \le \tau \int_0^T\int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^{n} \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2 \right)\, dxdt$$

$$+ \frac{c}{\tau}\int_0^T\int_\Omega |\nabla_\varepsilon \eta|^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \sum_{i=1}^{n} \lambda_{i,\delta}(\nabla_\varepsilon u)\; dxdt.$$

As for I_2^ℓ, again the structure condition (3.6) and the Young inequality implies that

$$\begin{aligned}\sum_{\ell=1}^{2n+1} I_\ell^2 \leq \tau \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right) \, dxdt \\ + \frac{c}{\tau} \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)|Zu|^2 \, dxdt.\end{aligned} \tag{4.8}$$

Let us estimate I_ℓ^3, for any $\ell \in \{1,\dots,2n+1\}$. Computing the derivative and integrating by parts, one has

$$\begin{aligned}I_\ell^3 =& -2s_\ell \int_0^T \int_\Omega \eta Z\eta D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} X_\ell^\varepsilon u \, dxdt \\ & - s_\ell \beta \int_0^T \int_\Omega \eta^2 X_\ell^\varepsilon u D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta-2}{2}} \sum_{k=1}^{2n+1} X_k^\varepsilon u X_k^\varepsilon Zu \, dxdt \\ & - s_\ell \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u) X_\ell^\varepsilon Zu \, dxdt \\ =& -2s_\ell \int_0^T \int_\Omega \eta Z\eta D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} X_\ell^\varepsilon u \, dxdt \\ & + s_\ell \beta \sum_{k=1}^{2n+1} \int_0^T \int_\Omega X_k^\varepsilon \left(\eta^2 X_\ell^\varepsilon u D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta-2}{2}} X_k^\varepsilon u\right) Zu \, dxdt \\ & + s_\ell \int_0^T \int_\Omega X_\ell^\varepsilon \left(\eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} D_{\ell+s_\ell n} f_\delta(\nabla_\varepsilon u)\right) Zu \, dxdt.\end{aligned}$$

Therefore, computing derivatives and using commutation rules analogous to (4.4) in the last two integrals, using the structure condition (3.6), one has

$$\begin{aligned}\sum_{\ell=1}^{2n+1} I_\ell^3 \leq& \tau \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right) \, dxdt \\ & + c(\beta+1) \int_0^T \int_\Omega \left(|\nabla_\varepsilon \eta|^2 + \eta|Z\eta|\right)(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \, dxdt \\ & + \frac{c(\beta+1)^2}{\tau} \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)|Zu|^2 \, dxdt.\end{aligned} \tag{4.9}$$

In the end,

$$\begin{aligned}\sum_{\ell=1}^{2n+1} |I_\ell^4| \leq& \tau \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right) \, dxdt \\ & + \frac{c(\beta+1)^2}{\tau} \int_0^T \int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)|Zu|^2 \, dxdt.\end{aligned} \tag{4.10}$$

The thesis follows putting together (4.6), (4.7), (4.8), (4.9) and (4.10), choosing for instance $\tau = \frac{1}{8}$ and using the Leibniz rule on the first term in the left hand side, eventually redefining the constant $c = c(n,p,L)$. $\square$

5. An integrability estimate for the vertical derivative of the approximating solutions

In this section we prove a uniform integrability estimate for the vertical derivative of approximating solutions Zu, see Proposition 5.3. We will use it in the next section to remove the presence of the term containing Zu in (4.1), see Theorem 6.1.

The first result we need to prove Lemma 5.3 is the following uniform and standard Caccioppoli-type inequality for Zu.

Lemma 5.1. *Let u be a weak solution of (3.3) in Q. For any $\beta \geq 0$ and any non-negative function $\eta \in C^1([0,T], C_0^\infty(\Omega))$, vanishing on the parabolic boundary of Q, one has*

$$\begin{aligned} &\int_0^T \int_\Omega \eta^2 |Zu|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left((X_i^\varepsilon Zu)^2 + (X_{n+i}^\varepsilon Zu)^2 + (X_{2n+1}^\varepsilon Zu)^2 \right) \, dxdt \\ &\leq \frac{2L^2}{(\beta+1)^2} \int_0^T \int_\Omega |\nabla_\varepsilon \eta|^2 |Zu|^{\beta+2} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \, dxdt + \frac{2}{(\beta+1)^2} \int_0^T \int_\Omega \eta |\partial_t \eta| |Zu|^{\beta+2} \, dxdt. \end{aligned} \tag{5.1}$$

Proof. We use $\phi = \eta^2 |Zu|^\beta Zu$ as a test function in the equation satisfied by Zu, i.e. (3.9) with $\varepsilon = 1$, to obtain

$$\int_0^T \int_\Omega \partial_t Zu \eta^2 |Zu|^\beta Zu \, dxdt = \int_0^T \int_\Omega \sum_{i,j=1}^{2n+1} X_i^\varepsilon (D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu) \eta^2 |Zu|^\beta Zu \, dxdt. \tag{5.2}$$

Integrating by parts, the left-hand side of (5.2) can be expressed as

$$\int_0^T \int_\Omega \eta^2 \partial_t Zu |Zu|^\beta Zu \, dxdt = \frac{1}{\beta+2} \int_0^T \int_\Omega \eta^2 \partial_t \left(|Zu|^{\beta+2} \right) \, dxdt = -\frac{2}{\beta+2} \int_0^T \int_\Omega \eta \partial_t \eta |Zu|^{\beta+2} \, dxdt.$$

An integration by parts in the right-hand side of (5.2) gives

$$\begin{aligned} &\int_0^T \int_\Omega \sum_{i,j=1}^{2n+1} X_i^\varepsilon (D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu) \eta^2 |Zu|^\beta Zu \, dxdt = \\ &\qquad - \int_0^T \int_\Omega \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu X_i^\varepsilon (\eta^2 |Zu|^\beta Zu) \, dxdt \\ &\qquad = -2 \int_0^T \int_\Omega \eta |Zu|^\beta Zu \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu X_i^\varepsilon \eta \, dxdt \\ &\qquad\qquad - (\beta+1) \int_0^T \int_\Omega \eta^2 |Zu|^\beta \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu X_i^\varepsilon Zu \, dxdt. \end{aligned}$$

Combining the previous equations we obtain

$$\begin{aligned} &(\beta+1) \int_0^T \int_\Omega \eta^2 |Zu|^\beta \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu X_i^\varepsilon Zu \, dxdt \\ &= -2 \int_0^T \int_\Omega \eta |Zu|^\beta Zu \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon Zu X_i^\varepsilon \eta \, dxdt + \frac{2}{\beta+2} \int_0^T \int_\Omega \eta \partial_t \eta |Zu|^{\beta+2} \, dxdt. \end{aligned}$$

The structure condition (3.6) and the Young inequality imply

$$\begin{aligned}
&\int_0^T\int_\Omega \eta^2|Zu|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)\left((X_i^\varepsilon Zu)^2+(X_{n+i}^\varepsilon Zu)^2+(X_{2n+1}^\varepsilon Zu)^2\right)\,dxdt\\
&\le \frac{2}{(\beta+1)}\int_0^T\int_\Omega \eta|Zu|^{\beta+1}\left|\sum_{i,j=1}^{2n+1} D^2_{i,j}f_\delta(\nabla_\varepsilon u)X_j^\varepsilon Zu X_i^\varepsilon\eta\right|dxdt+\frac{2}{(\beta+1)^2}\int_0^T\int_\Omega \eta|\partial_t\eta||Zu|^{\beta+2}\,dxdt\\
&\le \frac{\tau}{2}\int_0^T\int_\Omega \eta^2|Zu|^\beta\sum_{i=1}^n\lambda_{i,\delta}(\nabla_\varepsilon u)\left((X_i^\varepsilon Zu)^2+(X_{n+i}^\varepsilon Zu)^2+(X_{2n+1}^\varepsilon Zu)^2\right)\,dxdt\\
&+\frac{2L^2}{\tau(\beta+1)^2}\int_0^T\int_\Omega |\nabla_\varepsilon\eta|^2|Zu|^{\beta+2}\sum_{i=1}^n\lambda_{i,\delta}(\nabla_\varepsilon u)\,dxdt+\frac{2}{(\beta+1)^2}\int_0^T\int_\Omega \eta|\partial_t\eta||Zu|^{\beta+2}\,dxdt,
\end{aligned}$$

where in the last inequality we used the structure condition (3.6) and the Young inequality. The thesis follows by choosing $\tau=1$. □

A key result for the proof of Proposition 5.3 is the following Poincaré-type inequality, which was first established in [15]. This result does not depend on the equation (3.1), but it holds for any $u\in C^2_{loc}(Q)$ and it is the only point were the restriction $2\le p\le 4$ plays a role in the paper. We do not use this result directly but a key estimate in its proof. However, we state the result (in a slightly different way) for the reader's convenience.

Lemma 5.2. *Let $2\le p\le 4$ and $u\in C^2_{loc}(Q)$, and let us denote by $\nabla_{H,\ell}=(X_\ell^\varepsilon,X_{n+\ell}^\varepsilon)$, for any $\ell\in\{1,\dots,2n\}$. Then, there exists a constant $c=c(n,p)$ such that for any $\beta\ge 0$ and any non negative $\eta\in C^1([0,T],C_0^\infty(\Omega))$ vanishing on the parabolic boundary of Q, we have*

$$\begin{aligned}
\int_0^T\int_\Omega \eta^{\beta+p}|Zu|^{\beta+p}\,dxdt &\le c(\beta+p)\|\nabla_H\eta\|_\infty\int\int_{\mathrm{supp}(\eta)}(\delta+|\nabla_{H,\ell}u|^2)^{\frac{\beta+p}{2}}\,dxdt\\
&+c(\beta+p)\int_0^T\int_\Omega \eta^{4+\beta}|Zu|^\beta(\delta+|\nabla_{H,\ell}u|^2)^{\frac{p-2}{2}}|\nabla_{H,\ell}Zu|^2\,dxdt.
\end{aligned}$$

For the proof of this result see [8, Lemma 4.1].

Proposition 5.3. *Let $2\le p\le 4$ and let u be a weak solution to (3.3). Then,*

$$Zu\in L^q_{loc}(\Omega\times(0,T)),$$

for any $q\ge 1$. Moreover, there exists a constant $c=c(n,p,L)$ such that for any $\beta\ge 0$ and any non negative $\eta\in C^1([0,T],C_0^\infty(\Omega))$ vanishing on the parabolic boundary of Q, we have

$$\begin{aligned}
(5.3)\quad \left(\int_0^T\int_\Omega \eta^{\beta+p}|Zu|^{\beta+p}\,dxdt\right)^{\frac{1}{\beta+p}} &\le c(\beta+p)\|\nabla_\varepsilon\eta\|_\infty\left(\int\int_{\mathrm{supp}(\eta)}(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta+p}{2}}\,dxdt\right)^{\frac{1}{\beta+p}}\\
&+c(\beta+p)\|\eta\partial_t\eta\|_\infty^{\frac12}\,|\mathrm{supp}(\eta)|^{\frac{p-2}{2(\beta+p)}}\left(\int\int_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\,dxdt\right)^{\frac{4-p}{2(\beta+p)}}.
\end{aligned}$$

Proof. If we denote by $\nabla_{H,\ell} = (X^\varepsilon_\ell, X^\varepsilon_{n+\ell})$, for any $\ell \in \{1, \dots, 2n\}$, and by

$$\begin{aligned} I &:= \int_0^T \int_\Omega \eta^{\beta+p} |Zu|^{\beta+p}\ dxdt \\ R &:= \int\int_{\mathrm{supp}(\eta)} (\delta + |\nabla_{H,\ell} u|^2)^{\frac{\beta+p}{2}}\ dxdt \\ M &:= \int_0^T \int_\Omega \eta^{4+\beta} |Zu|^\beta (\delta + |\nabla_{H,\ell} u|^2)^{\frac{p-2}{2}} |\nabla_{H,\ell} Zu|^2\ dxdt, \end{aligned} \tag{5.4}$$

then from the proof of Lemma 5.2 it follows that

$$I \le 2(\beta+p) \left(M^{\frac{1}{2}} R^{\frac{4-p}{2(\beta+p)}} I^{\frac{2p-4+\beta}{2(\beta+p)}} + \|\nabla_{H,\ell}\eta\|_{L^\infty} R^{\frac{1}{\beta+p}} I^{\frac{\beta+p-1}{\beta+p}} \right). \tag{5.5}$$

First of all we observe that

$$R \le \int\int_{\mathrm{supp}(\eta)} (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+p}{2}}\ dxdt. \tag{5.6}$$

On the other hand, we apply the inequality (5.1) to estimate the integral M in the following way:

$$\begin{aligned} M &\le \int_0^T \int_\Omega \left(\eta^{\frac{4+\beta}{2}}\right)^2 |Zu|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left((X^\varepsilon_i Zu)^2 + (X^\varepsilon_{n+i} Zu)^2 + (X^\varepsilon_{2n+1} Zu)^2\right)\ dxdt \\ &\le c \int_0^T \int_\Omega \eta^{2+\beta} |\nabla_\varepsilon \eta|^2 |Zu|^{\beta+2} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)\ dxdt \\ &+ c \int_0^T \int_\Omega \eta^{3+\beta} |\partial_t \eta| |Zu|^{\beta+2}\ dxdt \\ &\le c \|\nabla_\varepsilon \eta\|^2_\infty\ I^{\frac{\beta+2}{\beta+p}} \left(\int\int_{\mathrm{supp}(\eta)} \left(\sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \right)^{\frac{\beta+p}{p-2}} dxdt \right)^{\frac{p-2}{\beta+p}} + c\|\eta \partial_t \eta\|_\infty\ I^{\frac{\beta+2}{\beta+p}} |\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}} \\ &\le c \|\nabla_\varepsilon \eta\|^2_\infty\ I^{\frac{\beta+2}{\beta+p}} \left(\int\int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dxdt \right)^{\frac{p-2}{\beta+p}} + c\|\eta \partial_t \eta\|_\infty\ |\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}}\ I^{\frac{\beta+2}{\beta+p}}, \end{aligned} \tag{5.7}$$

where we used the Hölder inequality and the fact that

$$\sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \le n \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{p-2}{2}}, \tag{5.8}$$

eventually changing the constant $c = c(n, p, L)$. Putting together estimates (5.5) and (5.7), and using (5.6) and the fact that $\|\nabla_{H,\ell}\eta\|_\infty \le \|\nabla_\varepsilon \eta\|_\infty$, we obtain

$$\begin{aligned} I \le c(\beta+p) \Bigg(&\|\nabla_\varepsilon \eta\|_\infty\ I^{\frac{\beta+p-1}{\beta+p}} \left(\int_0^T \int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dxdt \right)^{\frac{1}{\beta+p}} \\ &+ \|\eta \partial_t \eta\|^{\frac{1}{2}}_\infty\ |\mathrm{supp}(\eta)|^{\frac{p-2}{2(\beta+p)}}\ I^{\frac{\beta+p-1}{\beta+p}} \left(\int_0^T \int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dxdt \right)^{\frac{4-p}{2(\beta+p)}} \Bigg). \end{aligned}$$

The thesis follows by dividing both sides by the quantity $I^{\frac{\beta+p-1}{\beta+p}}$. $\square$

6. Main Caccioppoli-type estimate for the first derivatives of approximating solutions

The next theorem contains a standard Caccioppoli-type inequality for the Riemannian gradient of approximating solutions $\nabla_\varepsilon u$, where the vertical derivative Zu is not involved anymore.

Theorem 6.1. *Let $2 \le p \le 4$ and let u be a weak solution to (3.3). There exists a constant $c = c(n,p,L)$ such that for any $\beta \ge 0$ and any non-negative $\eta \in C^1([0,T], C_0^\infty(\Omega))$ vanishing on the parabolic boundary of Q, we have*

$$
\begin{aligned}
&\sup_{0<t<T} \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+2}{2}} \, dx \\
&+ \int_0^T \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right) \, dxdt \\
&\le c(\beta+p)^5 \left(\|\nabla_\varepsilon \eta\|_\infty^2 + \|\eta Z \eta\|_\infty\right) \int \int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}} \, dx \\
&+ c(\beta+p)^5 \|\eta \partial_t \eta\|_\infty |\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}} \left(\int \int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}} \, dx \right)^{\frac{\beta+2}{\beta+p}} .
\end{aligned} \tag{6.1}
$$

Proof. Lemma 4.1 implies that

$$
\begin{aligned}
&\sup_{0<t<T} \int_\Omega \eta^2 \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+2}{2}} \, dx \\
&+ \int_0^T \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left(|\nabla_\varepsilon X_i^\varepsilon u|^2 + |\nabla_\varepsilon X_{n+i}^\varepsilon u|^2 + |\nabla_\varepsilon X_{2n+1}^\varepsilon u|^2\right) \, dxdt \\
&\le c(\beta+p)^2 \int_0^T \int_\Omega \left(|\nabla_\varepsilon \eta|^2 + \eta |Z\eta|\right) (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \, dxdt \\
&+ c \int_0^T \int_\Omega \eta |\partial_t \eta| (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \, dx \\
&+ c(\beta+p)^3 \int_0^T \int_\Omega \eta^2 (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta}{2}} \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) |Zu|^2 \, dxdt =: I_1 + I_2 + I_3.
\end{aligned}
$$

It is obvious that I_1 is bounded by the first term in right hand side of (6.1). Hence, the thesis follows if the other two terms are also bounded by the right hand side of (6.1). Let us estimate I_2: the Hölder's inequality gives us

$$
\int_0^T \int_\Omega \eta |\partial_t \eta| (\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+2}{2}} \, dx \le \|\eta \partial_t \eta\|_\infty |\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}} \left(\int \int_{\mathrm{supp}(\eta)} \left(\delta + |\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}} \, dx \right)^{\frac{\beta+2}{\beta+p}},
$$

which is obviously bounded by the second term in the right hand side of (6.1). In the end, for I_3 we use the Hölder inequality, (5.3) and (5.8) to obtain

$$\int_0^T\int_\Omega \eta^2(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}}\sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u)|Zu|^2\ dxdt$$
$$\le \left(\int_0^T\int_\Omega \eta^{\beta+p}|Zu|^{\beta+p}\ dxdt\right)^{\frac{2}{\beta+p}}\left(\int\int_{\mathrm{supp}(\eta)}\left((\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta}{2}}\sum_{i=1}^n\lambda_{i,\delta}(\nabla_\varepsilon u)\right)^{\frac{\beta+p}{\beta+p-2}}\ dxdt\right)^{\frac{\beta+p-2}{\beta+p}}$$
$$\le c\left(\int_0^T\int_\Omega \eta^{\beta+p}|Zu|^{\beta+p}\ dxdt\right)^{\frac{2}{\beta+p}}\left(\int\int_{\mathrm{supp}(\eta)}(\delta+|\nabla_\varepsilon u|^2)^{\frac{\beta+p}{2}}\ dxdt\right)^{\frac{\beta+p-2}{\beta+p}}$$
$$\le c(\beta+p)^2\|\nabla_\varepsilon\eta\|_\infty^2\int\int_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dxdt$$
$$+\,c(\beta+p)^2\|\eta\partial_t\eta\|_\infty|\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}}\left(\int\int_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dxdt\right)^{\frac{\beta+2}{\beta+p}},$$

and the thesis follows. □

The following result is an easy consequence of the previous theorem.

Corollary 6.2. *Let $2\le p\le 4$ and let u be a weak solution to* (3.3)*. There exists a constant $c=c(n,p,L)$ such that for any $\beta\ge 0$ and any non-negative $\eta\in C^1([0,T],C_0^\infty(\Omega))$ vanishing on the parabolic boundary of Q, we have*

$$\sup_{0<t<T}\int_\Omega \eta^2\left(\delta+(X_k^\varepsilon u)^2\right)^{\frac{\beta+2}{2}}\ dx+\int_0^T\int_\Omega\eta^2(\delta+(X_k^\varepsilon u)^2)^{\frac{\beta+p-2}{2}}|\nabla_\varepsilon X_k^\varepsilon u|^2\ dxdt$$
$$\le c(\beta+p)^5\left(\|\nabla_\varepsilon\eta\|_\infty^2+\|\eta Z\eta\|_\infty\right)\int\int_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dx \tag{6.2}$$
$$+\,c(\beta+p)^5\|\eta\partial_t\eta\|_\infty|\mathrm{supp}(\eta)|^{\frac{p-2}{\beta+p}}\left(\int\int_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\beta+p}{2}}\ dx\right)^{\frac{\beta+2}{\beta+p}}.$$

7. Uniform Lipschitz estimate for approximating solutions

In the next theorem we establish a uniform local Lipschitz bound for approximating solutions $u=u_{\delta,\varepsilon}$: the argument is based on the Moser iteration scheme and relies the observation that the quantity $\delta+|\nabla_\varepsilon u|$ is bounded from below by δ and that for every $\beta\ge0$ it is locally bounded in $L^{\beta+p}$ in a parabolic cylinder, uniformly in ε. In the iteration, we will consider the Riemannian balls B_ε introduced in the subsection 2.2: we recall here that the balls B_ε converge to the sub-Riemannian balls in terms of Hausdorff distance and therefore the estimate in the following theorem is stable as $\delta,\varepsilon\to0$.

Theorem 7.1. *Let $2\le p\le 4$ and u be a weak solution to* (3.3) *in $\Omega\times(0,T)$. Then, for any $Q^\varepsilon_{\mu,2r}\subset Q$, it holds*

$$\|\nabla_\varepsilon u\|_{L^\infty(Q^\varepsilon_{\mu,r})}\le c\max\left\{\left(\frac{1}{\mu r^{N+2}}\int\int_{Q^\varepsilon_{\mu,2r}}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{p}{2}}\ dxdt\right)^{\frac1p},\mu^{\frac{p}{2(2-p)}}\right\}, \tag{7.1}$$

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

Proof. We consider a non negative cut-off function $\eta \in C^1([0,T], C_0^\infty(\Omega))$, vanishing on the parabolic boundary of Q, such that $|\eta| \le 1$ in Q. For any $\beta \ge 0$, we denote by

$$v_k = \eta(\delta + (X_k^\varepsilon u)^2)^{\frac{\beta+p}{4}}, \quad k = 1, \dots, 2n+1,$$
$$v = \eta(\delta + |\nabla_\varepsilon u|^2)^{\frac{\beta+p}{4}}.$$

The Caccioppoli inequality (6.2) implies

$$\begin{aligned}\sup_{0<t<T} \int_\Omega v_k^m \, dx + \int_0^T \int_\Omega |\nabla_\varepsilon v_k|^2 \, dxdt &\le c(\beta+p)^5 \left(\|\nabla_\varepsilon \eta\|_\infty^2 + \|\eta Z \eta\|_\infty\right) \int\int_{\operatorname{supp}(\eta)} v^2 \, dx \\ &+ c(\beta+p)^5 \|\eta \partial_t \eta\|_\infty |\operatorname{supp}(\eta)|^{\frac{p-2}{\beta+p}} \left(\int\int_{\operatorname{supp}(\eta)} v^2 \, dx\right)^{\frac{\beta+2}{\beta+p}},\end{aligned} \tag{7.2}$$

where $m = \frac{2(\beta+2)}{\beta+p}$. We note that $4/p < m \le 2$.

Moreover,

$$\begin{aligned}\int_0^T \int_\Omega v_k^q \, dxdt &\le \int_0^T \left(\int_\Omega v_k^m \, dx\right)^{\frac{2}{N}} \left(\int_\Omega v_k^{\frac{2N}{N-2}} \, dx\right)^{\frac{N-2}{N}} dt \\ &\le c \left(\sup_{0<t<T} \int_\Omega v_k^m \, dx\right)^{\frac{2}{N}} \left(\int_0^T \int_\Omega |\nabla_\varepsilon v_k|^2 \, dxdt\right),\end{aligned}$$

where we denoted by

$$q = \frac{2(m+N)}{N} = 2 + \frac{4(\beta+2)}{N(\beta+p)}$$

and in the second inequality we used the Sobolev inequality in the space variable, with $c = c(n)$. Raising both sides of the previous inequality to the power $\frac{N}{N+2}$, using the Young inequality, (7.2) and summing over $k = 1, \dots, 2n+1$, we get

$$\begin{aligned}\sum_{k=1}^{2n+1} \left(\int_0^T \int_\Omega v_k^q \, dxdt\right)^{\frac{N}{N+2}} &\le c(\beta+p)^5 \left(\|\nabla_\varepsilon \eta\|_\infty^2 + \|\eta Z \eta\|_\infty\right) \int\int_{\operatorname{supp}(\eta)} v^2 \, dxdt \\ &+ c(\beta+p)^5 \|\eta \partial_t \eta\|_\infty |\operatorname{supp}(\eta)|^{\frac{p-2}{\beta+p}} \left(\int\int_{\operatorname{supp}(\eta)} v^2 \, dx\right)^{\frac{\beta+2}{\beta+p}}.\end{aligned} \tag{7.3}$$

On the other hand,

$$\left(\int_0^T \int_\Omega v^q \, dxdt\right)^{\frac{N}{N+2}} \le (2n+1)^{\beta+p} \sum_{k=1}^{2n+1} \left(\int_0^T \int_\Omega v_k^q \, dxdt\right)^{\frac{N}{N+2}}, \tag{7.4}$$

then putting together (7.3) and (7.4) we get the following inequality that we will iterate:

$$\begin{aligned}\left(\int_0^T \int_\Omega v^q \, dxdt\right)^{\frac{N}{N+2}} &\le c(\beta+p)^5 \left(\|\nabla_\varepsilon \eta\|_\infty^2 + \|\eta Z \eta\|_\infty\right) \int\int_{\operatorname{supp}(\eta)} v^2 \, dxdt \\ &+ c(\beta+p)^5 \|\eta \partial_t \eta\|_\infty |\operatorname{supp}(\eta)|^{\frac{p-2}{\beta+p}} \left(\int\int_{\operatorname{supp}(\eta)} v^2 \, dx\right)^{\frac{\beta+2}{\beta+p}}.\end{aligned} \tag{7.5}$$

We consider $Q_{\mu,2r}\subset Q$ and we define a sequence of radii $r_i=(1+2^{-i})r$ and a sequence of exponents β_i, such that $\beta_0=0$ and

$$\beta_{i+1}+p=(p+\beta_i)\left(1+\frac{2(\beta_i+2)}{N(\beta_i+p)}\right),$$

that is,

$$\beta_i=2(k^i-1),\quad \text{with } k=\frac{N+2}{N}.$$

We denote by $Q_i=Q^\varepsilon_{\mu,r_i}$, so that $Q_0=Q^\varepsilon_{\mu,2r}$ and $Q_\infty=Q^\varepsilon_{\mu,r}$. Moreover we choose a standard parabolic cut-off function $\eta_i\in C^\infty(Q_i)$ such that $\eta_i=1$ in Q_{i+1} and

$$|\nabla_\varepsilon\eta_i|\le\frac{2^{i+8}}{r},\quad |Z\eta_i|\le\frac{2^{2i+8}}{r^2},\quad |\partial_t\eta_i|\le\frac{2^{2i+8}}{\mu r^2}\quad \text{in } Q_i.$$

Now we take $\eta=\eta_i$ and $\beta=\beta_i$ in (7.5) and we obtain

$$\begin{aligned}\left(\int\int_{Q_{i+1}}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\alpha_i+1}{2}}\,dxdt\right)^{\frac{N}{N+2}} &\le c2^{2i}\alpha_i^7r^{-2}\left[\left(\int\int_{Q_i}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\alpha_i}{2}}\,dxdt\right)^{\frac{p-2}{\alpha_i}}\right.\\ &\left.\quad+\mu^{-1}\left(\mu r^{N+2}\right)^{\frac{p-2}{\alpha_i}}\right]\left(\int\int_{Q_i}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\alpha_i}{2}}\,dxdt\right)^{\frac{\alpha_i-p+2}{\alpha_i}},\end{aligned}\tag{7.6}$$

where $c=c(n,p,L)>0$ and $\alpha_i=\beta_i+p=p-2+2k^i$. We denote by

$$M_i=\left(\frac{1}{\mu r^{N+2}}\int\int_{Q_i}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{\alpha_i}{2}}\,dxdt\right)^{\frac{1}{\alpha_i}}.$$

Then we can rewrite (7.6) as

$$M_{i+1}^{\frac{\alpha_i+1}{k}}\le c\mu^{\frac{2}{N+2}}2^{2i}\alpha_i^7\left(M_i^{p-2}+\mu^{-1}\right)M_i^{\alpha_i-p+2}.$$

We set

$$\overline{M}_i=\max\left\{M_i,\mu^{\frac{1}{2-p}}\right\},$$

so that the above inequality implies

$$\overline{M}_{i+1}^{\frac{\alpha_i+1}{k}}\le c\mu^{\frac{2}{N+2}}2^{2i}\alpha_i^7\overline{M}_i^{\alpha_i}\tag{7.7}$$

Iterating (7.7) we obtain

$$\overline{M}_{i+1}\le\left(\Pi_{j=0}^iK_j^{\frac{k^{i+1-j}}{\alpha_{i+1}}}\right)\overline{M}_0^{\frac{\alpha_0k^{i+1}}{\alpha_{i+1}}},$$

where

$$K_i=c\mu^{\frac{2}{N+2}}2^{2i}\alpha_i^7.$$

Letting i tend to ∞, we obtain

$$\sup_{Q^\varepsilon_{\mu,r}}|\nabla_\varepsilon u|\le\overline{M}_\infty=\limsup_{i\to\infty}\overline{M}_i\le c\mu^{\frac12}\overline{M}_0^{\frac{p}{2}},$$

where

$$\overline{M}_0=\max\left\{\left(\frac{1}{\mu r^{N+2}}\int\int_{Q^\varepsilon_{\mu,2r}}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{p}{2}}\,dxdt\right)^{\frac1p},\mu^{\frac{1}{2-p}}\right\}$$

which concludes the proof. □

8. Higher integrability for the time derivative of approximating solutions

In this section we prove a uniform local bound for the L^q norm of the time derivative of the approximating solutions $\partial_t u$, for any $q \geq 1$.

We need the following Caccioppoli-type inequality for the time derivative of approximating solutions.

Lemma 8.1. *Let u be a weak solution of (3.3) in Q. There exists $c = c(n, p, L) > 0$ such that, for any $\beta \geq 0$ and any non-negative function $\eta \in C^1([0,T], C_0^\infty(\Omega))$, vanishing on the parabolic boundary of Q, one has*

$$\begin{aligned} &\int_0^T \int_\Omega \eta^{\beta+4} |\partial_t u|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left((X_i^\varepsilon \partial_t u)^2 + (X_{i+n}^\varepsilon \partial_t u)^2 + (X_{2n+1}^\varepsilon \partial_t u)^2 \right) \, dxdt \\ &\leq \int_0^T \int_\Omega \eta^{\beta+2} |\partial_t u|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) |\nabla_\varepsilon \eta|^2 \, dxdt \\ &+ c \int_0^T \int_\Omega \eta^{\beta+3} |\partial_t \eta| \, |\partial_t u|^{\beta+2} \, dxdt. \end{aligned} \tag{8.1}$$

Proof. First of all we notice that the function $\partial_t u$ solves the following equation:

$$\partial_t (\partial_t u) = \sum_{i,j=1}^{2n+1} X_i^\varepsilon \left(D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon \partial_t u \right). \tag{8.2}$$

Let $\eta \in C^1([0,T], C_0^\infty(\Omega))$ be a non-negative cut off function vanishing on the parabolic boundary of Q, we use $\phi = \eta^{\beta+4} |\partial_t u|^\beta \partial_t u$ as test function in the weak formulation of (8.2): integrating by parts on the right hand side and dividing both sides by $\beta + 1$ we obtain

$$\begin{aligned} &\int_0^T \int_\Omega \eta^{\beta+4} |\partial_t u|^\beta \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon \partial_t u X_i^\varepsilon \partial_t u \, dxdt \\ &= -\frac{\beta+4}{\beta+1} \int_0^T \int_\Omega \eta^{\beta+3} |\partial_t u|^\beta \partial_t u \sum_{i,j=1}^{2n+1} D_{i,j}^2 f_\delta(\nabla_\varepsilon u) X_j^\varepsilon \partial_t u X_i^\varepsilon \eta \, dxdt - \frac{1}{\beta+1} \int_0^T \int_\Omega \eta^{\beta+4} |\partial_t u|^\beta \partial_t u \, \partial_t (\partial_t u) \, dxdt \\ &= I_1 + I_2. \end{aligned} \tag{8.3}$$

Using the structure condition (3.5), the left hand side of (8.3) can be bounded from below by

$$(LHS) \geq \int_0^T \int_\Omega \eta^{\beta+4} |\partial_t u|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left((X_i^\varepsilon \partial_t u)^2 + (X_{i+n}^\varepsilon \partial_t u)^2 + (X_{2n+1}^\varepsilon \partial_t u)^2 \right) \, dxdt.$$

Let us bound by above the right hand side of (8.3).

For the first term we use (3.6) and the Young inequality:

$$\begin{aligned} I_1 \leq \tau &\int_0^T \int_\Omega \eta^{\beta+4} |\partial_t u|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) \left((X_i^\varepsilon \partial_t u)^2 + (X_{i+n}^\varepsilon \partial_t u)^2 + (X_{2n+1}^\varepsilon \partial_t u)^2 \right) \, dxdt \\ + \frac{c}{\tau} &\int_0^T \int_\Omega \eta^{\beta+2} |\partial_t u|^\beta \sum_{i=1}^n \lambda_{i,\delta}(\nabla_\varepsilon u) |\nabla_\varepsilon \eta|^2 \, dxdt, \end{aligned}$$

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

Instead the second term in the right hand side can be handled in the following way

$$I_2 = -\frac{1}{(\beta+1)(\beta+2)}\int_0^T\int_\Omega \eta^{\beta+4}\,\partial_t\left(|\partial_t u|^{\beta+2}\right)\,dxdt = \frac{\beta+4}{(\beta+1)(\beta+2)}\int_0^T\int_\Omega \eta^{\beta+3}\partial_t\eta\,|\partial_t u|^{\beta+2}\,dxdt$$
$$\le c_0\int_0^T\int_\Omega \eta^{\beta+3}|\partial_t\eta|\,|\partial_t u|^{\beta+2}\,dxdt,$$

where c_0 is a fixed constant. The thesis follows by choosing $\tau = \frac{1}{2}$. □

Proposition 8.2. *Let u be a weak solution to* (3.3)*, then*

$$\partial_t u \in L^q_{loc}\left(\Omega\times(0,T)\right),$$

for any $q \ge 1$. Moreover, there exists a constant $c = c(n,p,L) > 0$ such that for any $\beta \ge 0$ and any non-negative $\eta \in C^1\left([0,T], C^\infty_0(\Omega)\right)$, vanishing on the parabolic boundary, it holds

$$\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^{\beta+2}\,dxdt \le c\,|\mathrm{supp}(\eta)|\left(M^{2p-2}\|\nabla_\varepsilon\eta\|^2_\infty + M^p\|\eta\partial_t\eta\|_\infty\right)^{\frac{\beta+2}{2}}, \tag{8.4}$$

where $M = \sup_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{1}{2}}$.

Proof. Since u solves (3.3), we can rewrite

$$|\partial_t u|^{\beta+2} = |\partial_t u|^\beta\partial_t u\sum_{i=1}^{2n+1}X_i^\varepsilon\left(D_i f_\delta(\nabla_\varepsilon u)\right). \tag{8.5}$$

From the previous identity and an integration by parts it follows

$$\begin{aligned} L :&= \int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^{\beta+2}\,dxdt = \sum_{i=1}^{2n+1}\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^\beta\partial_t u X_i^\varepsilon\left(D_i f_\delta(\nabla_\varepsilon u)\right)\,dxdt \\ &= -(\beta+2)\sum_{i=1}^{2n+1}\int_0^T\int_\Omega \eta^{\beta+1}|\partial_t u|^\beta\partial_t u D_i f_\delta(\nabla_\varepsilon u)X_i^\varepsilon\eta\,dxdt \\ &\quad -(\beta+1)\sum_{i=1}^{2n+1}\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^\beta D_i f_\delta(\nabla_\varepsilon u)X_i^\varepsilon\left(\partial_t\eta\right)\,dxdt = I_1+I_2. \end{aligned} \tag{8.6}$$

To estimate them we use (3.5) and the Hölder inequality. As for I_1, we obtain

$$\begin{aligned} |I_1| &\le (\beta+2)\int_0^T\int_\Omega \eta^{\beta+1}|\partial_t u|^{\beta+1}\left|\sum_{i=1}^{2n+1}\lambda_{i,\delta}(\nabla_\varepsilon u)X_i^\varepsilon u X_i\eta\right|\,dxdt \\ &\le (2n+1)(\beta+2)\int_0^T\int_\Omega \eta^{\beta+1}|\partial_t u|^{\beta+1}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{p-1}{2}}|\nabla_\varepsilon\eta|\,dxdt \\ &\le (2n+1)(\beta+2)\,|\mathrm{supp}(\eta)|^{\frac{1}{\beta+2}}\,\|\nabla_\varepsilon\eta\|_\infty\,M^{p-1}\,L^{\frac{\beta+1}{\beta+2}}, \end{aligned} \tag{8.7}$$

where $M = \sup_{\mathrm{supp}(\eta)}\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac{1}{2}}$.

As for I_2, we obtain

$$
\begin{aligned}
|I_2| &\le (\beta+1)\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^\beta \left|\sum_{i=1}^{2n+1}\lambda_{i,\delta}(\nabla_\varepsilon u)X_i^\varepsilon u X_i\,(\partial_t u)\right|\,dxdt\\
&\le (\beta+1)\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^\beta \left(\sum_{i=1}^{2n+1}\lambda_{i,\delta}(\nabla_\varepsilon u)(X_i^\varepsilon u)^2\right)^{\frac12}\left(\sum_{i=1}^{2n+1}\lambda_{i,\delta}(\nabla_\varepsilon u)(X_i^\varepsilon \partial_t u)^2\right)^{\frac12}\,dxdt\\
&(2n+1)^{\frac12}(\beta+1)\int_0^T\int_\Omega \eta^{\beta+2}|\partial_t u|^\beta\left(\delta+|\nabla_\varepsilon u|^2\right)^{\frac p4}\left(\sum_{i=1}^{2n+1}\lambda_{i,\delta}(\nabla_\varepsilon u)(X_i^\varepsilon \partial_t u)^2\right)^{\frac12}\,dxdt\\
&\le (2n+1)^{\frac12}(\beta+1)|\mathrm{supp}(\eta)|^{\frac{1}{\beta+2}}\;M^{\frac p2}\;L^{\frac{\beta}{2(\beta+2)}}\;J^{\frac12},
\end{aligned}
\tag{8.8}
$$

where

$$
J := \int_0^T\int_\Omega \eta^{\beta+4}|\partial_t u|^\beta \sum_{i=1}^{n}\lambda_{i,\delta}(\nabla_\varepsilon u)\left((X_i^\varepsilon\partial_t\eta)^2+(X_{i+n}^\varepsilon\partial_t\eta)^2+(X_{2n+1}^\varepsilon\partial_t\eta)^2\right)\,dxdt.
$$

From (8.1) it follows that

$$
J \le c(M^{p-2}\|\nabla_\varepsilon\eta\|_\infty^2+\|\eta\partial_t\eta\|_\infty)L, \tag{8.9}
$$

where $c=c(n,p,L)>0$. Hence, combining (8.8) and (8.9) we obtain

$$
|I_2| \le c\;(\beta+1)\;|\mathrm{supp}(\eta)|^{\frac{1}{\beta+2}}\;M^{\frac p2}\left(M^{p-2}\|\nabla_\varepsilon\eta\|_\infty^2+\|\eta\partial_t\eta\|_\infty\right)^{\frac12}L^{\frac{\beta+1}{\beta+2}}, \tag{8.10}
$$

eventually changing the constant $c=c(n,p,L)>0$.

Combining (8.6), (8.7) and (8.10) the thesis follows

□

9. Proof of the main theorem

This section is devoted to the proof of Theorem 1. The proof relies on the passage to the limit of the estimates (5.3), (7.1) and (8.4), that are all stable as the approximating parameters ε and δ approach zero.

Proof. First of all let us recall that

$$
\partial_t u \in L^{p'}\left((0,T),HW^{1,-p}(\Omega)\right),
$$

where $p'=\frac{p}{p-1}$ and $HW^{1,-p}(\Omega)$ is the dual space of $HW_0^{1,p}(\Omega)$.

Let us fix a parabolic cylinder $Q_{\mu,r}(x_0,t_0)$ such that $Q_{\mu,2r}(x_0,t_0)\subset\Omega$.

For the sake of notation in the rest of the proof we drop the centers of the balls and of the parabolic cylinders. By density, we consider a sequence of smooth functions $(\varphi_k)_{k\in\mathbb{N}}\subset C^\infty(Q_{\mu,r})$ such that $\varphi_k\to u$ in $L^p\left((0,T),HW^{1,p}(\Omega)\right)$, as $k\to\infty$, and $\partial_t\varphi_k\to\partial_t u$ in $L^{p'}\left((0,T),HW^{1,-p}(\Omega)\right)$. We denote by $u^k_{\delta,\varepsilon}$ the unique weak solution to

$$
\begin{cases}
\partial_t u^k_{\delta,\varepsilon} = \mathrm{div}_\varepsilon\left(Df_\delta(\nabla_\varepsilon u^k_{\delta,\varepsilon})\right), & \text{in } Q_{\mu,r},\\
u^k_{\delta,\varepsilon}=\varphi_k, & \text{in } \partial Q_{\mu,r}.
\end{cases}
\tag{9.1}
$$

We use $\phi = u^k_{\delta,\varepsilon} - \varphi_k$ as test function in the weak formulation of (9.1): an integration by parts in the first term gives us

$$\int\int_{Q_{\mu,r}} \partial_t u^k_{\delta,\varepsilon}\left(u^k_{\delta,\varepsilon} - \varphi_k\right)\, dxdt + \sum_{i=1}^{2n+1}\int\int_{Q_{\mu,r}} D_i f_\delta(\nabla_\varepsilon u^k_{\delta,\varepsilon}) X^\varepsilon_i u^k_{\delta,\varepsilon}\, dxdt$$
$$= \sum_{i=1}^{2n+1}\int\int_{Q_{\mu,r}} D_i f_\delta(\nabla_\varepsilon u^k_{\delta,\varepsilon}) X^\varepsilon_i \varphi_k\, dxdt.$$

We note that the first term in the left hand side can be written as

$$\int\int_{Q_{\mu,r}} \partial_t u^k_{\delta,\varepsilon}\left(u^k_{\delta,\varepsilon} - \varphi_k\right)\, dxdt = \frac{1}{2}\int_{B(x_0,r)} |u^k_{\delta,\varepsilon} - \varphi_k|^2(\cdot,t_0)\, dx + \int\int_{Q_{\mu,r}} \partial_t \varphi_k\left(u^k_{\delta,\varepsilon} - \varphi_k\right)\, dxdt,$$

and therefore we end up with

$$\frac{1}{2}\int_{B(x_0,r)} |u^k_{\delta,\varepsilon} - \varphi_k|^2(\cdot,t_0)\, dx + \sum_{i=1}^{2n+1}\int\int_{Q_{\mu,r}} D_i f_\delta(\nabla_\varepsilon u^k_{\delta,\varepsilon}) X^\varepsilon_i u^k_{\delta,\varepsilon}\, dxdt$$
$$= \sum_{i=1}^{2n+1}\int\int_{Q_{\mu,r}} D_i f_\delta(\nabla_\varepsilon u^k_{\delta,\varepsilon}) X^\varepsilon_i \varphi_k\, dxdt - \int\int_{Q_{\mu,r}} \partial_t \varphi_k\left(u^k_{\delta,\varepsilon} - \varphi_k\right)\, dxdt.$$

Using (3.4) and (3.5), we obtain

$$\frac{1}{2}\int_{B(x_0,r)} |u^k_{\delta,\varepsilon} - \varphi_k|^2(\cdot,t_0)\, dx + \int\int_{Q_{\mu,r}} \left(\delta + |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\right)^{\frac{p-2}{2}} |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\, dxdt$$
$$\leq \int\int_{Q_{\mu,r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p-1}{2}} |\nabla_\varepsilon \varphi_k|\, dxdt + \tau\int\int_{Q_{\mu,r}} |\nabla_H u^k_{\varepsilon,\delta} - \nabla_H \varphi_k|^p\, dxdt + c_\tau \|\partial_t \varphi_k\|_{L^{p'}((0,T),HW^{1,-p}(\Omega))}$$
$$\leq c\tau \int\int_{Q_{\mu,r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p}{2}}\, dxdt + c_\tau\int\int_{Q_{\mu,r}} |\nabla_H \varphi_k|^p\, dxdt + c_\tau \|\partial_t \varphi_k\|_{L^{p'}((0,T),HW^{1,-p}(\Omega))} + o(\varepsilon)$$
$$\leq c\tau \int\int_{Q_{\mu,r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p}{2}}\, dxdt + c_\tau\int\int_{Q_{\mu,r}} |\nabla_H u|^p\, dxdt + c_\tau \|\partial_t u\|_{L^{p'}((0,T),HW^{1,-p}(\Omega))} + o(\varepsilon) + o(1/k),$$

where $\tau > 0$ and $c, c_\tau > 0$ depend only on p. Moreover, from a simple computation

$$\int\int_{Q_{\mu,r}} \left(\delta + |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\right)^{\frac{p}{2}}\, dxdt = \delta\int\int_{Q_{\mu,r}} \left(\delta + |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\right)^{\frac{p-2}{2}}\, dxdt + \int\int_{Q_{\mu,r}} \left(\delta + |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\right)^{\frac{p-2}{2}} |X^\varepsilon_i u^k_{\delta,\varepsilon}|^2\, dxdt$$
$$\leq c\tau \int\int_{Q_{\mu,r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p}{2}}\, dxdt + c_\tau\int\int_{Q_{\mu,r}} |\nabla_H u|^p\, dxdt + c_\tau \|\partial_t u\|_{L^{p'}((0,T),HW^{1,-p}(\Omega))} + o(\varepsilon) + o(1/k).$$

Summing up over $i = 1, \dots, 2n+1$ and choosing τ small enough we obtain

$$\frac{1}{2}\int_{B(x_0,r)} |u^k_{\delta,\varepsilon} - \varphi_k|^2(\cdot,t_0)\, dx + \int\int_{Q_{\mu,r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p}{2}}\, dxdt \tag{9.2}$$
$$\leq c\int_{Q_{\mu,r}} |\nabla_H u|^p\, dxdt + c\|\partial_t u\|_{L^{p'}((0,T),HW^{1,-p}(\Omega))} + o(\varepsilon) + o(1/k),$$

where $c = c(n,p) > 0$. Therefore $\left(u^k_{\varepsilon,\delta}\right)_{\varepsilon,\delta,k}$ is a bounded sequence in $L^p\left((t_0 - \mu r^2, t_0), HW^{1,p}(B(x_0,r)\right)$, uniformly in ε, δ, k, and hence there is $u_0 \in L^p\left((t_0 - \mu r^2, t_0), HW^{1,p}(B(x_0,r)\right)$ such that $u^k_{\varepsilon,\delta} \rightharpoonup u_0$, up to subsequences. Since $u^k_{\varepsilon,\delta} - \varphi_k \in L^p\left((t_0-\mu r^2, t_0), HW^{1,p}_0(B(x_0,r))\right)$, passing to the limit it follows that

$$u_0 - u \in L^p\left((t_0-\mu r^2, t_0), HW^{1,p}_0(B(x_0,r))\right),$$

and therefore $u = u_0$ on the lateral boundary of $Q_{\mu,r}$. From (9.2) it also follows that $u(\cdot, t_0 - \mu r^2) = u_0(\cdot, t_0 - \mu r^2)$, eventually testing (9.1) with $\phi = \chi(t)(u^k_{\varepsilon,\delta} - \varphi_k)$, with $\chi \in C_0(t - \mu r^2, t_0)$ $\chi \equiv 1$ on $(t - \mu r^2, \bar{t})$ and $\chi \equiv 0$ on $(\bar{t}, t_0)$. The comparison principle implies that $u = u_0$ in $Q^\varepsilon_{\mu,r}$. Moreover, there exists $\tau \in (0,1)$ such that $B_\varepsilon(x_0, 2\tau r) \subset B(x_0, r)$, for $\varepsilon > 0$ small enough, and hence from Theorem 7.1 it follows that

$$\|\nabla_\varepsilon u^k_{\varepsilon,\delta}\|_{L^\infty(Q^\varepsilon_{\mu,\tau r})} \le c \max\left\{ \left(\frac{1}{\mu r^{N+2}} \int\int_{Q^\varepsilon_{\mu,2\tau r}} \left(\delta + |\nabla_\varepsilon u^k_{\varepsilon,\delta}|^2\right)^{\frac{p}{2}} \, dxdt \right)^{\frac{1}{p}}, \mu^{\frac{p}{2(2-p)}} \right\},$$

where the constant $c > 0$ is uniform in ε, δ and k. Passing to the limit we obtain

$$\|\nabla_H u\|_{L^\infty(Q_{\mu,\tau r})} \le c \max\left\{ \left(\frac{1}{\mu r^{N+2}} \int\int_{Q_{\mu,r}} |\nabla_H u|^p \, dxdt \right)^{\frac{1}{p}}, \mu^{\frac{p}{2(2-p)}} \right\}.$$

Now the estimate (1.4) follows by a simple covering argument.

Finally, Proposition 5.3 establishes the local L^q integrability of $Zu^k_{\varepsilon,\delta}$ with uniform L^q bounds, for all $1 \le q \le \infty$. This implies that, up to subsequences, $Zu^k_{\varepsilon,\delta}$ weakly converges to a L^q_{loc} function, which in view of the definition of weak derivative,is also a derivative along the center of the limit of $u^k_{\varepsilon,\delta}$, which is a solution to (1.3). The same argument also applies to the sequence $\partial_t u^k_{\varepsilon,\delta}$ for which the local L^q uniform bound (8.4) holds. □

References

[1] P. Bousquet and L. Brasco. "C^1 regularity of orthotropic p-harmonic functions in the plane". In: *Anal. PDE* 11.4 (2018), pp. 813–854.

[2] P. Bousquet and L. Brasco. "Lipschitz regularity for orthotropic functionals with nonstandard growth conditions". In: *Rev. Mat. Iberoam.* 36.7 (2020), pp. 1989–2032.

[3] P. Bousquet, L. Brasco, and C. Leone. "Singular orthotropic functionals with nonstandard growth conditions". In: *Rev. Mat. Iberoam.* 40.2 (2024), pp. 753–802.

[4] P. Bousquet et al. "Gradient estimates for an orthotropic nonlinear diffusion equation". In: *Adv. Calc. Var.* 16.3 (2023), pp. 705–730.

[5] P. Bousquet et al. "On the Lipschitz character of orthotropic p-harmonic functions". In: *Calc. Var. Partial Differential Equations* 57.3 (2018), Paper No. 88, 33.

[6] L. Capogna and G. Citti. "Regularity for subelliptic PDE through uniform estimates in multi-scale geometries". In: *Bull. Math. Sci.* 6.2 (2016), pp. 173–230.

[7] L. Capogna, G. Citti, and N. Garofalo. "Regularity for a class of quasilinear degenerate parabolic equations in the Heisenberg group". In: *Math. Eng.* 3.1 (2021), Paper No. 008, 31.

[8] L. Capogna, G. Citti, and X. Zhong. "Lipschitz regularity for solutions of the parabolic p-Laplacian in the Heisenberg group". In: *Ann. Fenn. Math.* 48.2 (2023), pp. 411–428.

[9] L. Capogna, G. Citti, and X. Zhong. "Regularity theory of quasilinear elliptic and parabolic equations in the Heisenberg group". In: *Vietnam J. Math.* 52.4 (2024), pp. 807–827.

[10] L. Capogna et al. "Conformality and Q-harmonicity in sub-Riemannian manifolds". In: *J. Math. Pures Appl. (9)* 122 (2019), pp. 67–124.

[11] W. L. Chow. “Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung”. In: *Math. Ann.* 117 (1939), pp. 98–105.

[12] M. Circelli and G. Citti. “Transport densities and congested optimal transport problem in the Heisenberg group”. In: *arXiv preprint arXiv:2303.16052* (2023).

[13] M. Circelli, G. Citti, and A. Clop. “Lipschitz regularity for solutions to an orthotropic q-Laplacian-type equation in the Heisenberg group”. In: *arXiv preprint arXiv:2407.07548* (2024).

[14] M. Circelli and A. Clop. “A continuous model of transportation in the Heisenberg group”. In: *arXiv preprint arXiv:2406.09380* (2024).

[15] G. Citti, E. Lanconelli, and A. Montanari. “Smoothness of Lipchitz-continuous graphs with nonvanishing Levi curvature”. In: *Acta Math.* 188.1 (2002), pp. 87–128.

[16] F. Demengel. “Lipschitz interior regularity for the viscosity and weak solutions of the pseudo p-Laplacian equation”. In: *Adv. Differential Equations* 21.3-4 (2016), pp. 373–400.

[17] E. DiBenedetto. *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993, pp. xvi+387.

[18] M. Gromov. *Structures métriques pour les variétés riemanniennes.* Vol. 1. Textes Mathématiques [Mathematical Texts]. Edited by J. Lafontaine and P. Pansu. CEDIC, Paris, 1981, pp. iv+152.

[19] P. Lindqvist and D. Ricciotti. “Regularity for an anisotropic equation in the plane”. In: *Nonlinear Anal.* 177 (2018), pp. 628–636.

[20] D. Ricciotti. “Regularity of the derivatives of p-orthotropic functions in the plane for $1 < p < 2$”. In: *Ann. Acad. Sci. Fenn. Math.* 44.2 (2019), pp. 1093–1099.

[21] X. Zhong. “Regularity for variational problems in the Heisenberg group”. In: *arXiv preprint arXiv:1711.03284* (2017).