published in the Journal de Mathématiques Pures et Appliquées (Ser. IX) 98 (2012) 390-427

NEW PERTURBATION METHODS FOR NONLINEAR PARABOLIC PROBLEMS

TUOMO KUUSI AND GIUSEPPE MINGIONE

ABSTRACT. We develop methods aimed at deriving regularity results for solutions to nonlinear degenerate parabolic equations and systems via local perturbation; as a consequence we obtain, in a unified way, Lipschitz continuity of solutions under weak parabolicity assumptions, and gradient continuity results in borderline cases. Nonlinear Schauder estimates as those of Misawa [29] are recovered and extended to more general settings.

RÉSUMÉ - Nous développons des méthodes dont l'objectif est d'obtenir des résultats de régularité pour les solutions d'équations et de systèmes paraboliques dégénérés non linéaires par des techniques de perturbations locales. Comme conséquence, nous obtenons, de manière unifiée, le caractère lipschitzien des solutions, sous des conditions de parabolicité faible, et des résultats de continuité du gradient dans les cas limites. Des estimations de Schauder non linéaires comme celles de Misawa [29] sont ainsi retrouvées et étendues des situations plus générales.

Keywords: Degenerate parabolic systems, regularity, intrinsic geometry

1. Introduction and results

The aim of this paper is to develop the parabolic analog of a series of regularity results that, although being rather classical in the elliptic setting, remained open in the parabolic one, mainly due to the lack of suitable perturbations techniques. We shall deal with model problems of the type

$$u_t - \operatorname{div}(\gamma(x,t)a(Du)) = -\operatorname{div} G(x,t)\,, \tag{1.1}$$

which in the particular case $a(z) = |z|^{p-2}z$ gives back the non-homogenous p-Laplacean system with coefficients

$$u_t - \operatorname{div}(\gamma(x,t)|Du|^{p-2}Du) = -\operatorname{div} G(x,t)\,. \tag{1.2}$$

A peculiarity appearing in above problems, detectable already in the case

$$u_t - \operatorname{div}(|Du|^{p-2}Du) = 0\,, \tag{1.3}$$

is the lack of homogeneity: *multiplying a solution by a constant does not yield a solution to a similar equation.* This is mainly due to different scaling properties of the evolutionary and the diffusive parts, ultimately reflecting in total lack of homogeneous a priori estimates on standard parabolic cylinders. In turn, this fact does not allow to apply standard perturbation and iteration methods which, as such, need a set of homogeneous estimates to be worked out.

On the other hand, in a couple of recent papers [20, 21], the authors succeeded in establishing new regularity techniques aimed at proving nonlinear potential estimates for solutions to nonlinear parabolic equations. In this paper we will show how the basic ideas of such techniques, when combined with new arguments, can be applied to obtain a series of regularity results which, typically dealt with via perturbation methods in the elliptic case, did not find up to now a parabolic analog – at least when $p \neq 2$. Perturbation techniques for degenerate parabolic problems have already been introduced by Misawa in [29, 30]; the ones presented here are rather different and allow, for instance, to obtain gradient boundedness and continuity results without necessarily assuming Hölder continuous coefficients unlike in [29]. This has been actually a common point in almost all classical perturbation techniques, even in the elliptic case: Hölder continuous coefficients are used to prove first a Morrey regularity result for Du, and then its Hölder continuity. The method exploited here allows instead for a more direct approach, catching those borderline regularity estimates unreachable otherwise. We summarize three basic type of results:

- Local gradient boundedness for solutions u to systems which are not everywhere parabolic, but rather become parabolic only in an asymptotic sense, i.e. for large values of the gradient norm $|Du|$. Known in the elliptic case, the extension to the parabolic case of the available elliptic techniques has not been found. This is basically due to the above mentioned lack of homogeneous estimates.

- Continuity of Du when space variable coefficients are Dini continuous. This is also a classical result in the elliptic case, while the available parabolic techniques do not seem to catch this borderline case.

- Hölder continuity results for Du when coefficients are themselves Hölder continuous. This fact, originally obtained by Misawa for the p-Laplacean system, allows to recover the results obtained, by means of a different type of perturbation methods, by Misawa [29, 30] himself and Manfredi [24, 25] for the elliptic case. The result is here valid for general parabolic equations and quasi-diagonal parabolic systems.

We shall very often deal with model problems for the sake of brevity, eventually providing the indications for more general extensions.

1.1. **Asymptotic regularity.** We start with the missing parabolic version of certain classical elliptic results which have been extensively developed over the last years; see for instance [7, 32, 11, 12, 23, 33] and related references. These results, in the standard elliptic version, amount to prove the Lipschitz regularity of solutions to elliptic systems of the type $\operatorname{div} a(Du) = 0$, with $u\colon \Omega \to \mathbb{R}^N$, under the main assumption that the vector field $a\colon \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ is asymptotically close, in C^1-sense, to the regular vector field $|z|^{p-2}z$; see (1.6) below. The heuristic of the proof of this result is rather natural: either the gradient stays bounded, and in this case there is nothing to prove. Otherwise $|Du|$ must be assumed to be very large. But then, in this last case, the vector field $a(Du)$ is close enough to $|Du|^{p-2}Du$ and this means that Du almost solves the p-Laplacean system, and therefore is still bounded. The rigorous implementation of such alternatives is of course far from being straightforward. Let us remark that asymptotic regularity results of the type just described are often crucial in establishing dimension estimates for singular sets of solutions to elliptic system (see for instance [17, 18, 27] and the recent survey [28] for a general overview) and in several problems coming from mathematical materials science (see for instance the interesting applications to the integrability of minimizing gradient Young measures in [11]).

The first result of this paper shows that such a parabolic version of the classical elliptic results actually holds. Specifically we consider a model problem of the type in (1.1), considered in the cylindrical domain $\Omega_T = \Omega \times (-T, 0)$ where $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is a bounded domain and $T > 0$. The solution u is in general a vector valued map

$$u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N)) \cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))\,, \qquad N \geq 1 \tag{1.4}$$

and solves (1.1) in the distributional sense

$$\int_{\Omega_T} \Big(-u\varphi_t + \langle \gamma(x,t)a(Du), D\varphi\rangle \Big)\, dx\, dt = \int_{\Omega_T} \langle G, D\varphi\rangle\, dx\, dt$$

whenever $\varphi \in C_c^\infty(\Omega_T,\mathbb{R}^N)$. In this section we make no other assumption on the C^1-vector field $a(\cdot)$ than

$$|a(z)| + |\partial a(z)|(|z|+1) \leq L(|z|+1)^{p-1}\,, \tag{1.5}$$

which has to hold whenever $z \in \mathbb{R}^{Nn}$, and the following C^1-asymptotic closeness condition:

$$\lim_{|z|\to\infty} \frac{|\partial a(z) - \partial b(z)|}{|z|^{p-2}} = 0\,, \qquad \text{where } \ b(z) := |z|^{p-2}z\,. \tag{1.6}$$

In particular, we are not assuming that the system considered is parabolic in that parabolicity only holds at infinity. Here, as in the rest of the paper, we shall always assume the standard lower bound

$$\frac{2n}{n+2} < p \tag{1.7}$$

that is in fact necessary to obtain all the regularity results stated below, and already in the case of solutions to the model case (1.3) (see [9, 1]). As for the function $\gamma(\cdot)$ and the map $G(\cdot)$, we assume that they are measurable and satisfy the non-degeneracy conditions

$$0 < \nu \leq \gamma(\cdot) \leq L\,. \tag{1.8}$$

We shall assume that the partial maps $x \mapsto \gamma(x,\cdot)$ and $x \mapsto G(x,\cdot)$ are Dini continuous in a suitable sense. More precisely, by defining the modulus of continuity

$$\omega(\varrho) := \sup_{\substack{t\in(-T,0),\, x,y\in B_\varrho \\ B_\varrho \subset \Omega}} |\gamma(x,t)-\gamma(y,t)| + |G(x,t)-G(y,t)|^{\min\{1,p/[2(p-1)]\}}\,,$$

we assume that

$$\int_0 \omega(\varrho)\, \frac{d\varrho}{\varrho} < \infty\,. \tag{1.9}$$

We then have

Theorem 1.1 (Asymptotic regularity)**.** *Let u be a solution to* (1.1) *under the assumptions* (1.5)-(1.9)*; then $Du \in L^\infty_{\mathrm{loc}}(\Omega_T)$. Moreover, there exists a constant c depending only on n, N, p, ν, L and the rate of convergence in* (1.6) *such that*

$$|Du(x_0,t_0)| \leq c \left[⨍_{Q_r(x_0,t_0)} (|Du|^p + 1)\, dx\, dt \right]^{d/p} \tag{1.10}$$

holds whenever $Q_r(x_0,t_0) \subset \Omega_T$ is a parabolic cylinder with vertex (x_0,t_0), where (x_0,t_0) is a Lebesgue point for Du. Here

$$d := \begin{cases} \frac{p}{2} & \text{if} \quad p \geq 2 \\ \frac{2p}{p(n+2)-2n} & \text{if} \quad \frac{2n}{n+2} < p < 2 \end{cases}$$

is the scaling deficit exponent of the p-Laplacean system.

Let us notice that the Dini continuity assumed on the map $\gamma(\cdot)$ is indeed necessary. Counterexamples (see [26]) valid already in the case of linear elliptic equations of the type $\operatorname{div}(A(x)Du) = 0$ show that when coefficients $A(\cdot)$ (i.e. its entries as a matrix) are merely continuous, but not necessary Dini continuous, the gradient might be unbounded and even does not belong to BMO. As a matter of fact, in this respect Theorem 1.1 is new already in the case (1.2) and extends to the parabolic case classical elliptic results. As for the improved Dini continuity on the right hand side datum $G(\cdot)$, this type of result appears to be new already in the elliptic case. Notice that for homogeneity reasons, the correction to the standard Dini continuity due to the presence of the exponent $\min\{1, p/[2(p-1)]\}$ in the definition of $\omega(\varrho)$ appears to be the natural one. We shall go back to Dini continuity in the next section, where we shall show that when considering everywhere (not only asymptotically) parabolic systems of the type in (1.2), the gradient is not only locally bounded, but, rather, continuous. Further optimality of Theorem 1.1 is featured by estimate (1.10). This indeed shows an optimal scaling - essentially linked to the anisotropicity of the evolutionary p-Laplacean structure - and reduces to the one of DiBenedetto [9, Chapter 8, Theorems 5.1, 5.2] and DiBenedetto & Friedman [10] for the case (1.3); this is in turn reproduced in Theorem 4.2 below (where one has to take $\lambda = 1$). Also compare estimate (1.10) with the ones in [1] that show the occurrence of the same scaling deficit exponent d precisely reflecting the anisotropicity of the operator considered. In this connection, we actually remark that the Theorem 1.1 will be derived as a consequence of a more general *intrinsic gradient bound* obtained in Theorem 4.1 below that involves an optimal extension of DiBenedetto intrinsic estimates (see Theorem 4.2 below).

1.2. **Borderline conditions for continuity.** When dealing with truly parabolic systems - as for instance in (1.2) - Dini continuity of coefficients actually implies the continuity of the (spatial) gradient. This fact, being classical and sharp in the elliptic case, was still an open issue in the parabolic one and it is hereby established both for general equations and for systems with quasi-diagonal structure as the one in (1.2). In this last respect, we have

Theorem 1.2 (Borderline gradient continuity)**.** *Let u be a solution to* (1.2) *under the assumptions* (1.7)-(1.9)*. Then Du is continuous in Ω_T.*

The previous theorem extends to general classes of quasilinear parabolic equations of the type

$$u_t - \operatorname{div} a(x,t,Du) = \operatorname{div} G(x,t)\,, \tag{1.11}$$

with the vector field $a\colon \Omega_T \times \mathbb{R}^n \to \mathbb{R}^n$ satisfying the assumptions

$$\begin{cases} |a(x,t,z)| + |\partial a(x,t,z)|(|z|^2+s^2)^{1/2} \le L(|z|^2+s^2)^{(p-1)/2} \\ \nu(|z|^2+s^2)^{(p-2)/2}|\xi|^2 \le \langle \partial a(x,t,z)\xi,\xi\rangle \\ |a(x,t,z)-a(x_0,t,z)| \le L\omega(|x-x_0|)(|z|^2+s^2)^{(p-1)/2} \end{cases} \tag{1.12}$$

whenever $z,\xi \in \mathbb{R}^n$ and $(x,t) \in \Omega_T$. Here ∂a denotes the partial derivative of $a(\cdot)$ with respect to the gradient variable z. Numbers s,ν,L are assumed to satisfy $0 < \nu \le L$ and $s \ge 0$. Here $\omega(\cdot) \le 1$ is nondecreasing functions which is assumed to satisfy (1.9) and it describes the rate of oscillations of coefficients.

Theorem 1.3. *Let u be a solution to* (1.11) *under the assumptions* (1.12) *and* (1.9)*; here $N = 1$. Then Du is continuous in Ω_T.*

Note that the previous theorem only holds for equations as it is generally false for general systems, unless, as usual, a quasi-diagonal structure is assumed. For general systems only so called partial regularity is available - i.e. continuity of the gradient outside a negligible closed set - and for the parabolic case we refer for instance to the recent paper of Baroni [2].

1.3. **Nonlinear Schauder estimates.** A major gap in the regularity theory of quasilinear parabolic equations as (1.11) is the lack of the so-called *nonlinear Schauder estimates*. This, in turn, amounts to the following: when considering an equation as (1.11) with Hölder continuous "data", spatial gradients of solutions are Hölder continuous. More precisely, let us assume that the vector field $G\colon \Omega_T \to \mathbb{R}^n$ is Hölder continuous w.r.t. to the variable x and that so is also the partial map

$$x \mapsto \frac{a(x,\cdot)}{(|z|^2+1)^{(p-1)/2}}\,,$$

that is

$$\omega(\varrho) \leq c\varrho^h \tag{1.13}$$

holds for some $h > 0$. Then, in analogy to the elliptic case, one expects that Du is locally Hölder continuous in Ω_T. While Misawa [29, 30, 31] has shown this fact for the model case (1.2) (and also when solutions are considered to be vector valued), the result for the general equations as (1.11) was still missing, as a consequence of the lack of a priori regularity estimates for general equations of the type (1.11). Such a result has been recently obtained in [20, 21] in the context of pointwise estimates via nonlinear potentials (see also the announcement in [22] and [19] for nonlinear potentials). There a new approach to the Hölder continuity of the spatial gradient of solutions to equations as $u_t - \operatorname{div} a(Du) = 0$ is proved. Starting from the arguments in [20, 21], we are then able to establish the expected regularity results:

Theorem 1.4 (Nonlinear Schauder estimates)**.** *Let u be a solution to* (1.11) *under the assumptions* (1.12) *and* (1.13). *Then there exists an exponent $h_0 \in (0,1)$, depending only on n, p, ν, L, h such that $Du \in C^{0,h_0}_{\rm loc}(\Omega_T, \mathbb{R}^n)$.*

The proof of Theorem 1.4 applies to solutions to the p-Laplacean system in (1.1) as well, and in this case we recover the result of Misawa in [29].

2. Main notation and definitions

In what follows we denote by c a general positive constant, possibly varying from line to line; special occurrences will be denoted by c_1, c_2 etc; relevant dependencies on parameters will be emphasized using parentheses. *All such constants, with exception of the constant in this paper denoted by c_0, will be larger or equal than one.* We also denote by

$$B(x_0, r) := \{x \in \mathbb{R}^n \,:\, |x - x_0| < r\}$$

the open ball with center x_0 and radius $r > 0$; when not important, or clear from the context, we shall omit denoting the center as follows: $B_r \equiv B(x_0, r)$. Unless otherwise stated, different balls in the same context will have the same center. We shall also denote $B \equiv B_1 = B(0,1)$ if not differently specified. In a similar fashion we shall denote by $Q_r(x_0,t_0) := B(x_0,r) \times (t_0 - r^2, t_0)$ the standard parabolic cylinder with vertex (x_0, t_0) and width $r > 0$. When the vertex will not be important in the context or it will be clear that all the cylinders occurring in a proof will share the same vertex, we shall omit to indicate it, simply denoting Q_r. With $\lambda > 0$ being a free parameter, we shall often consider cylinders of the type

$$Q_r^\lambda(x_0,t_0) := B(x_0,r) \times (t_0 - \lambda^{2-p} r^2, t_0)\,. \tag{2.1}$$

These will be called "intrinsic cylinders" as they will be usually employed in a context when the parameter λ is linked to the behavior of the solution of some equation on the same cylinder Q_r^λ according to the standard intrinsic geometry techniques (see for instance [9, 1, 14, 15, 16]). Again, when specifying the vertex will not be essential we shall simply denote $Q_r^\lambda \equiv Q_r^\lambda(x_0,t_0)$. Observe that the intrinsic cylinders reduce to the standard parabolic ones when either $p=2$ or $\lambda=1$. In the rest of the paper λ *will always denote a constant larger than zero* and will be considered in connection to intrinsic cylinders as (2.1). We shall often denote

$$\delta Q_r^\lambda(x_0,t_0) \equiv Q_{\delta r}^\lambda(x_0,t_0) = B(x_0,\delta r)\times(t_0-\lambda^{2-p}\delta^2 r^2, t_0)$$

the intrinsic cylinder with width magnified of a factor $\delta>0$. Finally, with $Q = A\times(t_1,t_2)$ being a cylindrical domain, we denote by

$$\partial_{\mathrm{par}} Q := (A\times\{t_1\})\cup(\partial A\times[t_1,t_2])$$

the usual parabolic boundary of Q, and this is nothing else but the standard topological boundary without the upper cap $A\times\{t_2\}$.

With $A\subset\mathbb{R}^{n+1}$ being a measurable subset with positive measure, and with $g\colon A\to\mathbb{R}^n$ being a measurable map, we shall denote by

$$\fint_A g(x)\,dx\,dt := \frac{1}{|A|}\int_A g(x)\,dx\,dt$$

its integral average; here $|A|$ denotes the Lebesgue measure of A. A similar notation is adopted if the integral is only in space or time. The oscillation of g on A is instead defined as

$$\operatorname*{osc}_A g := \sup_{(x,t),(x_0,t_0)\in A} |g(x,t)-g(x_0,t_0)|\,.$$

Remark 2.1. When dealing with parabolic equations, a standard difficulty in using test functions arguments involving the solution is that we start with solutions that, enjoying the regularity in (1.4), do not have in general time derivatives in any reasonable sense. In the following, we shall argue on a formal level, that is, arguing as the solutions is differentiable with respect to time. The argument can be made rigorous in a standard way via Steklov averages as for instance in [9].

2.1. **The map $V_s(z)$, and the monotonicity of $a(x,z)$.** With $s\geq 0$, we define

$$V_s(z) := (s^2+|z|^2)^{(p-2)/4}z\,,\qquad V(z)\equiv V_0(z) = |z|^{(p-2)/2}z \tag{2.2}$$

whenever $z\in\mathbb{R}^n$, which is easily seen to be a locally bi-Lipschitz bijection of $\mathbb{R}^n$. A basic property of V_s, whose proof can be found in [13, Lemma 2.1], is the following: For any $z_1,z_2\in\mathbb{R}^n$, and any $s\geq 0$, it holds

$$c^{-1}\Big(s^2+|z_1|^2+|z_2|^2\Big)^{\frac{p-2}{2}} \leq \frac{|V(z_2)-V(z_1)|^2}{|z_2-z_1|^2} \leq c\Big(s^2+|z_1|^2+|z_2|^2\Big)^{\frac{p-2}{2}}\,, \tag{2.3}$$

where $c\equiv c(n,N,p)$ is independent of s. The strict monotonicity properties of the vector field $a(\cdot)$ implied by the left hand side in $(1.12)_1$ can be recast using the map V_s. Indeed combining $(1.12)_2$ and (2.3) yields, for $c\equiv c(n,N,\nu)>0$, and whenever $z_1,z_2\in\mathbb{R}^n$

$$c^{-1}|V_0(z_2)-V_0(z_1)|^2 \leq \langle b(z_2)-b(z_1), z_2-z_1\rangle\,. \tag{2.4}$$

We recall that the vector field $b(\cdot)$ has been defined in (1.6) as $b(z)=|z|^{p-2}z$. We also notice the following inequalities:

$$\begin{cases} |z|^p \leq |V_s(z)|^2 \leq 2(s+|z|)^{p-2}|z|^2 & \text{if} \quad p\geq 2\\ |z|^p \leq s^p + 2^{(2-p)/2}|V_s(z)|^2 & \text{if} \quad p\in[1,2)\\ |V_s(z)|^2 \leq |z|^p & \text{if} \quad p\in[1,2]\,. \end{cases} \tag{2.5}$$

Remark 2.2. Given a vector valued, weakly differentiable map w, beside the usual Hilbert norm given by

$$|Dw|^2 := \sum_{\alpha,i} |D_i w^\alpha|^2,$$

when dealing with the scalar case of equations in (1.11), we shall also consider the equivalent one defined by

$$\|Dw\| := \max_{\alpha,i} |D_i w^\alpha| \,. \tag{2.6}$$

3. $C^{0,\alpha}$ SPATIAL GRADIENT ESTIMATES

This section is dedicated to extend to the vectorial case of the p-Laplacean system

$$w_t - \operatorname{div}(|Dw|^{p-2}Dw) = 0 \tag{3.1}$$

a decay excess result proved in [20, 21, 22] for equations of the type

$$w_t - \operatorname{div}\tilde{a}(t, Dw) = 0 \tag{3.2}$$

where the vector field satisfies (1.12) (when suitably recast with no x-dependence). An additional novelty with respect to [20, 21, 22], is that the results will be formulated in terms of the new excess functional

$$E_s(G, Q_\varrho^\lambda) := \left(\fint_{Q_\varrho^\lambda} |V_s(G) - (V_s(G))_{Q_\varrho^\lambda}|^2 \, dx \, dt \right)^{1/2} \tag{3.3}$$

whenever $G \in L^p(Q_\varrho^\lambda, \mathbb{R}^{Nn})$, and this will require additional delicate estimates. We shall in the following very often use the following property of integral averages

$$E_s(G, Q_\varrho^\lambda) \leq \left(\fint_{Q_\varrho^\lambda} |V_s(G) - \Gamma|^2 \, dx \, dt \right)^{1/2}, \qquad \forall\, \Gamma \in R^{Nn} \,. \tag{3.4}$$

The following theorems shall be proven:

Theorem 3.1. *Suppose that w is a weak solution to* (3.1) *in a cylinder Q_r^λ and consider numbers $A, B \geq 1$ and $\gamma \in (0,1)$. Then there exists a constant $\delta_\gamma \in (0, 1/2)$ depending only on n, p, A, B, γ, such that if*

$$\frac{\lambda}{B} \leq \sup_{Q_{\delta_\gamma r}^\lambda} |Dw| \leq \sup_{Q_r^\lambda} |Dw| \leq A\lambda \,, \tag{3.5}$$

then

$$E_s(Dw, \delta_\gamma Q_r^\lambda) \leq \gamma E_s(Dw, Q_r^\lambda) \tag{3.6}$$

holds for every number $s \geq 0$. Moreover, there exist constants $\alpha_0 \in (0,1)$ and $c(A) \geq 1$, depending only on n, N, p, A, but not on B, such that

$$\delta_\gamma = \frac{1}{c(A)} \left(\frac{\gamma}{B} \right)^{1/\alpha_0} \,. \tag{3.7}$$

In the case of the general parabolic equations (1.11) we instead have

Theorem 3.2. *Suppose that w is a weak solution to* (3.2) *in a cylinder Q_r^λ, under the assumptions* (1.12)*, and consider numbers $A, B \geq 1$ and $\gamma \in (0,1)$. Then there exists a constant $\delta_\gamma \in (0, 1/2)$ depending only on $n, p, \nu, L, A, B, \gamma$, such that, with s fixed in* (1.12)*, if*

$$\frac{\lambda}{B} \leq s + \sup_{Q_{\delta_\gamma r}^\lambda} \|Dw\| \leq s + \sup_{Q_r^\lambda} \|Dw\| \leq A\lambda \,,$$

then (3.6) *holds for the same s. Moreover, there exist constants $\alpha \in (0,1)$ and $c(A) \geq 1$, depending only on n, N, p, ν, L, A, such that also* (3.7) *holds.*

We shall start with the proof of Theorem 3.1 and then, also taking into account the results from [20, 21], we shall describe the necessary modifications to get Theorem 3.2.

3.1. **The vectorial case and Theorem 3.1.** We start with a preliminary result that encodes the fundamental regularity results obtained by DiBenedetto for the system in (3.1) in [9]; we refer to [20, Theorem 3.2] and [21, Theorem 3.2] for the scalar case and for more details on the specific formulations used here.

Theorem 3.3. *Suppose that w is, in a given cylinder Q_r^λ, either a weak solution to* (3.2) *under the assumptions* (1.12)*, or a solution to* (3.1) *(in this case w is vector valued). Then Dw is locally Hölder continuous in Q_r^λ. Moreover if*

$$s + \sup_{Q_r^\lambda} |Dw| \leq A\lambda \tag{3.8}$$

holds for a certain constant $A \geq 1$ with s fixed in (1.12) *when considering* (3.2) *(and with $s = 0$ when considering the system* (3.1)*), then*

$$|Dw(x,t) - Dw(x_1,t_1)| \leq \tilde{c}_h \lambda \left(\frac{\varrho}{r}\right)^\alpha \tag{3.9}$$

holds whenever $(x,t), (x_1,t_1) \in Q_\varrho^\lambda$ for constants $\tilde{c}_h \equiv \tilde{c}_h(n,N,p,\nu,L,A) \geq 1$ and $\alpha \equiv \alpha(n,N,p,\nu,L,A) \in (0,1)$ which are independent of s. Here $Q_\varrho^\lambda \subset Q_r^\lambda$ are intrinsic cylinders sharing the same vertex.

Let us immediately record a

Corollary 3.1. *Under the assumption of Theorem 3.3 we have that*

$$|V_s(Dw(x,t)) - V_s(Dw(x_1,t_1))| \leq c_h \lambda^{p/2} \left(\frac{\varrho}{r}\right)^{p\alpha/2} \tag{3.10}$$

holds whenever $(x,t), (x_1,t_1) \in Q_\varrho^\lambda$, for constants $c_h \equiv c_h(n,N,p,\nu,L,A)$ and α appearing in Theorem 3.3. Here $Q_\varrho^\lambda \subset Q_r^\lambda$ are intrinsic cylinders sharing the same vertex.

Proof. We use Theorem 3.3; in the case $p \geq 2$, by (2.3), (3.8) and (3.9), it follows:

$$\begin{aligned} &|V_s(Dw(x,t)) - V_s(Dw(x_1,t_1))| \\ &\qquad \leq c(s + |Dw(x,t)| + |Dw(x_1,t_1)|)^{(p-2)/2} |Dw(x,t) - Dw(x_1,t_1)| \\ &\qquad \leq c(s+\lambda)^{(p-2)/2} \lambda (\varrho/r)^\alpha \leq c\lambda^{p/2} (\varrho/r)^\alpha \,. \end{aligned} \tag{3.11}$$

In the case $2n/(n+2) < p \leq 2$ we distinguish two cases. The first is when one of the following three inequalities holds: $|Dw(x,t)| \geq |Dw(x,t) - Dw(x_1,t_1)|$, $|Dw(x_1,t_1)| \geq |Dw(x,t) - Dw(x_1,t_1)|$, $s \geq |Dw(x,t) - Dw(x_1,t_1)|$. Say, for instance, that it is the first one, the case of one of the others being similar. In this case then, using the first of the inequalities in (3.11), we come up with

$$\begin{aligned} |V_s(Dw(x,t)) - V_s(Dw(x_1,t_1))| &\leq c|Dw(x,t)|^{(p-2)/2} |Dw(x,t) - Dw(x_1,t_1)| \\ &\leq c|Dw(x,t) - Dw(x_1,t_1)|^{p/2} \end{aligned}$$

so that the statement follows directly from (3.9). Finally, when all the three inequalities fail we estimate, again starting from the second inequality in (3.11)

$$\begin{aligned} |V_s(Dw(x,t)) - V_s(Dw(x_1,t_1))| &\leq c(s + |Dw(x,t)| + |Dw(x_1,t_1)|)^{p/2} \\ &\leq c|Dw(x,t) - Dw(x_1,t_1)|^{p/2} \end{aligned}$$

and again the assertion follows using (3.9). □

Proposition 3.1. *Suppose that w solves (3.1) in a cylinder Q_r^λ, and that*

$$\sup_{Q_r^\lambda} |Dw| \leq A\lambda \tag{3.12}$$

holds; there exist constants $\sigma \in (0,1)$ and $H \geq 1$, both depending only on n, N, p, A, such that if

$$|Q_r^\lambda \cap \{|Dw| < \lambda/2\}| \leq \sigma |Q_r^\lambda| \tag{3.13}$$

holds, then $|Dw| \geq \lambda/4$ a.e. in $Q_{r/H}^\lambda$.

Proof. We shall use Theorem 3.3; therefore let us first determine a number $H \equiv H(n,N,p,A) > 1$ is such a way that $\tilde{c}_h H^{-\alpha} = 1/4$ and in turn we take $\sigma := (2H)^{-(n+2)}$, with $\tilde{c}_h$ as in (3.10). With such a choice it follows that

$$\{(x,t) \in Q_r^\lambda \,:\, |Dw(x,t)| \geq \lambda/2\} \cap Q_{r/H}^\lambda \neq \emptyset$$

and therefore there exists $(x_0,t_0) \in Q_{r/H}^\lambda$ such that

$$|Dw(x_0,t_0)| \geq \lambda/2.$$

Therefore, if $(x,t) \in Q_{r/H}^\lambda$, then

$$\begin{aligned} |Dw(x,t)| &\geq |Dw(x_0,t_0)| - |Dw(x,t) - Dw(x_0,t_0)| \\ &\geq \lambda/2 - \tilde{c}_h \lambda H^{-\alpha} = \lambda/2 - \lambda/4 = \lambda/4\,. \end{aligned}$$

□

Proposition 3.2. *Suppose that w solves (3.1) in a cylinder Q_r^λ such that $0 < \lambda/4 \leq |Dw(x,t)| \leq A\lambda$ for every $(x,t) \in Q_r^\lambda$, where $A \geq 1$. Then there exists an exponent $\beta \in (0,1)$, depending only on the parameters n, N, p, A, such that*

$$E_s(Dw, Q_{\delta r}^\lambda) \leq c\delta^\beta E_s(Dw, Q_r^\lambda) \tag{3.14}$$

holds whenever $\delta \in (0,1)$ and $s \geq 0$, for a constant $c \equiv c(n,N,p,A) \geq 1$, which is in turn independent of the number s.

Proof. Without loss of generality we shall assume that the vertex of the cylinder coincides with the origin. We now make the standard intrinsic scaling by defining $v(x,t) := r^{-1} w(rx, \lambda^{2-p} r^2 t)$ whenever $(x,t) \in Q_1$ so that the newly defined function v solves

$$\lambda^{p-2} v_t - \operatorname{div}(|Dv|^{p-2} Dv) = 0\,. \tag{3.15}$$

This change of variables allows to prove the statement only for v; the corresponding will then follow by scaling back to w. With the new definition we still have

$$0 < \lambda/4 \leq |Dv(x,t)| \leq A\lambda \qquad \forall\ (x,t) \in Q_1\,. \tag{3.16}$$

Now, first observe that (3.16) implies

$$Dv \in L^2_{\mathrm{loc}}(-1,0; W^{1,2}_{\mathrm{loc}}(B_1, \mathbb{R}^{Nn})) \cap C^0(-1,0; L^2_{\mathrm{loc}}(B_1, \mathbb{R}^{Nn}))\,. \tag{3.17}$$

Indeed, for degenerate elliptic and parabolic systems as the one we are considering here, the existence of second spatial derivatives fails in general, as $|Dv|$ might vanish at some points. On the other hand the lower inequality in (3.16) rules out this possibility and in this case the differentiability in (3.17) follows. Therefore we differentiate (3.15) with respect to x_i, thereby obtaining

$$\lambda^{p-2} (v_{x_i})_t - \operatorname{div}\,(\partial b(Dv(x,t)) Dv_{x_i}) = 0\,.$$

In turn, dividing the latest system by λ^{p-2} we see that each component v_{x_i} solves the system

$$(v_{x_i})_t - \operatorname{div}\,(B(x,t) Dv_{x_i}) = 0 \qquad \text{where} \quad B(x,t) := \lambda^{2-p} \partial b(Dv(x,t))\,.$$

By virtue of (3.16) the matrix $B(x,t)$ is uniformly elliptic in the sense that the inequalities

$$c^{-1}|\xi|^2 \leq \langle B(x,t)\xi,\xi\rangle \leq c|\xi|^2 \qquad \text{for every} \quad \xi \in \mathbb{R}^{Nn}$$

where $c \equiv c(n,N,p,A) \geq 1$. Moreover, we observe that the matrix $B(x,t)$ has Hölder continuous entries (see Lemma 3.1 below), and ultimately has a modulus of continuity which depends only on n,N,p,A. We can therefore invoke the standard Campanato's perturbation theory for linear parabolic systems with continuous coefficients (see for instance [6]) yielding the following decay estimate:

$$\fint_{Q_\delta} |v_{x_i} - (v_{x_i})_{Q_\delta}|^2\,dx\,dt \leq c\delta^{2\beta} \fint_{Q_1} |v_{x_i} - (v_{x_i})_{Q_1}|^2\,dx\,dt$$

that holds whenever $\delta \in (0,1)$, for a constant c which depends only on n,N,p,A. Ultimately, since $i \in \{1,\ldots,n\}$ is arbitrary, we arrive at

$$\fint_{Q_\delta} |Dv - (Dv)_{Q_\delta}|^2\,dx\,dt \leq c\delta^{2\beta} \fint_{Q_1} |Dv - (Dv)_{Q_1}|^2\,dx\,dt \tag{3.18}$$

and again this holds whenever $\delta \in (0,1)$, for a (new) constant c which depends only on n,N,p,A. We are now ready for the proof of estimate (3.14). Let us fix $\xi \in \mathbb{R}^{Nn}$ such that $V_s(\xi) = (V_s(Dv))_{Q_1}$ - this is possible as $V_s(\cdot)$ is bijective. We then have, using (2.3), (3.4), (3.16) and (3.18), that

$$\begin{aligned}
&\fint_{Q_\delta} |V_s(Dv) - (V_s(Dv))_{Q_\delta}|^2\,dx\,dt \\
&\quad \leq \fint_{Q_\delta} |V_s(Dv) - V_s((Dv)_{Q_\delta})|^2\,dx\,dt \\
&\quad \leq c\fint_{Q_\delta} (s^2 + |Dv|^2 + |(Dv)_{Q_\delta}|^2)^{(p-2)/2}|Dv - (Dv)_{Q_\delta}|^2\,dx\,dt \\
&\quad \leq c(s+\lambda)^{p-2}\fint_{Q_\delta} |Dv - (Dv)_{Q_\delta}|^2\,dx\,dt \\
&\quad \leq c\delta^{2\beta}(s+\lambda)^{p-2}\fint_{Q_1} |Dv - (Dv)_{Q_1}|^2\,dx\,dt \\
&\quad \leq c\delta^{2\beta}(s+\lambda)^{p-2}\fint_{Q_1} |Dv - \xi|^2\,dx\,dt \\
&\quad \leq c\delta^{2\beta}\fint_{Q_1} (s^2 + |Dv|^2 + |\xi|^2)^{(p-2)/2}|Dv - \xi|^2\,dx\,dt \\
&\quad \leq c\delta^{2\beta}\fint_{Q_1} |V_s(Dv) - V_s(\xi)|^2\,dx\,dt \\
&\quad = c\delta^{2\beta}\fint_{Q_1} |V_s(Dv) - (V_s(Dv))_{Q_1}|^2\,dx\,dt\,.
\end{aligned} \tag{3.19}$$

Observe that we have used, when $p \leq 2$, the inequality $|\xi| \leq A\lambda$, that we prove as follows. Set $z_1 = (V_s(Dv))_{Q_1} = V_s(\xi)$, and we have to prove that $V_s^{-1}(z_1) \in B_{A\lambda}$; it is sufficient to show that $z_1 \in V_s(B_{A\lambda})$. In turn, this is implied by

$$|z_1| \leq (s^2 + A^2\lambda^2)^{(p-2)/4}|A\lambda|$$

that holds as

$$|z_1| \leq \fint_{Q_1} |V_s(Dv)|\,dx\,dt \leq (s^2 + A^2\lambda^2)^{(p-2)/4}|A\lambda|$$

as the map $t \mapsto (s^2+t^2)^{(p-2)/4}t$ is increasing on the positive part of the real line. Now, scaling back the inequality in (3.19) from v and w finally yields (3.14). $\square$

Lemma 3.1. *In the framework of Proposition 3.2, it holds that*

$$|B(x,t)-B(x_0,t_0)| \leq c\left(|x-x_0|+\sqrt{|t-t_0|}\right)^{\beta_0}, \qquad \forall\ (x,t),(x_0,t_0)\in Q_{1/2}$$

where $c \geq 1$, $\beta_0 \in (0,1)$ depend on n, N, p, A.

Proof. Indeed, by scaling (3.9) to v we have that

$$|Dv(x,t)-Dv(x_0,t_0)| \leq c\lambda\left(|x-x_0|+\sqrt{|t-t_0|}\right)^{\alpha} \tag{3.20}$$

holds whenever $(x,t),(x_0,t_0)\in Q_{1/2}$ with $c \equiv c(n,N,p,A)$. Observe now that

$$B(x,t)=\frac{|Dv|^{p-2}}{\lambda^{p-2}}\left[I+(p-2)\frac{Dv\otimes Dv}{|Dv|^2}\right].$$

Therefore the statement follows by mean value theorem, with (3.20) and (3.16). □

Propositions 3.1-3.2 combined give in turn

Proposition 3.3. *Suppose that w solves* (3.1) *in a cylinder Q_r^λ, where* (3.12) *is satisfied. There exists a positive number $\sigma \equiv \sigma(n,N,p,A) \in (0,1/2)$ such that if* (3.13) *holds, then it holds that*

$$E_s(Dw,Q^\lambda_{\delta r}) \leq c_d\delta^\beta E_s(Dw,Q_r^\lambda) \qquad \forall\ \delta\in(0,1)\,, \tag{3.21}$$

for constants $\beta \in (0,1)$ and $c_d \geq 1$ depending only on n, N, p, A.

Proof. Proposition 3.1 applies here, thereby yielding $\lambda/4 \leq |Dw(x,t)|$ in $Q^\lambda_{r/H}$; this in turn allows to apply Proposition 3.2 (in the cylinder $Q^\lambda_{r/H}$). As an outcome we get that

$$E_s(Dw,Q^\lambda_{\delta r/H}) \leq c\delta^\beta E_s(Dw,Q^\lambda_{r/H}) \qquad \text{holds whenever } \delta\in(0,1)\,.$$

To estimate the right hand side of the last inequality we note that

$$\begin{aligned} E_s(Dw,Q^\lambda_{r/H}) &\leq \left(\fint_{Q^\lambda_{r/H}} |V_s(Dw)-(V_s(Dw))_{Q_r^\lambda}|^2\,dx\,dt\right)^{1/2} \\ &\leq H^{(n+2)/2}\left(\fint_{Q^\lambda_{r}} |V_s(Dw)-(V_s(Dw))_{Q_r^\lambda}|^2\,dx\,dt\right)^{1/2}. \end{aligned}$$

This means that now we have that (3.21) holds for $\delta \in (0,1/H)$; the case $\delta \in [1/H,1)$ follows enlarging again the constant of a factor $H^{(n+2)/2}$. □

The next result analyzes the case ruled out by the previous Proposition 3.3. For this we refer to [9, Proposition 1.2] and [20, Proposition 3.4].

Proposition 3.4. *Assume that* (3.12) *holds, while* (3.13) *does not hold. Then there exist $\sigma_1 \in (0,1)$ and $\eta \in (1/2,1)$, depending only on n, N, p, A, such that*

$$\sup_{Q^\lambda_{\sigma_1 r}} |Dw| \leq \eta A\lambda\,. \tag{3.22}$$

Proof of Theorem 3.1. The proof of Theorem 3.1 goes now in several steps and it is based on the one for an analogous result given in [20, 21]; since there are several points to modify, we shall report here the full argument for the sake of the reader. In turn, for brevity we shall confine ourselves to the case $p \geq 2$; the case $2n/(n+2) < p < 2$ can be obtained combining the modifications introduced here with those in the proofs in [20], and finally with the proof in [21].

Step 1: Iteration. Given a cylinder Q_r^λ such that (3.12) holds, by Propositions 3.3 and 3.4 one of the following occurs:

- **The Nondegenerate Alternative**. This means that we can apply Proposition 3.3 and therefore (3.21) holds for every $\delta \in (0,1)$, where the constants $\beta \equiv \beta(n,N,p,A) \in (0,1)$ and $c_d \equiv c_d(n,N,p,A) \geq 1$ are those defined in Proposition 3.2.
- **The Degenerate Alternative**. In this case we can instead apply Proposition 3.4 that in turn yields (3.22), where $\eta \equiv \eta(n,N,p,A) \in (0,1)$ and $\sigma_1 \equiv \sigma_1(n,N,p,A) \in (0,1)$.

The rest of the proof is based on a combination of the previous alternatives. By starting with a condition as (3.12) in an intrinsic cylinder Q_r^λ, we consider the number $\eta \equiv \eta(n,N,p,A) \in (0,1)$ defined in Proposition 3.4 and then define the sequences

$$
\begin{cases} \lambda_{j+1} := \eta\lambda_j \\ \lambda_0 := \lambda\,, \end{cases} \qquad \begin{cases} R_{j+1} := d_0 R_j \\ R_0 := r\,, \end{cases} \qquad d_0 := \frac{\sigma_1 \eta^{(p-2)/2}}{2} \in (0,1/2)\,, \tag{3.23}
$$

so that $d_0 \equiv d_0(n,N,p,A)$. With such a choice, and since we are here considering the case $p \geq 2$, the following inclusions hold:

$$
Q^{\lambda}_{R_{j+1}} \subset Q^{\lambda_{j+1}}_{R_{j+1}} \subset Q^{\lambda_j}_{\sigma_1 R_j} \subset Q^{\lambda_j}_{R_j} \subset Q^{\lambda}_{r}\,, \qquad \forall\, j \in \mathbb{N}\,. \tag{3.24}
$$

Here, as in the following, all the cylinders share the same vertex. From now on we shall also denote $Q_i := Q^{\lambda_i}_{R_i}$. We now proceed building the iteration scheme by induction: to this aim, let us assume that the Degenerate Alternative holds in the cylinders $Q^{\lambda_i}_{R_i}$ for $i \in \{1,\ldots,j\}$ for some integer j. Therefore we have that

$$
\sup_{Q^{\lambda_j}_{R_j}} |Dw| \leq A\lambda_j \qquad \text{and} \qquad \sup_{Q^{\lambda_j}_{\sigma_1 R_j}} |Dw| \leq \eta A \lambda_j = A\lambda_{j+1}
$$

hold. It follows from the last inequality and (3.24) that the intrinsic condition (3.12) is still satisfied on $Q^{\lambda_{j+1}}_{R_{j+1}}$. We can therefore check again whether or not the Degenerate Alternative holds on $Q^{\lambda_{j+1}}_{R_{j+1}}$ and so on. This procedure defines an iteration that stops in the case we reach a cylinder where the Nondegerate Alternative holds. We now have to find a suitable number δ_γ such that the statement of the theorem is true. We shall do this assuming that the lower bound in (3.5) holds for a suitably small number δ_γ that we shall determine in due course of the proof according to various restrictions, finally leading to the dependence on the various constants described in the statement. We define $m \in \mathbb{N}$ as the smallest integer such that

$$
\eta^m A\lambda < \lambda/2B\,. \tag{3.25}
$$

Observe that this determines $m \geq 1$ as a function of the parameters n,N,p,A,B and, more precisely, it satisfies

$$
m \approx \frac{\log 4AB}{-\log \eta} =: \tilde{c}_*(A)\log(AB) = \log(AB)^{\tilde{c}_*(A)}\,, \tag{3.26}
$$

for suitable constant $c_*(A)$, which is non-decreasing in A, and also depends on n,N,p. We now start taking $\delta_\gamma \leq d_0^{m+1}$, where d_0 has been introduced in (3.23), and show that, as an effect of the assumed lower bound in (3.5), the iteration always stops after a controllable number of steps. Indeed, by (3.25) we notice that

$$
A\lambda_m \equiv \eta^m A\lambda < \lambda/(2B) \leq \sup_{Q^{\lambda}_{\delta_\gamma r}} |Dw| \leq \sup_{Q^{\lambda}_{m+1}} |Dw|\,. \tag{3.27}
$$

Then, let us now define

$$
\tilde{m} := \min\left\{k \in \mathbb{N} \,:\, \text{The Degenerate Alternative does not occur on } Q^{\lambda_k}_{R_k}\right\}.
$$

Observe that by definition this means that the Degenerate Iteration can be performed $\tilde m$ times, but that the Degenerate Alternative doesn't hold on the cylinder $Q^{\lambda_{\tilde m}}_{R_{\tilde m}}$. We have now

$$\tilde m \leq m\,. \tag{3.28}$$

Indeed, were $\tilde m < m$ not the case we observe that $\tilde m = m$, as in fact we would otherwise have

$$\sup_{Q^{\lambda}_{m+1}} |Dw| \leq \sup_{Q^{\lambda m+1}_{R_{m+1}}} |Dw| \leq \eta^{m+1} A\lambda\,,$$

contradicting (3.27). Thus (3.28) holds. In the next step we shall find further smallness conditions on δ_γ.

Step 2: Estimates at the exit time. By the definition of $\tilde m$, the Nondegenerate Alternative holds in the cylinder $Q^{\lambda_{\tilde m}}_{R_{\tilde m}}$; observe that here it may happen that $\tilde m = 0$. We can therefore use (3.21) in such a cylinder, that is, we apply (3.21) with the choice $Q^\lambda_r \equiv Q^{\lambda_{\tilde m}}_{R_{\tilde m}}$. Let us define

$$\tilde\delta_\gamma := \tilde\delta d_0^m \quad \text{with} \quad \tilde\delta \in (0, d_0)\,. \tag{3.29}$$

The number $\tilde\delta$ will be chosen in a few lines, in a way that will make it depending on γ, and this justifies the notation in the line above. Recalling (3.24), we observe the following inclusions:

$$Q^\lambda_{\tilde\delta_\gamma r} = \tilde\delta d_0^{m-\tilde m} Q^\lambda_{\tilde m} \subset \tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}} \subset Q^{\lambda_{\tilde m}}_{R_{\tilde m}} \subset Q^\lambda_r \tag{3.30}$$

hold as a consequence of (3.24) and (3.29). Therefore

$$\begin{aligned}
E_s(Dw, Q^\lambda_{\tilde\delta_\gamma r}) &\leq \left(\fint_{Q^\lambda_{\tilde\delta_\gamma r}} |V_s(Dw) - (V_s(Dw))_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} \\
&\leq c \left(\frac{|\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}|}{|Q^\lambda_{\tilde\delta_\gamma r}|} \fint_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |V_s(Dw) - (V_s(Dw))_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} \\
&\leq c \left(\frac{|Q^{\lambda_{\tilde m}}_{R_{\tilde m}}|}{|Q^\lambda_{\tilde m}|} \fint_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |V_s(Dw) - (V_s(Dw))_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} .
\end{aligned}$$

On the other hand, using (3.21) with $\delta = \tilde\delta d_0^{m-\tilde m}$ and in the cylinder $Q^{\lambda_{\tilde m}}_{R_{\tilde m}}$, and keeping again (3.30) in mind, we have

$$\begin{aligned}
&\left(\fint_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |V_s(Dw) - (V_s(Dw))_{\tilde\delta d_0^{m-\tilde m} Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} \\
&\qquad \leq c(\tilde\delta d_0^{m-\tilde m})^\beta \left(\fint_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |V_s(Dw) - (V_s(Dw))_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} \\
&\qquad \leq c\tilde\delta^\beta \left(\fint_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |V_s(Dw) - (V_s(Dw))_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|^2 \, dx\, dt \right)^{1/2} ,
\end{aligned}$$

where in the last estimate we used that $d_0 \leq 1$ and (3.28); the constant c depends only on n, N, p, A. Connecting the last two groups of inequalities and continuing with the estimate, and again keeping (3.30) in mind, we have

$$\left(\fint_{Q^\lambda_{\tilde\delta_\gamma r}} |V_s(Dw) - (V_s(Dw))_{Q^\lambda_{\tilde\delta_\gamma r}}|^2 \, dx\, dt \right)^{1/2}$$

$$\leq c\tilde{\delta}^{\beta}\left(\frac{|Q^{\lambda_{\tilde m}}_{R_{\tilde m}}|}{|Q^{\lambda}_{\tilde m}|}\fint_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}}|V_s(Dw)-(V_s(Dw))_{Q^{\lambda}_r}|^2\,dx\,dt\right)^{1/2}$$

$$\leq c\tilde{\delta}^{\beta}\left(\frac{|Q^{\lambda}_{r}|}{|Q^{\lambda}_{\tilde m}|}\fint_{Q^{\lambda}_{r}}|V_s(Dw)-(V_s(Dw))_{Q^{\lambda}_r}|^2\,dx\,dt\right)^{1/2}\leq\frac{\tilde c\tilde{\delta}^{\beta}E(Dw,Q^{\lambda}_r)}{d_0^{m(n+2)/2}},$$

where $\tilde c\equiv\tilde c(n,N,p,A)$ and we have used (3.28). Now, if we impose that

$$\tilde\delta\leq\left(\frac{d_0^{m(n+2)/2}\gamma}{\tilde c(n,N,p,A)}\right)^{1/\beta} \tag{3.31}$$

then we have (3.6).

Step 3: Final choice of δ_γ *and verification of* (3.7). By using (3.29) and (3.31), we are led to define

$$\delta_\gamma:=\left(\frac{d_0^{m(n+2)/2}\gamma}{\tilde c(n,N,p,A)}\right)^{1/\beta}d_0^m$$

so that (3.6) follows as in Step 1 and 2. It remains to check the validity of (3.7); this, in turn, easily follows for suitable values of c and α_0, recalling that d_0 depends only on n,N,p,A and using (3.26). The proof is complete. □

Remark 3.1. Although we have stated results for solutions to the standard p-Laplacean system, Theorems 3.1 and 3.3 remain valid for systems with measurable time dependent coefficients of the type

$$w_t-\operatorname{div}(\tilde\gamma(t)|Dw|^{p-2}Dw)=0\,, \tag{3.32}$$

where the function $\tilde\gamma(\cdot)$ is a just a measurable function satisfying bounds as in (1.8). Indeed, at every stage, in the proof of the gradient regularity of solutions to systems as (3.1), the only point where the regularity of coefficients is needed is in the first step, that consists of differentiating the system with respect to the space variable. At this stage the regularity of coefficients with respect to the time variable is irrelevant and therefore measurable dependent coefficients can be allowed. See for instance [3, 4, 5]. Summarizing, we have that Theorems 3.1 and 3.3 hold for solutions to (3.32), while the various constants depend now also on ν, L.

3.2. **The scalar case and Theorem 3.2.** The proof in the scalar case has been obtained in [20, 21] for a different, actually simpler, notion of excess functional. As a matter of fact, Proposition 3.1 is already present in [20, 21] with a different proof, suited for the scalar case, while the only thing to change is Proposition 3.2, where the excess functional appearing in (3.14) should be considered. In turn this can be achieved by reasoning as in the proof of [20, Lemma 3.2 and Proposition 3.3] and then estimating as in (3.19). This eventually leads, as here, to the analog of Proposition 3.3, while for Proposition 3.4 we again refer to [20, 21] in the scalar case. Finally, in order to achieve Theorem 3.2, Propositions 3.3 and 3.4, in their scalar formulation, can be combined exactly as here.

4. Proof of Theorem 1.1

4.1. **An intrinsic estimate.** Theorem 1.1 actually follows from

Theorem 4.1 (Intrinsic gradient bound)**.** *Let u be a solution to* (1.1) *under the assumptions* (1.5)-(1.9)*; then $Du\in L^\infty_{\rm loc}(\Omega_T)$. Moreover, let $Q^{\lambda}_{2r}\equiv Q^{\lambda}_{2r}(x_0,t_0)\subset\Omega_T$ be an intrinsic cylinder with (x_0,t_0) being a Lebesgue point for Du. There exists a constant $c_i>1$, depending only on n,N,p,ν,L, and on the rate of convergence* *in* (1.6)*, and a positive radius* $r_0 > 0$ *depending only on* n, N, p, ν, L *and* $\tilde{\omega}(\cdot)$*, such that if* $\lambda > 0$ *satisfies*

$$c_i \left(\fint_{Q_r^\lambda} (|Du| + 1)^p \, dx \, dt \right)^{1/p} \leq \lambda \tag{4.1}$$

and if

$$r \leq r_0 \,, \tag{4.2}$$

then

$$|Du(x_0, t_0)| \leq \lambda \,. \tag{4.3}$$

By saying that λ is a generalized root of (4.1), we mean a (the smallest can be taken) positive solution of the previous equation, with the word generalized referring to the possibility that no root exists in which case we simply set $\lambda = \infty$. We shall then show, when proving Theorem 1.1, that it is always possible to find generalized roots. Theorem 1.1 extends to the non-homogeneous case the classical estimates of DiBenedetto [9, Ch. 8, Sec. 5] in turn reported in Theorem 4.2 below.

Theorem 4.2. *Suppose that* w *is, in a given cylinder* Q_ϱ^λ*, either a weak solution to* (3.2)*, under the assumptions* (1.12)*, or a solution to* (3.32) *(and in this case it is vector valued and in the rest of the statement we take* $s = 0$*). Then there exists a constant* $c \geq 1$*, depending only on* n, N, p, ν, L*, but otherwise independent of* s*, of the solution* w *considered and of the vector field* $a(\cdot)$*, such that*

$$\sup_{\frac{1}{2}Q_\varrho^\lambda} |Dw| \leq c\lambda + c\lambda^{(2-p)/2} \left(\fint_{Q_\varrho^\lambda} (|Dw| + s)^p \, dx \, dt \right)^{1/2}$$

when $p \geq 2$*. In the case* $2n/(n+2) < p < 2$ *we instead have*

$$\sup_{\frac{1}{2}Q_\varrho^\lambda} |Dw| \leq c\lambda + c\lambda^{\frac{n(p-2)}{p(n+2)-2n}} \left(\fint_{Q_\varrho^\lambda} (|Dw| + s)^p \, dx \, dt \right)^{\frac{2}{p(n+2)-2n}} .$$

Therefore if

$$s^{p/2} + \left(\fint_{Q_\varrho^\lambda} |V_s(Dw)|^2 \, dx \, dt \right)^{1/2} \leq 2\lambda^{p/2}$$

holds, then

$$s + \sup_{\frac{1}{2}Q_\varrho^\lambda} |Dw| \leq c_b \lambda$$

also holds, where $c_b \geq 1$ *is a constant depending only on* n, N, p, ν, L.

4.2. **A few lemmas.** In this section we provide a few preliminary arguments that will be useful for the proof of Theorem 4.1.

Important notational remark. From now on, for the rest of the entire Section 4, we shall simply the notation by denoting

$$V(z) = V_0(z) = |z|^{(p-2)/2} z \,.$$

Accordingly, when using the notion in (3.3) we shall simply denote $E(\cdot) \equiv E_0(\cdot)$.

The following lemma can be retrieved from [23, 32], with minor modifications, due to the assumptions (1.5) we are using here.

Lemma 4.1. *Under the assumption* (1.6)*, for every* $\varepsilon_1 > 0$ *there exists* $\Sigma \equiv \Sigma(\varepsilon_1) \geq 1/\varepsilon_1$*, depending only on* n, N, p, ε_1*, such that*

$$\left| \int_0^1 (\partial b(\sigma z + (1-\sigma) z_0) - \partial a(\sigma z + (1-\sigma) z_0)) \, d\sigma \, (z - z_0) \right|$$

$$\leq \varepsilon_1 (|z-z_0|+K) \left[|z|^2+|z_0|^2+(1-\chi_2)^2\right]^{(p-2)/2}$$

holds whenever $(x_0,t_0)\in\Omega_T$, $u_0\in\mathbb{R}^n$, $z,z_0\in\mathbb{R}^n$ *and* $K\geq 0$, *and provided either* $|z_0|\geq \Sigma(\varepsilon_1)$ *or* $K\geq\Sigma(\varepsilon_1)$ *hold. We have denoted*

$$\chi_2\equiv\chi_2(p):=\begin{cases} 0 & \text{if} \quad p\geq 2\\ 1 & \text{if} \quad p\in(1,2)\,. \end{cases} \tag{4.4}$$

Let us now consider, in a fixed parabolic cylinder $Q\equiv Q_\varrho^\lambda(x_0,t_0)\Subset\Omega_T$, the unique solution

$$\begin{aligned} w &\in C^0(t_0-\lambda^{2-p}\varrho^2,t_0;L^2(B(x_0,\varrho),\mathbb{R}^N))\\ &\cap L^p(t_0-\lambda^{2-p}\varrho^2,t_0;W^{1,p}(B(x_0,\varrho),\mathbb{R}^N)) \end{aligned} \tag{4.5}$$

to the Cauchy-Dirichlet problem

$$\begin{cases} w_t-\operatorname{div}\,(\gamma(x_0,t)b(Dw))=0 & \text{in } Q_\varrho^\lambda\\ \qquad\qquad\qquad\qquad w=u & \text{on } \partial_{\text{par}}Q_\varrho^\lambda\,. \end{cases} \tag{4.6}$$

In the following we shall consider

$$\omega_\gamma(\varrho):=\sup_{\substack{t\in(-T,0),x,y\in B_\varrho\\ B_\varrho\subset\Omega}}|\gamma(x,t)-\gamma(y,t)|$$

and

$$\omega_G(\varrho):=\sup_{\substack{t\in(-T,0),x,y\in B_\varrho\\ B_\varrho\subset\Omega}}|G(x,t)-G(y,t)|\,.$$

Observe that, recalling the definition of $\omega(\cdot)$ given before (1.9), we have

$$[\omega_\gamma(\varrho)]^2+[\omega_G(\varrho)]^{\min\{2,p/(p-1)\}}\leq c(p)[\omega(\varrho)]^2\,. \tag{4.7}$$

The central result in this Section 4.2 is the following:

Lemma 4.2. *Let* u *be as in Theorem 1.1 and* w *defined as in* (4.6)*; let* $\varepsilon_1\in(0,1)$ *with* $\Sigma(\varepsilon_1)$ *being the corresponding number provided by Lemma 4.1. Finally, fix* $z_0\in\mathbb{R}^{Nn}$ *and, accordingly, set*

$$\tilde{E}:=\left(\fint_{Q_\varrho^\lambda}|V(Du)-V(z_0)|^2\,dx\,dt\right)^{1/2}$$

and

$$K:=\begin{cases}\left(\tilde{E}^2+|z_0|^p\right)^{(2-p)/(2p)}\tilde{E} & \text{when } p\geq 2\\ \tilde{E}^{2/p} & \text{when } p<2\,.\end{cases}$$

There exists a constant $\tilde{c}_0$ *depending only on* n,N,p,ν,L *such that if either* $|z_0|\geq\Sigma(\varepsilon_1)$ *or* $K\geq\Sigma(\varepsilon_1)$*, then*

$$\begin{aligned} &\fint_{Q_\varrho^\lambda}|V(Du)-V(Dw)|^2\,dx\,dt\\ &\leq \tilde{c}_0(\varepsilon_1+\omega(\varrho))\tilde{E}^2+\tilde{c}_0[\omega(\varrho)]^2\fint_{Q_\varrho^\lambda}(|V(Du)|+1)^2\,dx\,dt\,, \end{aligned} \tag{4.8}$$

where $\omega(\cdot)$ *appears in* (1.9).

Proof. Start the proof by testing the weak form of the difference equation

$$(u-w)_t-\operatorname{div}\,(\gamma(x,t)a(Du)-\gamma(x_0,t)b(Dw))=\operatorname{div}\,(G(x,t)-G(x_0,t))$$

with $u-w$; this is possible modulo a standard use of Steklov averages. After performing elementary manipulations it follows that

$$\sup_{t_0-\lambda^{2-p}\varrho^2<t<t_0} \frac{1}{|Q_\varrho^\lambda|}\int_{B(x_0,\varrho)\times\{t\}} |u-w|^2(x,t)\,dx + \fint_{Q_\varrho^\lambda} \langle \gamma(x,t)a(Du)-\gamma(x_0,t)b(Dw), Du-Dw\rangle\,dx\,dt \leq \fint_{Q_\varrho^\lambda} |G(x,t)-G(x_0,t)||Du-Dw|\,dx\,dt\,.$$

As a consequence we also have

$$\begin{aligned}\fint_{Q_\varrho^\lambda} &\langle \gamma(x_0,t)b(Du)-\gamma(x_0,t)b(Dw), Du-Dw\rangle\,dx\,dt \\ &\leq \fint_{Q_\varrho^\lambda} \langle \gamma(x_0,t)b(Du)-\gamma(x,t)a(Du), Du-Dw\rangle\,dx\,dt \\ &\quad + \fint_{Q_\varrho^\lambda} |G(x,t)-G(x_0,t)||Du-Dw|\,dx\,dt\end{aligned}$$

and, thanks to (2.4), that

$$\begin{aligned}\fint_{Q_\varrho^\lambda} &|V(Du)-V(Dw)|^2\,dx\,dt \\ &\leq c\fint_{Q_\varrho^\lambda} \langle \gamma(x_0,t)b(Du)-\gamma(x,t)a(Du), Du-Dw\rangle\,dx\,dt \\ &\quad +c\fint_{Q_\varrho^\lambda} |G(x,t)-G(x_0,t)||Du-Dw|\,dx\,dt\,,\end{aligned}$$

with $c\equiv c(n,N,p,\nu,L)$. In turn, we rewrite the previous inequality as follows:

$$\begin{aligned}\fint_{Q_\varrho^\lambda} &|V(Du)-V(Dw)|^2\,dx\,dt \\ &\leq c\fint_{Q_\varrho^\lambda} \langle \gamma(x_0,t)b(Du)-\gamma(x_0,t)a(Du), Du-Dw\rangle\,dx\,dt \\ &\quad +c\fint_{Q_\varrho^\lambda} \langle \gamma(x_0,t)a(Du)-\gamma(x,t)a(Du), Du-Dw\rangle\,dx\,dt \\ &\quad +c\fint_{Q_\varrho^\lambda} |G(x,t)-G(x_0,t)||Du-Dw|\,dx\,dt =: I+II+III\,.\end{aligned} \tag{4.9}$$

We now proceed with suitable manipulations of the terms I, II and III, actually in reverse order. When $p\geq 2$, we have

$$|Du-Dw| \leq c|V(Du)-V(Dw)|^{2/p}\,, \tag{4.10}$$

while, when $p<2$, by Young's inequality we obtain

$$\begin{aligned}|Du-Dw| &\leq c(|Du|^2+|Dw|^2)^{(2-p)/4}|V(Du)-V(Dw)| \\ &\leq c(|Du-Dw|+|Du|)^{(2-p)/2}|V(Du)-V(Dw)| \\ &\leq \frac{1}{2}|Du-Dw| \\ &\quad +c|Du|^{(2-p)/2}|V(Du)-V(Dw)|+c|V(Du)-V(Dw)|^{2/p}\end{aligned}$$

so that

$$|Du-Dw| \leq c\chi_2|Du|^{(2-p)/2}|V(Du)-V(Dw)|+c|V(Du)-V(Dw)|^{2/p} \tag{4.11}$$

holds in any case, whenever $p > 1$. Here χ_2 is as in (4.4). Therefore we have

$$
\begin{aligned}
III &\leq c\omega_G(\varrho) \fint_{Q_\varrho^\lambda} |Du - Dw| \, dx \, dt \\
&\leq \frac{1}{8} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt \\
&\qquad + c[\omega_G(\varrho)]^{\min\{2,p/(p-1)\}} \fint_{Q_\varrho^\lambda} (|Du|^{\max\{0,2-p\}} + 1) \, dx \, dt \\
&\leq \frac{1}{8} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt + c[\omega(\varrho)]^2 \fint_{Q_\varrho^\lambda} (|V(Du)| + 1)^2 \, dx \, dt \,.
\end{aligned}
$$

In the last line we have used (4.7) and (2.5). We proceed with the estimation of II; for this we have to distinguish the case $p \geq 2$ from the one in which $p < 2$. In this last case, using (4.11) and that $p/(p-1) \geq 2$, by Young's inequality it then follows

$$
\begin{aligned}
II &\leq c\omega_\gamma(\varrho) \fint_{Q_\varrho^\lambda} (|Du| + 1)^{p-1} |Du - Dw| \, dx \, dt \\
&\leq \frac{1}{8} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt + c[\omega_\gamma(\varrho)]^2 \fint_{Q_\varrho^\lambda} (|V(Du)| + 1)^2 \, dx \, dt \,.
\end{aligned}
$$

In the case $p \geq 2$ we instead use the fact that by assumption

$$
\text{either} \qquad |z_0| \geq \Sigma(\varepsilon_1) \geq 1/\varepsilon_1 \qquad \text{or} \qquad K \geq \Sigma(\varepsilon_1) \geq 1/\varepsilon_1 \qquad \text{hold}\,, \tag{4.12}
$$

giving

$$
1 \leq \frac{|z_0| + K}{\Sigma(\varepsilon_1)} \leq \varepsilon_1 \left(|z_0| + K\right) . \tag{4.13}
$$

This and the fact $\varepsilon_1 \leq 1$ further imply that if the bound $|Du| \leq 1/2$ is in force, also $|Du| + 1 \leq 4(|Du - z_0| + \varepsilon_1 K)$ holds; therefore

$$
\begin{aligned}
II &\leq c\omega_\gamma(\varrho) \fint_{Q_\varrho^\lambda} (|Du| + 1)^{p-1} |Du - Dw| \, dx \, dt \\
&\leq c\frac{\omega_\gamma(\varrho)}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap \{|Du| \geq 1/2\}} |Du|^{p-1} |Du - Dw| \, dx \, dt \\
&\qquad + c\frac{\omega_\gamma(\varrho)}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap \{|Du| < 1/2\}} (|Du - z_0| + \varepsilon_1 K)^{p-1} |Du - Dw| \, dx \, dt \\
&=: II_1 + II_2 \,.
\end{aligned}
$$

Using (2.3) and Young's inequality we obtain

$$
\begin{aligned}
II_1 &\leq c\omega_\gamma(\varrho) \fint_{Q_\varrho^\lambda} |Du|^{p/2} (|Du|^2 + |Dw|^2)^{(p-2)/4} |Du - Dw| \, dx \, dt \\
&\leq \frac{1}{16} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt + c[\omega_\gamma(\varrho)]^2 \fint_{Q_\varrho^\lambda} (|V(Du)| + 1)^2 \, dx \, dt \,.
\end{aligned}
$$

On the other hand, (4.10) and Young's inequality give

$$
\begin{aligned}
II_2 &\leq \frac{1}{16} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt \\
&\qquad + c[\omega_\gamma(\varrho)]^{p/(p-1)} \fint_{Q_\varrho^\lambda} |V(Du) - V(z_0)|^2 \, dx \, dt + c\varepsilon_1^p K^p \\
&\leq \frac{1}{16} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt + c(\varepsilon_1 + \omega_\gamma(\varrho))\tilde{E}^2 \,.
\end{aligned}
$$

Here we have also appealed to the obvious inequality $K^p \leq \tilde{E}^2$. Combining (4.12) and the estimates for II_1 and II_2 – together with the one for III in (4.7) – we conclude with

$$\begin{aligned} II + III \quad &\leq \quad \frac{1}{4} \fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx \, dt \\ &\quad + c(\varepsilon_1 + \omega(\varrho))\tilde{E}^2 + c[\omega(\varrho)]^2 \fint_{Q_\varrho^\lambda} (|V(Du)| + 1)^2 \, dx \, dt \, . \end{aligned} \tag{4.14}$$

The constant c depends only on n, N, p, ν, L. As for I, we have

$$\begin{aligned} I \quad &= \quad c \fint_{Q_\varrho^\lambda} \gamma(x_0, t) \langle (b(Du) - b(z_0)) - (a(Du) - a(z_0)), Du - Dw \rangle \, dx \, dt \\ &\leq \quad \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap A_M} |\langle (b(Du) - b(z_0)) - (a(Du) - a(z_0)), Du - Dw \rangle| \, dx \, dt \\ &\quad + \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \setminus A_M} |\langle (b(Du) - b(z_0)) - (a(Du) - a(z_0)), Du - Dw \rangle| \, dx \, dt \\ &=: \quad I_1 + I_2 \, , \end{aligned}$$

where

$$A_M := \left\{ (x,t) \in Q_\varrho^\lambda \ : \ |Dw(x,t)|^2 > M^2 \left(|Du(x,t)|^2 + |z_0|^2 \right) \right\} \, , \qquad M > 2 \, ,$$

with M to be chosen in a few lines. We now estimate I_1, using definitions and properties of A_M and $V(\cdot)$ together with Young's inequality and (4.13), as

$$\begin{aligned} I_1 \quad &\leq \quad \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap A_M} \left(|Du|^2 + |z_0|^2 + \varepsilon_1^2 K^2 \right)^{(p-1)/2} (|Du| + |Dw|) \, dx \, dt \\ &\leq \quad \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap A_M} M^{1-p} |V(Dw)|^2 \, dx \, dt \\ &\quad + \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \cap A_M} (\varepsilon_1 K)^{p-1} \, |V(Dw)|^{2/p} \, dx \, dt \\ &\leq \quad \left(\tilde{c}_1 M^{1-p} + 1/16 \right) \fint_{Q_\varrho^\lambda} |V(Dw) - V(Du)|^2 \, dx \, dt + \tilde{c}_1 \varepsilon_1 K^p \end{aligned}$$

for a constant $\tilde{c}_1 \equiv \tilde{c}_1(n, N, p, \nu, L) \geq 2^p$. We fix

$$M = (16\tilde{c}_1)^{1/(p-1)} \tag{4.15}$$

so that

$$I_1 \leq \frac{1}{8} \fint_{Q_\varrho^\lambda} |V(Dw) - V(Du)|^2 \, dx \, dt + \tilde{c}_1 \varepsilon_1 \tilde{E}^2 \tag{4.16}$$

follows, invoking also the fact $K^p \leq \tilde{E}^2$. Observe that the choice in (4.15) fixes M as a quantity depending only on n, p, N, ν, L. We now focus on I_2 and rewrite it as

$$\begin{aligned} I_2 \quad = \quad \frac{c}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \setminus A_M} \Bigg| \langle \int_0^1 \left(\partial a(\sigma Du + (1-\sigma) z_0) - \partial b(\sigma Du + (1-\sigma) z_0) \right) d\sigma \\ (Du - z_0), Du - Dw \rangle \Bigg| \, dx \, dt \, . \end{aligned}$$

We then use Lemma 4.1, which is applicable by (4.12), and obtain

$$\begin{aligned} I_2 &\leq \frac{\varepsilon_1 \tilde{c}_2}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \setminus A_M} (|Du - z_0| + K) \left(|Du|^2 + |z_0|^2 \right)^{(p-2)/2} |Du - Dw| \, dx \, dt \\ &\quad + (1 - \chi_2) \frac{\varepsilon_1 \tilde{c}_2}{|Q_\varrho^\lambda|} \int_{Q_\varrho^\lambda \setminus A_M} \left(K^{p-1} |Du - Dw| + K^{p-2} |Du - Dw||Du - z_0| \right) dx \, dt \, , \end{aligned}$$

for a constant $\tilde{c}_2 \equiv \tilde{c}_2(n,N,p,\nu,L)$, with the last integral being non-null only when $p \geq 2$ by (4.4). Observe that to obtain the estimate of the last integral we have used (4.12) to estimate $(1-\chi_2) \leq |z_0| + (1-\chi_2)K$ (essentially only when $p > 2$). Consider now the integrands appearing in the latest display, that is

$$\mathcal{I}_2 := \Big[(|Du-z_0|+K)\left(|Du|^2+|z_0|^2\right)^{(p-2)/2}|Du-Dw|\Big]\,\chi_{Q_\varrho^\lambda\setminus A_M}$$

and

$$\mathcal{I}_K := (1-\chi_2)\left[K^{p-1}|Du-Dw| + K^{p-2}|Du-Dw||Du-z_0|\right]\chi_{Q_\varrho^\lambda\setminus A_M}\,,$$

where $\chi_{Q_\varrho^\lambda\setminus A_M}$ denotes the indicator function of the set $Q_\varrho^\lambda \setminus A_M$. Observe that (2.3) implies

$$\begin{aligned}
\mathcal{I}_2 \quad &\approx \quad |V(Du)-V(z_0)|\left(\frac{|Du|^2+|z_0|^2}{|Du|^2+|Dw|^2}\right)^{(p-2)/4}|V(Du)-V(Dw)|\chi_{Q_\varrho^\lambda\setminus A_M}\\
&\qquad +K\left(\frac{(|Du|^2+|z_0|^2)^2}{|Du|^2+|Dw|^2}\right)^{(p-2)/4}|V(Du)-V(Dw)|\chi_{Q_\varrho^\lambda\setminus A_M}\\
&=: \quad \mathcal{I}_{2,1}+\mathcal{I}_{2,2}\,,
\end{aligned}$$

i.e. there exists a constant $\tilde{c}_V \equiv \tilde{c}_V(n,N,p)$ such that

$$\mathcal{I}_2/\tilde{c}_V \leq \mathcal{I}_{2,1}+\mathcal{I}_{2,2} \leq \tilde{c}_V\mathcal{I}_2\,. \tag{4.17}$$

In order to establish the previous equivalence we can always assume that $|Du| + |Dw| \neq 0$ and $|Du| + |z_0| \neq 0$, in the cases $p \geq 2$ and $p < 2$, respectively. Indeed, otherwise it would immediately follow that $\mathcal{I}_2 = 0$ by the definition of A_M; in such a case an upper bound for $\mathcal{I}_2$ would follow immediately. To complete the estimate for I we shall now distinguish the cases $p \geq 2$ and $p < 2$ and will estimate $\mathcal{I}_2$ and $\mathcal{I}_K$ accordingly, the last one being nontrivial only in the case $p \geq 2$.

Case $p \geq 2$. We start with the estimation of $\mathcal{I}_K$, via (2.3), (4.10) and Young's inequality as follows:

$$\begin{aligned}
\mathcal{I}_K \quad &\leq \quad 2K^{p-1}|Du-Dw| + |Du-z_0|^{p-1}|Du-Dw|\\
&\leq \quad cK^{p-1}|V(Du)-V(Dw)|^{2/p}\\
&\qquad\qquad +c|V(Du)-V(z_0)|^{2(p-1)/p}|V(Du)-V(Dw)|^{2/p}\\
&\leq \quad \frac{1}{16\tilde{c}_2}|V(Du)-V(Dw)|^2 + c\tilde{E}^2 + c|V(Du)-V(z_0)|^2\,, \qquad (4.18)
\end{aligned}$$

where we have also used that $\tilde{E}^2 \geq K^p$. We now turn to the estimates for $\mathcal{I}_2$. First, if $|z_0|^2 > 4\left(|Du|^2+|Dw|^2\right)$, then Young's inequality gives

$$\begin{aligned}
\mathcal{I}_2 \quad &\leq \quad c\,(|z_0|+K)\,|z_0|^{p-1}\\
&\leq \quad c\Big(|V(Du)-V(z_0)|^{2/p}+K\Big)\,|V(Du)-V(z_0)|^{2(p-1)/p}\\
&\leq \quad cK^p + c|V(Du)-V(z_0)|^2 \leq c\tilde{E}^2 + c|V(Du)-V(z_0)|^2\,,
\end{aligned}$$

because, obviously, $|Du-Dw| \leq |z_0|$ and $|z_0| \leq c(p)|V(Du)-V(z_0)|^{2/p}$ in this case. We have also used in the last estimate, again, that $K^p \leq \tilde{E}^2$. We then analyze the case $|z_0|^2 \leq 4\left(|Du|^2+|Dw|^2\right)$ and look at (4.17); Young's inequality gives

$$\mathcal{I}_{2,1} \leq \frac{1}{32\tilde{c}_2\tilde{c}_V}|V(Du)-V(Dw)|^2 + c|V(Du)-V(z_0)|^2$$

and, similarly,

$$\mathcal{I}_{2,2} \leq \frac{1}{32\tilde{c}_2\tilde{c}_V}|V(Du)-V(Dw)|^2 + cK^2\left(|Du|^2+|z_0|^2\right)^{(p-2)/2}\,.$$

Estimating further as

$$\begin{aligned}\left(|Du|^2+|z_0|^2\right)^{(p-2)/2} &\leq c\,(|Du-z_0|^p+|z_0|^p)^{(p-2)/p}\\ &\leq c\left(|V(Du)-V(z_0)|^2+|z_0|^p\right)^{(p-2)/p}\end{aligned}$$

- in the second estimate we have again used (2.3) - gives

$$\mathcal{I}_{2,2} \leq \frac{1}{32\tilde{c}_2\tilde{c}_V}|V(Du)-V(Dw)|^2 + cK^2\left(|V(Du)-V(z_0)|^2+|z_0|^p\right)^{(p-2)/p}\,. \tag{4.19}$$

Combining estimates between (4.18) and (4.19), and recalling (4.17), we have

$$\begin{aligned}\mathcal{I}_2+\mathcal{I}_K &\leq \frac{1}{8\tilde{c}_2}|V(Du)-V(Dw)|^2 + c|V(Du)-V(z_0)|^2\\ &\quad +c\tilde{E}^2 + cK^2\left(|V(Du)-V(z_0)|^2+|z_0|^p\right)^{(p-2)/p}\,,\end{aligned} \tag{4.20}$$

with $c \equiv c(n,N,p,\nu,L)$. Averaging the last estimate, and then using Hölder's inequality and definitions of K and $\tilde{E}$, yields

$$K^2 \fint_{Q_\varrho^\lambda}\left(|V(Du)-V(z_0)|^2+|z_0|^p\right)^{(p-2)/p}\,dx\,dt \leq K^2\left(\tilde{E}^2+|z_0|^p\right)^{(p-2)/p} = \tilde{E}^2\,.$$

Using this last observation, and putting (4.16) and (4.20) together, gives

$$I \leq \frac{1}{4}\fint_{Q_\varrho^\lambda}|V(Du)-V(Dw)|^2\,dx\,dt + c\varepsilon_1\tilde{E}^2\,,$$

with $c \equiv c(n,N,p,\nu,L)$. In turn, combining this with (4.9) and (4.14) completes the proof of (4.8) in the case $p \geq 2$.

Case $p<2$. It remains to estimate $\mathcal{I}_2$ in the case $p<2$. As we have restricted our study to $Q_\varrho^\lambda \setminus A_M$, we have $|Dw|^2 \leq M^2(|Du|^2+|z_0|^2)$ with the choice of $M \equiv M(n,N,p,\nu,L)$ operated in (4.15). This in turn implies

$$\mathcal{I}_{2,1} \leq c|V(Du)-V(z_0)||V(Du)-V(Dw)|$$

with $c \equiv c(n,N,p,M) \equiv c(n,N,p,\nu,L)$ and

$$\mathcal{I}_{2,2} \leq cK\left(|Du|^2+|z_0|^2\right)^{\frac{p-2}{4}}|V(Du)-V(Dw)|\,.$$

Using again condition $|Dw|^2 \leq M^2(|Du|^2+|z_0|^2)$, we further estimate as

$$\begin{aligned}|V(Du)-V(Dw)| &\leq (|V(Du)|+|V(Dw)|)^{\frac{2-p}{p}}|V(Du)-V(Dw)|^{1-\frac{2-p}{p}}\\ &\leq c\left(|Du|^2+|Dw|^2\right)^{\frac{2-p}{4}}|V(Du)-V(Dw)|^{\frac{2(p-1)}{p}}\\ &\leq c\left(|Du|^2+|z_0|^2\right)^{\frac{2-p}{4}}|V(Du)-V(Dw)|^{\frac{2(p-1)}{p}}\,,\end{aligned}$$

so that the estimates in the last two displays give

$$\mathcal{I}_{2,2} \leq cK|V(Du)-V(Dw)|^{\frac{2(p-1)}{p}}\,,$$

again with $c \equiv c(n,N,p,\nu,L)$. Using (4.17) and Young's inequality we thus deduce

$$\mathcal{I}_2 \leq \frac{1}{8\tilde{c}_2}|V(Du)-V(Dw)|^2 + c|V(Du)-V(z_0)|^2 + cK^p\,.$$

Together with (4.16) and $K=\tilde{E}^{2/p}$ this gives, again for $c \equiv c(n,N,p,\nu,L)$, that

$$I \leq \frac{1}{4}\fint_{Q_\varrho^\lambda}|V(Du)-V(Dw)|^2\,dx\,dt + c\varepsilon_1\tilde{E}^2\,.$$

Combining (4.9), (4.14) and the last estimate gives (4.8) in the case $p<2$. □

Similarly to the previous lemma we have

Lemma 4.3. *Let u be as in Theorems 1.2 and w defined as in* (4.6). *There exists a constant c_V depending only on n, N, p, ν, L such that the following inequality holds:*

$$\fint_{Q_\varrho^\lambda} |V(Du) - V(Dw)|^2 \, dx\, dt \le c_V^2 [\omega(\varrho)]^2 \fint_{Q_\varrho^\lambda} |V(Du)|^2 \, dx\, dt\,,$$

where $\omega(\cdot)$ has been defined in (1.9).

Remark 4.1. Let us now consider the framework of Theorems 1.3-1.4; in a fixed parabolic cylinder $Q_\varrho^\lambda \Subset \Omega_T$, the unique solution w, as in (4.5), to the following Cauchy-Dirichlet problem:

$$\begin{cases} w_t - \operatorname{div} a(x_0, t, Dw) = 0 & \text{in } Q_\varrho^\lambda \\ w = u & \text{on } \partial_{\mathrm{par}} Q_\varrho^\lambda\,. \end{cases} \tag{4.21}$$

A slight but yet standard modification of the above arguments leads to see that Lemma 4.3 works exactly as in the case of (4.6), with $\omega(\cdot)$ being now defined in (1.12). More precisely, we have, with s introduced in (1.12), that

$$\fint_{Q_\varrho^\lambda} |V_s(Du) - V_s(Dw)|^2 \, dx\, dt \le c_V^2 [\omega(\varrho)]^2 \fint_{Q_\varrho^\lambda} (|Du| + s)^p \, dx\, dt\,.$$

4.3. **Proof of Theorem 4.1.** We shall use large (de)magnifying constants such as 600, 800, 1200, to clarify the role of certain passages in the proof. Now, define the set $\mathcal{L}_\lambda$ (of Lebesgue points) as

$$\mathcal{L}_\lambda = \left\{ (x_0, t_0) \in \Omega_T \ : \ \lim_{\varrho \to 0} \fint_{Q_\varrho^\lambda(x_0,t_0)} |V(Du)|^2 \, dx\, dt = |V(Du(x_0,t_0))|^2 \right\} \tag{4.22}$$

for $\lambda > 0$. Basic properties of maximal operators imply that this set is actually independent of λ and, in particular, $\mathcal{L}_\lambda = \mathcal{L}_1 =: \mathcal{L}$ for all $0 < \lambda < \infty$. Moreover, $\tilde{Q} \setminus \mathcal{L}$ has zero Lebesgue measure. Therefore, in the following, when referring to the statement of Theorem 4.1, we shall prove (4.1) whenever $(x_0, t_0) \in \mathcal{L}$.

Step 1: Setting of the quantities and exit time argument. In the following all the cylinders will have (x_0, t_0) as vertex, therefore we shall omit denoting the vertex simply writing $Q_\varrho^\lambda(x_0, t_0) \equiv Q_\varrho^\lambda$. Moreover, we recall the notation for the excess functional introduced in (3.3). We now start taking λ of the form

$$\lambda^{p/2} := H_1 \left(\fint_{Q_r^\lambda} |Du|^p \, dx\, dt \right)^{1/2} + H_2 = H_1 \left(\fint_{Q_r^\lambda} |V(Du)|^2 \, dx\, dt \right)^{1/2} + H_2\,, \tag{4.23}$$

with $r \le r_0$, and fix the constant $H_1, H_2 \ge 1$ and $r_0 > 0$ in due course of the proof in such way that they will depend only on n, N, p, ν, L and, quantitatively, on the rate of convergence in (1.6). We look at Theorems 3.1 and 4.2, and let

$$c_b =: A \equiv A(n, N, p, \nu, L)\,. \tag{4.24}$$

We then determine the constant $\delta_\gamma \equiv \delta_\gamma(n, N, p, \nu, L, A, B, \gamma) \in (0, 1/2)$ in Theorem 3.1 with such a choice of A and with

$$\gamma = 2^{-5-(n+2)/2}\,, \qquad B := 200^{2/p}\,. \tag{4.25}$$

Now define

$$Q_i := Q_{r_i}^\lambda\,, \qquad r_i = \delta_1^i r\,, \qquad \delta_1 := \delta_\gamma / 4 \tag{4.26}$$

whenever $i \ge 0$ is an integer; again $\delta_1 \equiv \delta_1(n, N, p, \nu, L) \in (0, 1/8)$. We also set

$$H_1 := 400 \delta_1^{-(n+2)/2} \tag{4.27}$$

so that

$$\left(\fint_{Q_0} |V(Du)|^2\, dx\, dt\right)^{1/2} + \delta_1^{-(n+2)/2} E(Du, Q_0) \leq \frac{\lambda^{p/2}}{100}\,. \tag{4.28}$$

Define now, whenever $i \geq 0$,

$$C_i := \left(\fint_{Q_i} |V(Du)|^2\, dx\, dt\right)^{1/2} + \delta_1^{-(n+2)/2} E(Du, Q_i)\,. \tag{4.29}$$

Now, observe that (4.28) reads also as

$$C_0 \leq \frac{\lambda^{p/2}}{100}\,.$$

Let us show that without loss of generality we may assume there exists an exit index $i_e \geq 0$ with respect to the previous inequality, that is an integer $i_e \geq 0$ such that

$$C_{i_e} \leq \frac{\lambda^{p/2}}{100}\,, \qquad C_{i_e+m} > \frac{\lambda^{p/2}}{100}\,, \qquad \forall\, m \geq 1\,. \tag{4.30}$$

Indeed, on the contrary, we could find an increasing subsequence $\{j_i\}$ such that $C_{j_i} \leq \lambda^{p/2}/100$, and then, as $(x_0, t_0) \in \mathcal{L}$, we have

$$|Du(x_0, t_0)|^{p/2} = \lim_{i\to\infty} \left(\fint_{Q_{j_i}} |V(Du)|^2\, dx\, dt\right)^{1/2} \leq \frac{\lambda^{p/2}}{100}\,,$$

and the proof would be finished. Therefore, from now on, for the rest of the proof, we shall argue under the additional assumption (4.30). Moreover, when considering the cylinders Q_i and related quantities, for the rest of the proof, *we shall always consider the case $i \geq i_e$, so that the inequalities* (4.30) *are in force.*

Next, we look at Corollary 3.1 and inequality (3.10), and with the choice of A made in (4.24) we consider the exponent $\alpha \in (0,1)$ and the constant $c_h \geq 1$ determined by A; again we observe the dependence $\alpha, c_h \equiv \alpha, c_h(n, N, p, \nu, L)$. We now take k as the smallest integer (larger or equal to 2) so that

$$c_h \delta_1^{(k-1)p\alpha/2} \leq \frac{\delta_1^{(n+2)/2}}{800}\,. \tag{4.31}$$

Then k depends only upon n, N, p, ν, L as also δ_1 and c_h do. With k and δ_1 fixed, we set

$$\varepsilon_1 := \frac{\delta_1^{(k+4)(n+2)}}{8\tilde{c}_0 800^4}\,, \tag{4.32}$$

where $\tilde{c}_0$ is the constant appearing in (4.8), and so ε_1 is still a function of n, N, p, ν, L. Then, looking at Lemma 4.1, we determine the quantity $\Sigma(\varepsilon_1) > 1/\varepsilon_1$ with the choice in (4.32), and therefore as a function of n, N, p, ν, L, and of course of the rate of convergence in (1.6). Finally, we fix $H_2 \equiv H_2(n, N, p, \nu, L)$ as follows:

$$H_2^2 := 2000^{4p} \delta_1^{-2(n+2)} [\Sigma(\varepsilon_1)]^p + 2000^2\,. \tag{4.33}$$

Once again, we have that H_2 depends on n, N, p, ν, L and on the rate of convergence in (1.6). Finally, we determine the value of r_0 so that a number of smallness conditions - determined only in dependence on the basic parameters n, N, p, ν, L - are satisfied. Specifically, we fix r_0 to be small enough to satisfy

$$\omega(r_0) + \int_0^{2r_0} \omega(\varrho)\, \frac{d\varrho}{\varrho} \leq \frac{\delta_1^{(k+4)(n+2)}}{8\tilde{c}_0 800^4}\,. \tag{4.34}$$

Notice that this makes r_0 being a constant depending only on n, N, p, ν, L and $\omega(\cdot)$. Next, for integers $i \geq i_e$, we define

$$A_i := E(Du, Q_i) \qquad \text{and} \qquad m_i = |(V(Du))_{Q_i}|\,, \tag{4.35}$$

and the numbers $\{K_i\}$ as

$$K_i := \begin{cases} \left(A_i^2 + |z_i|^p\right)^{(2-p)/(2p)} A_i & \text{when } p \geq 2 \\ A_i^{2/p} & \text{when } p < 2\,, \end{cases} \tag{4.36}$$

where $z_i \in \mathbb{R}^{Nn}$ - recall that $V(\cdot)$ is bijective - has been taken in order to satisfy

$$V(z_i) = (V(Du))_{Q_i}\,. \tag{4.37}$$

Observe that we may always assume that $A_i^2 + |z_i|^p > 0$ (otherwise all the kind of estimates we are bound to prove in the following trivialize), while the choice in (4.37) is possible as $V(\cdot)$ is a bijection of $\mathbb{R}^{Nn}$. Observe that $i \geq i_e$ and (4.30) give

$$\begin{aligned} &2|V(z_i)|^2 + \delta_1^{-(n+2)} \fint_{Q_i} |V(Du) - (V(Du))_{Q_i}|^2\,dx\,dt \\ &\quad \geq \fint_{Q_i} |V(Du)|^2\,dx\,dt + \left[\delta_1^{-(n+2)}/2\right] [E(Du, Q_i)]^2 \\ &\quad \geq \delta_1^{n+2} \left[\fint_{Q_{i+1}} |V(Du)|^2\,dx\,dt + \left[\delta_1^{-(n+2)}/2\right] [E(Du, Q_{i+1})]^2\right] \geq \frac{\delta_1^{n+2}\lambda^p}{40^4} \end{aligned}$$

and by the choice in (4.33) it also holds

$$2|z_i|^p + \left[\delta_1^{-(n+2)}/2\right] \fint_{Q_i} |V(Du) - (V(Du))_{Q_i}|^2\,dx\,dt \geq 50^p \delta_1^{-(n+2)} \left[\Sigma(\varepsilon_1)\right]^p\,.$$

Now, assume that $|z_i| < \Sigma(\varepsilon_1)$; it follows that $A_i^2 > 10^p\,[\Sigma(\varepsilon_1)]^p$ and by (4.36) also that

$$K_i^p \geq \frac{A_i^2}{\max\{1, 2^{(p-2)/2}\}} \geq [\Sigma(\varepsilon_1)]^p\,.$$

Summarizing, either $m_i^{2/p} = |z_i| \geq \Sigma(\varepsilon_1)$ or the inequality in the above display holds true. In any case we can apply Lemma 4.2 with $w_i \equiv w$, $z_i \equiv z_0$ and $K_i \equiv K$. Here w_i denotes the comparison map defined in (4.6) with $Q_\varrho^\lambda \equiv Q_i$, i.e. w_i solves

$$\begin{cases} (w_i)_t - \operatorname{div}\,(\gamma(x_0, t) b(Dw_i)) = 0 & \text{in } Q_i \\ w_i = u & \text{on } \partial_{\mathrm{par}} Q_i\,. \end{cases}$$

We obtain, after an elementary manipulation of (4.8), that, if $i \geq i_e$ then

$$\begin{aligned} &\fint_{Q_i} |V(Du) - V(Dw_i)|^2\,dx\,dt \\ &\quad \leq \tilde{c}_0(\varepsilon_1 + \omega(r_i))[E(Du, Q_i)]^2 + \tilde{c}_0[\omega(r_i)]^2 \fint_{Q_i} (|V(Du)| + 1)^2\,dx\,dt \\ &\quad \leq \frac{\delta_1^{(k+4)(n+2)}}{800^4}[E(Du, Q_i)]^2 + 2\tilde{c}_0[\omega(r_i)]^2(m_i + 1)^2\,, \end{aligned} \tag{4.38}$$

where we have used (4.32), (4.34) and (4.37). Here c depends only on n, N, p, ν, L.

Step 2: Intermediate Lemmas. In the following we present a series of Lemmas whose assumptions will be eventually verified when building up the final iteration procedure.

Lemma 4.4. *Assume that for $i \geq i_e$ it holds that*

$$1 + m_i + A_i \leq \frac{\lambda^{p/2}}{2}\,. \tag{4.39}$$

With $k \equiv k(n, N, p, \nu, L) \geq 2$ defined via (4.31)*,*

$$\left(\fint_{Q_{i+k}} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \leq \frac{\delta_1^{n+2}}{800^2} \lambda^{p/2} \tag{4.40}$$

and

$$\left(\fint_{Q_i} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \leq \frac{\delta_1^{n+2}}{800^2} \lambda^{p/2} \tag{4.41}$$

hold.

Proof. By (4.38), (4.39) and (4.34), we have

$$\begin{aligned}
&\left(\fint_{Q_{i+k}} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
&\quad \leq \left(\frac{|Q_i|}{|Q_{i+k}|} \right)^{1/2} \left(\fint_{Q_i} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
&\quad \leq \frac{\delta_1^{n+2}}{800^2} E(Du, Q_i) + \sqrt{2\tilde{c}_0} \delta_1^{-k(n+2)/2} \omega(r)(m_i + 1) \leq \frac{\delta_1^{n+2}}{800^2} \lambda^{p/2}
\end{aligned}$$

and (4.40) follows. The same argument also implies (4.41). □

Lemma 4.5. *If $i \geq i_e$ and*

$$\left(\fint_{Q_i} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \leq \lambda^{p/2} \tag{4.42}$$

holds, then

$$\sup_{Q_{i+1}} |Dw_i| \leq \sup_{\frac{1}{2} Q_i} |Dw_i| \leq c_b \lambda \equiv A\lambda\,. \tag{4.43}$$

Moreover, with $k \equiv k(n, N, p, \nu, L) \geq 2$ defined via (4.31)*, it holds that*

$$2\delta_1^{-(n+2)/2} E(Dw_i, Q_{i+k}) \leq \frac{\lambda^{p/2}}{400}\,. \tag{4.44}$$

Proof. Estimate (4.43) is a direct consequence of Theorem 4.2 and (4.42); recall that in this case we are taking $s = 0$ in Theorem 4.2. Notice also that $Q_{i+1} \subset (1/2)Q_i$ as $\delta_1 < 1/2$. At this point, as a consequence of Theorem 3.3 and Corollary 3.1 (applied with $Q_\varrho^\lambda \equiv Q_{i+k}$ and $Q_r^\lambda \equiv Q_{i+1}$, and recalling the choice in (4.24)), estimates (3.10) and (4.31) yield

$$\operatorname*{osc}_{Q_{i+k}} V(Dw_i) \leq c_h \delta_1^{(k-1)p\alpha/2} \lambda^{p/2} \leq \frac{\delta_1^{(n+2)/2}}{800} \lambda^{p/2}\,,$$

in turn implying (4.44). □

Lemma 4.6. *Assume that for $i \geq i_e$ estimate* (4.42) *holds together with* (4.39)*. Then it also holds*

$$\frac{\lambda}{200^{2/p}} \leq \sup_{\frac{\delta_\gamma}{2} Q_i} |Dw_i|\,. \tag{4.45}$$

Proof. By using (3.4), triangle inequality, (4.44) and (4.40), we have

$$
\begin{aligned}
C_{i+k} \ \leq\ & \left(\fint_{Q_{i+k}} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
& + \left(\fint_{Q_{i+k}} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
& + \delta_1^{-(n+2)/2} \left(\fint_{Q_{i+k}} |V(Du) - (V(Dw_i))_{Q_{i+k}}|^2 \, dx \, dt \right)^{1/2} \\
\leq\ & \left(\fint_{Q_{i+k}} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} + \delta_1^{-(n+2)/2} E(Dw_i, Q_{i+k}) \\
& + 2\delta_1^{-(n+2)/2} \left(\fint_{Q_{i+k}} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
\leq\ & \left(\fint_{Q_{i+k}} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} + \frac{\lambda^{p/2}}{200} \, .
\end{aligned}
$$

The previous inequality and (4.30) then give

$$
\frac{\lambda^{p/2}}{200} \leq \left(\fint_{Q_{i+k}} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \leq \sup_{Q_{i+1}} |Dw_i|^{p/2}
$$

and therefore

$$
\frac{\lambda}{200^{2/p}} \leq \sup_{Q_{i+1}} |Dw_i| \, .
$$

In turn, observe that by the definition of δ_1 in (4.26) we have

$$
Q_{i+1} = Q^{\lambda}_{\delta_1^{i+1} r} = Q^{\lambda}_{\delta_\gamma \delta_1^i r/2} = (\delta_\gamma/2) Q_i \tag{4.46}
$$

so that (4.45) follows and the lemma is proved. □

In the next lemma we exploit some decay properties of the excess functional.

Lemma 4.7. *Let $i \geq i_e$ and assume that* (4.39) *holds. Then it also holds*

$$
E(Du, Q_{i+1}) \leq \tfrac{1}{4} E(Du, Q_i) + 2\sqrt{\tilde{c}_0}\, \delta_1^{-(n+2)/2} \omega(r_i) \lambda^{p/2} \, , \tag{4.47}
$$

where the constant $\tilde{c}_0 \equiv \tilde{c}_0(n, N, p, \nu, L)$ is the one introduced in Lemma 4.2.

Proof. Let us first show that we are able to use both Lemma 4.5 and 4.6. In fact, by (4.39) we get (4.41), and therefore, again thanks (4.39), we have

$$
\begin{aligned}
& \left(\fint_{Q_i} |V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
& \quad \leq \left(\fint_{Q_i} |V(Du)|^2 \, dx \, dt \right)^{1/2} + \left(\fint_{Q_i} |V(Du) - V(Dw_i)|^2 \, dx \, dt \right)^{1/2} \\
& \quad \leq |(V(Du))_{Q_i}| + E(Du, Q_i) + \frac{\delta_1^{n+2}}{800^2} \lambda^{p/2} \leq \lambda^{p/2} \, .
\end{aligned}
$$

Since (4.42) is now satisfied, at this point we can apply both Lemma 4.5 and Lemma 4.6 to get (4.43) and (4.45), respectively; summarizing, we have

$$
\frac{\lambda}{200^{2/p}} \leq \sup_{\frac{\delta_\gamma}{2} Q_i} |Dw_i| \leq \sup_{\frac{1}{2} Q_i} |Dw_i| \leq A\lambda \, .
$$

The last inequality allows to apply Theorem 3.2 to $w_i(\equiv w)$, with the choice made in (4.25), in the cylinder $(1/2)Q_i(\equiv Q_r^\lambda$ in the notation of Theorem 3.2), thereby obtaining

$$E(Dw_i, Q_{i+1}) = E(Dw_i, (\delta_\gamma/2)Q_i) \leq \frac{1}{2^5 2^{(n+2)/2}} E(Dw_i, (1/2)Q_i)\,,$$

where we have kept (4.46) in mind. In turn, let us estimate as follows:

$$\begin{aligned} E(Dw_i, (1/2)Q_i) &\leq \left(\fint_{(1/2)Q_i} |V(Dw_i) - (V(Dw_i))_{Q_i}|^2\, dx\, dt \right)^{1/2} \\ &\leq 2^{(n+2)/2} E(Dw_i, Q_i)\,. \end{aligned}$$

Connecting the inequalities in the last two displays gives

$$E(Dw_i, Q_{i+1}) \leq \tfrac{1}{2^5} E(Dw_i, Q_i)\,. \tag{4.48}$$

On the other hand, by (4.38) and (3.4) we have

$$\begin{aligned} &E(Du, Q_{i+1}) \\ &\leq \left(\fint_{Q_{i+1}} |V(Du) - (V(Dw))_{Q_{i+1}}|^2\, dx\, dt \right)^{1/2} \\ &\leq E(Dw_i, Q_{i+1}) + \left(\fint_{Q_{i+1}} |V(Du) - V(Dw_i)|^2\, dx\, dt \right)^{1/2} \\ &\leq E(Dw_i, Q_{i+1}) + \delta_1^{-(n+2)/2} \left(\fint_{Q_i} |V(Du) - V(Dw_i)|^2\, dx\, dt \right)^{1/2} \\ &\leq E(Dw_i, Q_{i+1}) + \tfrac{1}{200} E(Du, Q_i) + \sqrt{2\tilde{c}_0}\,\delta_1^{-(n+2)/2} \omega(r_i)(m_i + 1) \\ &\leq E(Dw_i, Q_{i+1}) + \tfrac{1}{200} E(Du, Q_i) + \sqrt{2\tilde{c}_0}\,\delta_1^{-(n+2)/2} \omega(r_i)\lambda^{p/2}\,. \end{aligned} \tag{4.49}$$

Similarly

$$E(Dw_i, Q_i) \leq 2E(Du, Q_i) + \sqrt{2\tilde{c}_0}\,\omega(r_i)\lambda^{p/2}\,.$$

Connecting this last inequality with (4.48) and (4.49) yields (4.47). □

Step 3: Iteration and conclusion. Recall that by the definitions in (4.29) and (4.30), we have

$$m_{i_e} + \delta_1^{-(n+2)/2} A_{i_e} \leq C_{i_e} \leq \frac{\lambda^{p/2}}{100}\,. \tag{4.50}$$

We now prove, by induction, that

$$1 + m_j + A_j \leq \frac{\lambda^{p/2}}{4} \tag{4.51}$$

holds whenever $j \geq i_e$. Indeed, by (4.50) and the choice in (4.33), the case $j = i_e$ of the previous inequality holds. Then, assume by induction that (4.51) holds whenever $j \in \{i_e, \ldots, i\}$, and this implies that (4.39) is verified for all $j \in \{i_e, \ldots, i\}$. Applying Lemma 4.7 estimate (4.47) implies

$$A_{j+1} \leq \tfrac{1}{4} A_j + 2\sqrt{\tilde{c}_0}\,\delta_1^{-(n+2)/2} \omega(r_i)\lambda^{p/2} \tag{4.52}$$

for all $j \in \{i_e, \ldots, i\}$. It immediately follows by (4.51) (assumed for all $j \in \{i_e, \ldots, i\}$), and (4.34), that

$$A_{i+1} \leq \frac{\lambda^{p/2}}{14}\,. \tag{4.53}$$

Furthermore, summing up (4.52) for $j \in \{i_e, \ldots, i\}$ gives

$$\sum_{j=i_e}^{i+1} A_j \le A_{i_e} + \frac{1}{4}\sum_{j=i_e}^{i} A_j + 2\sqrt{\tilde{c}_0}\,\delta_1^{-(n+2)/2}\sum_{j=i_e}^{i}\omega(r_i)\lambda^{p/2}\,,$$

yielding

$$\sum_{j=i_e}^{i+1} A_j \le 2A_{i_e} + 4\sqrt{\tilde{c}_0}\,\delta_1^{-(n+2)/2}\sum_{i=0}^{\infty}\omega(r_i)\lambda^{p/2}\,. \tag{4.54}$$

Next, notice that

$$\begin{aligned}
\int_0^{2r}\omega(\varrho)\,\frac{d\varrho}{\varrho} &= \sum_{i=0}^{\infty}\int_{r_{i+1}}^{r_i}\omega(\varrho)\,\frac{d\varrho}{\varrho} + \int_r^{2r}\omega(\varrho)\,\frac{d\varrho}{\varrho}\\
&\ge \sum_{i=0}^{\infty}\omega(r_{i+1})\int_{r_{i+1}}^{r_i}\frac{d\varrho}{\varrho} + \omega(r)\int_r^{2r}\frac{d\varrho}{\varrho}\\
&= \log\left(\frac{1}{\delta_1}\right)\sum_{i=0}^{\infty}\omega(r_{i+1}) + \log 2\,\omega(r) \ge \log 2\sum_{i=0}^{\infty}\omega(r_i)\,. \qquad (4.55)
\end{aligned}$$

Using the last inequality together with (4.54), and recalling (4.34), gives

$$\sum_{j=i_e}^{i+1} A_j \le 2A_{i_e} + \delta_1^{(n+2)/2}\frac{\lambda^{p/2}}{800}\,.$$

In turn, the last estimate and Hölder's inequality give

$$\begin{aligned}
m_{i+1} - m_{i_e} &= \sum_{j=i_e}^{i}(m_{j+1}-m_j) \le \sum_{j=i_e}^{i}\fint_{Q_{j+1}}|V(Du)-(V(Du))_{Q_j}|\,dx\,dt\\
&\le \sum_{j=i_e}^{i}\left(\fint_{Q_{j+1}}|V(Du)-(V(Du))_{Q_j}|^2\,dx\,dt\right)^{1/2}\\
&\le \delta_1^{-(n+2)/2}\sum_{j=i_e}^{i}\left(\fint_{Q_j}|V(Du)-(V(Du))_{Q_j}|^2\,dx\,dt\right)^{1/2}\\
&= \delta_1^{-(n+2)/2}\sum_{j=i_e}^{i}A_j \le 2\delta_1^{-(n+2)/2}A_{i_e} + \frac{\lambda^{p/2}}{800}
\end{aligned}$$

and thus it follows that

$$m_{i+1} \le m_{i_e} + 2\delta_1^{-(n+2)/2}A_{i_e} + \frac{\lambda^{p/2}}{800}\,.$$

In turn, by (4.50) the previous estimate yields $m_{i+1} \le \lambda^{p/2}/25$. The last inequality together with (4.33) and (4.53) allows to verify the induction step, i.e.

$$1 + m_{i+1} + A_{i+1} \le \frac{\lambda^{p/2}}{2000} + \frac{\lambda^{p/2}}{14} + \frac{\lambda^{p/2}}{25} \le \frac{\lambda^{p/2}}{4}\,.$$

Therefore (4.51) holds for every $i \ge i_e$. Estimate (4.3) finally follows with the choice $c_i \approx H_1^{2/p} + H_2^{2/p}$, since by the definition of Lebesgue points in $\mathcal{L}$ it holds that

$$|Du(x_0,t_0)|^{p/2} = \lim_{i\to\infty} m_i \le \frac{\lambda^{p/2}}{4}\,.$$

The proof is complete.

4.4. **Proof of Theorem 1.1.** We first treat the case $p \geq 2$. We now show how the intrinsic formulation of Theorem 4.1 implies the general a priori estimate of Theorem 1.1. To this end, let us consider the function

$$h(\lambda) := \lambda - c_i \lambda^{(p-2)/p} A(\lambda)\,,$$

where

$$A(\lambda) := \left(\frac{1}{|Q_r|} \int_{Q_r^\lambda} (|Du| + 1)^p \, dx\, dt \right)^{1/p}$$

and c_i is the constant appearing in Theorem 4.1. We consider the function $h(\cdot)$ defined for all those λ such that $Q_r^\lambda \subset \Omega_T$ such that r satisfies (4.2); observe that the domain of definition of $h(\cdot)$ includes $[1, \infty)$ as $Q_{2r}^\lambda \subset Q_{2r} \subset \Omega_T$ when $\lambda \geq 1$. Again, observe that $h(\cdot)$ is a continuous function and moreover $h(1) < 0$ as $c_i > 1$. On the other hand, as $Q_{2r}^\lambda \subset Q_{2r}$ for all $\lambda \geq 1$, we have

$$\lim_{\lambda \to \infty} h(\lambda) \geq \lim_{\lambda \to \infty} \left(\lambda - c_i \lambda^{\frac{p-2}{p}} B \right) = \infty\,,$$

where

$$B := \left(⨍_{Q_r} (|Du| + 1)^p \, dx\, dt \right)^{1/p} .$$

It follows that there exists a finite number $\lambda > 1$ such that $h(\lambda) = 0$, that is λ satisfies (4.1). Therefore we can apply Theorem 4.1 that together with Young's inequality with conjugate exponents $(p/(p-2), p/2)$ (when $p > 2$) gives

$$\begin{aligned} \lambda + |Du(x_0, t_0)| &\leq 2c\lambda^{\frac{p-2}{p}} \left(\frac{1}{|Q_r|} \int_{Q_r^\lambda} (|Du| + 1)^p \, dx\, dt \right)^{1/p} \\ &\leq \frac{\lambda}{2} + c \left(⨍_{Q_r} (|Du| + 1)^p \, dx\, dt \right)^{1/2} \end{aligned}$$

with $c \equiv c(n, N, p, \nu, L)$, from which (1.10) readily follows when $p \geq 2$. The case $2n/(n+2) < p < 2$ is instead treated as follows. We do consider cylinders of the type $Q_{r_\lambda}^\lambda(x_0, t_0) := B(x_0, \lambda^{(p-2)/2} r) \times (t_0 - r^2, t_0)$, where $r_\lambda = \lambda^{(p-2)/2} r$, that we are eventually going to use in Theorem 1.1. Notice also that, as now $p < 2$, we have

$$Q_{r_\lambda}^\lambda(x_0, t_0) \subset Q_r^\lambda(x_0, t_0)\,, \qquad \text{for} \;\; \lambda \geq 1\,. \tag{4.56}$$

This time we consider the function

$$h(\lambda) := \lambda - c_i \lambda^{(2-p)n/(2p)} A(\lambda)\,,$$

where, in turn,

$$A(\lambda) := \left(\frac{1}{|Q_r|} \int_{Q_{r_\lambda}^\lambda} (|Du| + 1)^p \, dx\, dt \right)^{1/p} .$$

The function $h(\cdot)$ is again defined for all those λ such that $Q_{r_\lambda}^\lambda \subset \Omega_T$; observe that this time the domain of definition of $h(\cdot)$ includes $[1, \infty)$ by (4.56). Notice that

$$p > \frac{2n}{n+2} \iff \frac{(2-p)n}{2p} < 1 \tag{4.57}$$

therefore, proceeding as for the case $p \geq 2$ we find $\lambda > 1$ such that $h(\lambda) = 0$. That is to say that (4.1) holds for the cylinder $Q_{r_\lambda}^\lambda(x_0, t_0)$; therefore, applying Theorem 4.1 yields

$$\lambda + |Du(x_0, t_0)| \leq 2c\lambda^{\frac{(p-2)n}{2p}} \left(\frac{1}{|Q_r|} \int_{Q_{r_\lambda}^\lambda} (|Du| + 1)^p \, dx\, dt \right)^{1/p} .$$

Thanks to (4.57) we can apply Young's inequality with conjugate exponents

$$\left(\frac{2p}{(2-p)n}, \frac{2p}{p(n+2)-2n}\right)$$

thereby obtaining, using (4.56), that

$$\lambda + |Du(x_0,t_0)| \le \frac{\lambda}{2} + c\left(\frac{1}{|Q_r|}\int_{Q_r^\lambda} (|Du|+1)^p \, dx\,dt\right)^{2/[p(n+2)-2n]},$$

from which (1.10) follows in the subquadratic case and the proof is complete.

5. Proof of Theorems 1.2-1.4

As for the twin Theorems 1.2-1.3, we shall actually give the full proofs in the case of general equations as in (1.11), while the proof for Theorem 1.2, that is for the vectorial model case in (1.2), can be obtained as in the following lines, by using the corresponding estimates in Section 3, with $s = 0$.

Important notational remark. Since we are restricting to Theorems 1.3-1.4, for the rest of the entire Section 5, we shall only use the map

$$V_s(z) = (s^2 + |z|^2)^{(p-2)/4} z\,,$$

where s is the number introduced in (1.12). Accordingly, when using the notion in (3.3), we shall simply denote $E_s(\cdot)$ as defined in (3.3). Finally, for the rest of the section we recommend to keep in mind the notation in (2.6).

5.1. Two lemmas.

In this Section 5.1, let us consider, in a fixed parabolic cylinder $Q_\varrho^\lambda \equiv Q_\varrho^\lambda(x_0,t_0) \Subset \Omega_T$, the unique solution w as in (4.5) to (4.21).

Lemma 5.1. *Let $\delta, \theta \in (0,1)$. Suppose that*

$$s^{p/2} + \left(\fint_{Q_\varrho^\lambda} |V_s(Du)|^2 \, dx\,dt\right)^{1/2} \le \lambda^{p/2} \qquad \text{and} \qquad \omega(\varrho) \le \frac{\delta^{n+2}\theta}{4^p c_V}\,, \tag{5.1}$$

where $c_V \equiv c_V(n,p,\nu,L)$ is as in Lemma 4.3. Then

$$s + \sup_{\frac12 Q_\varrho^\lambda} ||Dw|| \le c_b\lambda\,, \tag{5.2}$$

where $c_b \equiv c_b(n,p,\nu,L)$ is as in Theorem 4.2, and, moreover, the lower bound

$$\left(\fint_{\delta Q_\varrho^\lambda} |V_s(Du)|^2 \, dx\,dt\right)^{1/2} - \theta\lambda^{p/2} \le \left(\fint_{\delta Q_\varrho^\lambda} |V_s(Dw)|^2 \, dx\,dt\right)^{1/2} \tag{5.3}$$

holds.

Proof. Lemma 4.3 with Remark 4.1, in view of (5.1) and of (2.5), give

$$\left(\fint_{Q_\varrho^\lambda} |V_s(Du)-V_s(Dw)|^2 \, dx\,dt\right)^{1/2} \le 2^{p/2}c_V\omega(\varrho)\lambda^{p/2} \le \delta^{n+2}\theta\lambda^{p/2}\,. \tag{5.4}$$

This, again together with (5.1), further implies the bound

$$s^{p/2} + \left(\fint_{Q_\varrho^\lambda} |V_s(Dw)|^2 \, dx\,dt\right)^{1/2} \le 2\lambda^{p/2}$$

and therefore Theorem 4.2 immediately yields (5.2). Applying then (5.4) together with the triangle inequality yields

$$\left(\fint_{\delta Q_\varrho^\lambda} |V_s(Du)|^2 \, dx\,dt\right)^{1/2} \quad \le \quad \delta^{-(n+2)/2}\left(\fint_{Q_\varrho^\lambda} |V_s(Du)-V_s(Dw)|\, dx\,dt\right)^{1/2}$$

$$+\left(\fint_{\delta Q_\varrho^\lambda} |V_s(Dw)|^2\,dx\,dt\right)^{1/2},$$

which, together with (5.4), finally gives (5.3). □

Lemma 5.2. *Let $\varepsilon,\delta\in(0,1/2)$. Suppose that Dw satisfies the decay estimate*

$$E_s(Dw,\delta Q_\varrho^\lambda)\le 2^{-(n+5)}\varepsilon E_s(Dw,2^{-1}Q_\varrho^\lambda) \tag{5.5}$$

and that the first inequality in (5.1) *holds. Then we have*

$$E_s(Du,\delta Q_\varrho^\lambda)\le \tfrac{\varepsilon}{4}E_s(Du,Q_\varrho^\lambda)+4^p c_V\delta^{-(n+2)/2}\omega(\varrho)\lambda^{p/2},$$

where $c_V\equiv c_V(n,p,\nu,L)$ is as in Lemma 4.3.

Proof. Applying the triangle inequality, (3.4), and assumption (5.5), and finally using Lemma 4.3, we arrive at the following chain of inequalities:

$$\begin{aligned}
E_s(Du,\delta Q_\varrho^\lambda) &\le \left(\fint_{\delta Q_\varrho^\lambda}|V_s(Du)-(V_s(Dw))_{\delta Q_\varrho^\lambda}|^2\,dx\,dt\right)^{1/2}\\
&\le E_s(Dw,\delta Q_\varrho^\lambda)+\left(\fint_{\delta Q_\varrho^\lambda}|V_s(Du)-V_s(Dw)|^2\,dx\,dt\right)^{1/2}\\
&\le \frac{\varepsilon E_s(Dw,2^{-1}Q_\varrho^\lambda)}{2^{n+5}}+\delta^{-\frac{n+2}{2}}\left(\fint_{Q_\varrho^\lambda}|V_s(Du)-V_s(Dw)|^2\,dx\,dt\right)^{1/2}\\
&\le \frac{\varepsilon E_s(Du,2^{-1}Q_\varrho^\lambda)}{2^{n+5}}+2\delta^{-\frac{n+2}{2}}\left(\fint_{Q_\varrho^\lambda}|V_s(Du)-V_s(Dw)|^2\,dx\,dt\right)^{1/2}\\
&\le \frac{\varepsilon E_s(Du,2^{-1}Q_\varrho^\lambda)}{2^{n+5}}+4^p c_V\delta^{-\frac{n+2}{2}}\omega(\varrho)\lambda^{p/2},
\end{aligned}$$

for a suitable constant $c_V\equiv c_V(n,p,\nu,L)$. The result follows by observing that

$$E_s(Du,2^{-1}Q_\varrho^\lambda)\le\left(\fint_{\frac12 Q_\varrho^\lambda}|V_s(Du)-(V_s(Du))_{Q_\varrho^\lambda}|\,dx\,dt\right)^{1/2}\le 2^{\frac{n+2}{2}}E_s(Du,Q_\varrho^\lambda).$$

□

5.2. **Proof of Theorem 1.3.** The proof of Theorem 1.3 will be given in the case the case $2n/(n+2)<p<2$, while the one when $p\ge 2$, which is slightly simpler, can be obtained by minor modifications. Now, to begin with, let us fix an open subcylinder $\tilde Q\Subset\Omega_T$ such that $\tilde Q=\tilde\Omega\times(t_1,t_2)$, where $\tilde\Omega\Subset\Omega$ is a smooth subdomain, and let us take an intermediate cylinder $\tilde Q'$ such that $\tilde Q\Subset\tilde Q'\Subset\Omega_T$ and $\bar R_0:=\operatorname{dist}_{\mathrm{par}}(\tilde Q,\partial_{\mathrm{par}}\tilde Q')/100\approx\operatorname{dist}_{\mathrm{par}}(\tilde Q',\partial_{\mathrm{par}}\Omega_T)/100>0$. The assumptions of Theorem 1.3 imply those of Theorem 1.1, so that the gradient is locally bounded in Ω_T; in particular, Du is bounded in $\tilde Q'$. Consequently, we denote

$$\lambda_M^{p/2}:=1+4s^{p/2}+4\sup_{\tilde Q'}|V_s(Du)| \qquad\text{and}\qquad R_0:=\lambda_M^{(p-2)/2}\bar R_0/4. \tag{5.6}$$

The number λ_M depends only on the quantities $n,p,\nu,L,s,\|Du\|_{L^p}$ and $\bar R_0$; this follows by estimate (1.10) and a simple covering argument. Moreover, it follows that $Q_r^{\lambda_M}(x_0,t_0)\subset\tilde Q'$ whenever $(x_0,t_0)\in\tilde Q$ and $r\le R_0$, and using (2.5), that

$$s+\sup_{Q_r^{\lambda_M}}||Du||\le\lambda_M \qquad\text{whenever}\quad r\le R_0.$$

First, a VMO-type estimate.

Lemma 5.3. *Under the assumptions of Theorem 1.3, and with the notation of Section 5.2, for every $\varepsilon > 0$, there exists a radius of the type*

$$r_\varepsilon = \frac{\varepsilon^{1/\alpha_1}}{c_3} R(\varepsilon)\,, \quad \text{with} \quad \alpha_1 \in (0,1)\,, \qquad c_3 \geq 1\,, \qquad R(\varepsilon) \in (0, R_0] \tag{5.7}$$

such that

$$E_s(Du, Q_\varrho^{\lambda_M}(x_0,t_0)) < \lambda_M^{p/2}\varepsilon \tag{5.8}$$

holds whenever $\varrho \in (0, r_\varepsilon]$. Here $c_3 \equiv c_3(n,p,\nu,L)$ and $\alpha_1 \equiv \alpha_1(n,p,\nu,L)$ are positive constants, and $R(\varepsilon)$ denotes yet another radius such that $R(\varepsilon) \equiv R(n,p,\nu,L,\omega(\cdot))$. The radius $R(\varepsilon)$ is determined in (5.10) *below.*

Proof. With $\varepsilon > 0$ fixed in the statement of the Lemma, we choose the number $\delta_\gamma \equiv \delta_\gamma(n,p,\nu,L,\varepsilon) \in (0,1/2)$ in Theorem 3.2 with parameters

$$\lambda \equiv \lambda_M\,, \quad A \equiv c_b, B \equiv \sqrt{n}10^5\varepsilon^{-2/p}, \quad \gamma \equiv \varepsilon 2^{-(n+5)}\,,$$

where $c_b \equiv c_b(n,p,\nu,L)$ is the constant fixed in Theorem 4.2. Set $\delta_1 := \delta_\gamma/2$; by taking (3.7) into account we have

$$\delta_1 = \frac{\varepsilon^{1/\alpha_1}}{c_3}\,, \qquad \alpha_1 = \frac{2\alpha_0}{p+2}\,, \qquad c_3 \geq 1\,, \tag{5.9}$$

where α_1 and c_3 depend only on n,p,ν,L. We then choose $R(\varepsilon) \equiv R(n,p,\nu,L,\varepsilon) \in (0,R_0]$ such that

$$\omega(R(\varepsilon)) \equiv \omega(R) \leq \frac{\delta_1^{n+2}\varepsilon}{4^p 100 c_V}\,, \tag{5.10}$$

where c_V has been defined in Lemma 4.3 and Remark 4.1. Next, with $(x_0,t_0) \in \tilde{Q}$, we define the chain of shrinking intrinsic cylinders

$$Q_i \equiv Q_{r_i}^{\lambda_M}(x_0,t_0)\,, \qquad r_i = \delta_1^i r, \qquad r \in (\delta_1 R, R]\,. \tag{5.11}$$

Our next aim is to prove that

$$E_s(Du, Q_h) < \lambda_M^{p/2}\varepsilon \qquad \text{holds for every } h \in \mathbb{N} \cap [1,\infty)\,. \tag{5.12}$$

Let us single out a generic index $h \geq 1$ and distinguish two cases; the first is when

$$\left(\fint_{Q_h} |V_s(Du)|^2\,dx\,dt\right)^{1/2} < \frac{\lambda_M^{p/2}\varepsilon}{50}\,,$$

so that (5.12) follows immediately. The other case is obviously

$$\left(\fint_{Q_h} |V_s(Du)|^2\,dx\,dt\right)^{1/2} \geq \frac{\lambda_M^{p/2}\varepsilon}{50}\,. \tag{5.13}$$

Keeping Remark 4.1 in mind, we define w_{h-1} as the solution to the Cauchy-Dirichlet problem

$$\begin{cases} (w_{h-1})_t - \operatorname{div} a(x_0,t,Dw_{h-1}) = 0 & \text{in } Q_{h-1} \\ w_{h-1} = u & \text{on } \partial_{\mathrm{par}} Q_{h-1}\,. \end{cases}$$

The next step now consists in applying Lemma 5.1 with choices of parameters $\lambda \equiv \lambda_M \geq 1$, $\delta \equiv \delta_1$, $\theta \equiv \varepsilon/100$, $Q_\varrho^\lambda \equiv Q_{h-1}$, and $\delta Q_\varrho^\lambda \equiv Q_h$. Therefore (5.13) implies

$$\left(\fint_{Q_h} |V_s(Dw_{h-1})|^2\,dx\,dt\right)^{1/2} \geq \frac{\lambda_M^{p/2}\varepsilon}{100}$$

while, in turn, (2.5) applied together with the last inequality gives

$$\frac{\lambda_M\varepsilon^{2/p}}{10^5} \leq s + \sup_{Q_h}|Dw_{h-1}| \Longrightarrow \frac{\lambda_M\varepsilon^{2/p}}{\sqrt{n}10^5} \leq s + \sup_{Q_h}||Dw_{h-1}||\,.$$

Moreover, by (5.2) we have also

$$s+\sup_{\frac{1}{2}Q_{h-1}}||Dw_{h-1}||\leq c_b\lambda_M\equiv A\lambda_M\,.$$

Theorem 3.2 then gives

$$E_s(Dw_{h-1},Q_h)=E_s(Dw_{h-1},(\delta_\gamma/2)Q_{h-1})\leq 2^{-(n+5)}\varepsilon E_s(Dw_{h-1},2^{-1}Q_{h-1})$$

and hence Lemma 5.2, together with (5.6) and (5.10), implies

$$\begin{aligned} E_s(Du,Q_h) &\leq \frac{\varepsilon}{4}E_s(Du,Q_{h-1})+4^pc_V\delta_1^{-(n+2)/2}\omega(r_{h-1})\lambda_M^{p/2} \\ &\leq \frac{\lambda_M^{p/2}\varepsilon}{2}+\frac{\lambda_M^{p/2}\varepsilon}{100}\leq\lambda_M^{p/2}\varepsilon\,. \end{aligned} \tag{5.14}$$

This completes the proof of (5.12). Now, since the reasoning is independent of the choice of $(x_0,t_0)\in\tilde{Q}$ and *of the initial radius* $r\in(\delta_1R,R]$ chosen to build the chain in (5.11), we obtain (5.8) with the choice $r_\varepsilon=\delta_1R$. Indeed, let $\varrho\leq\delta_1R$; this means there exists an integer $m\geq1$ such that $\delta_1^{m+1}R<\varrho\leq\delta_1^mR$. Therefore we have $\varrho=\delta_1^mr$ for some $r\in(\delta_1R,R]$ and (5.8) follows from (5.12). The form in (5.7), follows from $r_\varepsilon=\delta_1R$ together with (5.9). The proof of Lemma 5.3 is complete. □

We now proceed with the proof of Theorem 1.3. Since the map $V_s(\cdot)$ is locally bi-Lipschitz it will be sufficient to show that $V_s(Du)$ is continuous. In turn, this will be shown using the fact that $V_s(Du)$ can be obtained as the (locally) uniform limit of a net of continuous functions. Specifically, with $(x_0,t_0)\in\tilde{Q}$, consider - obviously continuous functions -

$$(x_0,t_0)\to(V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)}\qquad\text{with}\quad\varrho\leq R_0,$$

where the radius R_0 has been determined in (5.6). We then prove that for every $\varepsilon>0$ there exists a radius $r_\varepsilon\leq R_0$, independent of the point (x_0,t_0) considered, such that

$$|(V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)}-(V_s(Du))_{Q_\tau^{\lambda_M}(x_0,t_0)}|\leq\lambda_M^{p/2}\varepsilon\qquad\forall\ \varrho,\tau\in(0,r_\varepsilon] \tag{5.15}$$

and $(x_0,t_0)\in\tilde{Q}$. This implies the existence of a continuous function to which $\{(V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)}\}$ converges locally uniformly; since, on the other hand, we have that

$$(V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)}\to V_s(Du(x_0,t_0))\qquad\text{as }\varrho\to0$$

holds almost everywhere, this implies that the precise representative of $V_s(Du)$ is continuous. We stress that inequality (5.15) will be proved *for every point* (x_0,t_0). The rest of the proof is now dedicated to show the validity of (5.15). To this aim, with $\varepsilon>0$ fixed in (5.15), we choose the number $\delta_\gamma\equiv\delta_\gamma(n,p,\nu,L,\varepsilon)\in(0,1/2)$ in Theorem 3.2 corresponding to the choice of parameters $\lambda\equiv\lambda_M$, $A\equiv c_b$, $B\equiv\sqrt{n}10^5\varepsilon^{-2/p}$ and $\gamma\equiv2^{-(n+5)}$, where $c_b\equiv c_b(n,p,\nu,L)$ is the constant fixed in Theorem 4.2. Again, we set $\delta_1:=\delta_\gamma/2$. Next, we take a positive radius $R\leq R_0$ such that

$$\omega(R)+\int_0^{2R}\omega(\varrho)\,\frac{d\varrho}{\varrho}\leq\frac{\delta_1^{n+2}\varepsilon}{4^{p+2}800c_V} \tag{5.16}$$

and

$$\sup_{0<\varrho<R}\ \sup_{(x_0,t_0)\in\tilde{Q}}E_s(Du,Q_\varrho^{\lambda_M}(x_0,t_0))\leq\frac{\delta_1^{n+2}\lambda_M^{p/2}\varepsilon}{800}\,. \tag{5.17}$$

Let us observe that it is possible to assume (5.17) by Lemma 5.3. We shall eventually show that the radius R determined by the smallness conditions (5.16)-(5.17) will work as r_ε in (5.15). Next, we again define the chain of shrinking intrinsic cylinders as in (5.11), with the new value of δ_1. We then have the following result, whose proof is exactly similar to the one for (5.14) from Lemma 5.3:

Lemma 5.4. *Assume that*

$$\left(\fint_{Q_{i+1}} |V_s(Du)|^2\,dx\,dt\right)^{1/2} \geq \frac{\lambda_M^{p/2}\varepsilon}{50} \qquad \text{and} \qquad \omega(r_i) \leq \frac{\delta_1^{n+2}\varepsilon}{4^p 100 c_V}\,.$$

Then it holds that

$$E_s(Du,Q_{i+1}) \leq \tfrac{1}{2}E_s(Du,Q_i) + 4^p c_V \delta_1^{-(n+2)/2}\omega(r_i)\lambda_M^{p/2}\,. \tag{5.18}$$

As a next step, we shall prove that

$$|(V_s(Du))_{Q_h} - (V_s(Du))_{Q_k}| \leq \frac{\lambda_M^{p/2}\varepsilon}{12} \tag{5.19}$$

holds whenever $0 \leq k \leq h$. For the proof we need some terminology. Given a chain $\{Q_i\}$ of geometrically shrinking intrinsic cylinders as in (5.11), we consider the set $\mathcal{L}$ defined by

$$\mathcal{L} := \left\{ i \in \mathbb{N} \ : \ \left(\fint_{Q_i} |V_s(Du)|^2\,dx\,dt\right)^{1/2} < \frac{\lambda_M^{p/2}\varepsilon}{50} \right\}\,,$$

and, accordingly we then define the set

$$\mathcal{C}_i^m = \{ j \in \mathbb{N} : i \leq j \leq i+m,\ i \in \mathcal{L},\ i+m+1 \in \mathcal{L}, \quad j \notin \mathcal{L} \text{ if } j > i\}$$

and call it *maximal iteration chain* of length m, starting at i. In other words, we have $\mathcal{C}_i^m = \{i,\ldots,i+m\}$ and each element of $\mathcal{C}_i^m$ but i lies outside of $\mathcal{L}$; $\mathcal{C}_i^m$ is maximal in the sense that there cannot be another set of the same type properly containing it. Obviously, such sets do not exist when $\mathcal{L} = \mathbb{N}$. In the same way we define $\mathcal{C}_i^\infty = \{j \in \mathbb{N} : i \leq j < \infty,\ i \in \mathcal{L},\ j \notin \mathcal{L} \text{ if } j > i\}$ as *the infinite maximal chain* starting at i. Notice that, in every case, the smallest element of such a chain always belongs to $\mathcal{L}$, being then the only one of the chain to have such a property. Moreover, we define $i_e := \min \mathcal{L}$. Note that we set $i_e = \infty$ if $\mathcal{L} = \emptyset$. We are now ready for the proof of (5.19); for this we need to distinguish three cases. We shall, without losing the generality, assume $0 \leq k < h$.

Case 1: $k < h \leq i_e$. Keeping (5.16) in mind, notice that if $h-1 > k$, then we can apply Lemma 5.4 repeatedly, and this yields the validity of (5.18) for every $i \in \{k,\ldots,h-2\}$. Summing up the previous inequalities, and making manipulations similar to those in (4.52)-(4.54) - we have

$$\sum_{i=k}^{h-1} E_s(Du,Q_i) \leq 2E_s(Du,Q_k) + 4^{p+1}c_V\delta_1^{-(n+2)/2}\lambda_M^{p/2}\sum_{i=k}^{h-2}\omega(r_i)\,.$$

In turn, using (4.55) we have

$$\sum_{i=k}^{h-1} E_s(Du,Q_i) \leq 2E_s(Du,Q_k) + 4^{p+2}c_V\delta_1^{-(n+2)/2}\int_0^{2r}\omega(\varrho)\,\frac{d\varrho}{\varrho}\lambda_M^{p/2}\,,$$

and using directly (5.17) for the case $h-1=k$, we conclude that in any case (i.e. $h-1 \geq k$)

$$\sum_{i=k}^{h-1} E_s(Du,Q_i) \leq \frac{\delta_1^{(n+2)/2}\lambda_M^{p/2}\varepsilon}{50}$$

holds as a consequence of (5.16)-(5.17). In turn, (5.19) follows since

$$|(V_s(Du))_{Q_h} - (V_s(Du))_{Q_k}| \ \leq \ \sum_{i=k}^{h-1}\left(\fint_{Q_{i+1}} |V_s(Du) - (V_s(Du))_{Q_i}|^2\,dx\,dt\right)^{1/2}$$

$$
\begin{aligned}
&\le \sum_{i=k}^{h-1}\left(\frac{|Q_i|}{|Q_{i+1}|}\right)^{1/2} E_s(Du,Q_i) \\
&= \delta_1^{-(n+2)/2}\sum_{i=k}^{h-1} E_s(Du,Q_i) \le \frac{\lambda_M^{p/2}\varepsilon}{50}\,. \qquad (5.20)
\end{aligned}
$$

Notice that the case analyzed here includes the one when the index i_e is infinite, i.e. the set $\mathcal{L}$ is empty.

Case 2: $i_e \le k < h$. Let us prove that in this case we have

$$
|(V_s(Du))_{Q_h}| \le \frac{\lambda_M^{p/2}\varepsilon}{25} \qquad \text{and} \qquad |(V_s(Du))_{Q_k}| \le \frac{\lambda_M^{p/2}\varepsilon}{25}\,. \tag{5.21}
$$

We prove the former inequality in (5.21), the proof of the latter being the same. If $h \in \mathcal{L}$, the first inequality in (5.21) follows immediately from the definition of $\mathcal{L}$. On the other hand, if $h \notin \mathcal{L}$, then, as $h \ge i_e$, it is possible to consider the maximal iteration chain $\mathcal{C}_{i_h}^{m_h}$ such that $h \in \mathcal{C}_{i_h}^{m_h}$; notice that $h > i_h$ as $h \notin \mathcal{L} \ni i_h$. Then iterating Lemma 5.4 as done in Case 1 - i.e. replacing k by i_h - we gain the analogue of (5.20), that is

$$
|(V_s(Du))_{Q_h} - (V_s(Du))_{Q_{i_h}}| \le \frac{\lambda_M^{p/2}\varepsilon}{50}\,.
$$

In turn using that $|(V_s(Du))_{Q_{i_h}}| \le \lambda_M^{p/2}\varepsilon/50$ as $i_h \in L$, we again obtain the first inequality in (5.21) and in any case (5.21) follows. Estimating as

$$
\begin{aligned}
|(V_s(Du))_{Q_h} - (V_s(Du))_{Q_k}| &\le |(V_s(Du))_{Q_h}| + |(V_s(Du))_{Q_k}| \\
&\le \frac{\lambda_M^{p/2}\varepsilon}{25} + \frac{\lambda_M^{p/2}\varepsilon}{25} \le \frac{\lambda_M^{p/2}\varepsilon}{12}
\end{aligned}
$$

we have that (5.19) holds in the second case too.

Case 3: $k < i_e < h$. Here we prove that (1.7) still holds and then we conclude as in Step 2. Indeed, the first inequality in (5.21) follows as in Case 2. As for the second estimate in (5.21), let us remark that, as $i_e \in \mathcal{L}$, we have that

$$
|(V_s(Du))_{Q_{i_e}}| \le \frac{\lambda_M^{p/2}\varepsilon}{50}\,. \tag{5.22}
$$

On the other hand, we can argue exactly as in Case 1, i.e. this time replacing h by i_e, thereby obtaining

$$
|(V_s(Du))_{Q_{i_e}} - (V_s(Du))_{Q_k}| \le \frac{\lambda_M^{p/2}\varepsilon}{50}
$$

that, together with (5.22), gives the second inequality in (5.21). In turn, (5.19) follows also in this case. The proof of (5.19) is now complete.

Finally, the proof of (5.15) follows using (5.19) together with the already proved VMO-regularity of the gradient, that is (5.17). Indeed, by taking $r_\varepsilon = R$ and fixing $0 < \tau < \varrho \le R$, there exists two integers, $0 \le k \le h$, such that $\delta_1^{k+1}R < \varrho \le \delta_1^k R$ and $\delta_1^{h+1}R < \tau \le \delta_1^h R$. Observe that

$$
\begin{aligned}
&|(V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)} - (V_s(Du))_{Q_{k+1}}| \\
&\le \left(\fint_{Q_{k+1}} |V_s(Du) - (V_s(Du))_{Q_\varrho^{\lambda_M}(x_0,t_0)}|^2\,dx\,dt\right)^{1/2} \\
&\le \delta_1^{-(n+2)/2} E_s(Du, Q_\varrho^{\lambda_M}(x_0,t_0)) \le \frac{\lambda_M^{p/2}\varepsilon}{10}\,, \qquad (5.23)
\end{aligned}
$$

where in the last line we have used (5.17). In the same way we also obtain

$$|(V_s(Du))_{Q_\tau^{\lambda_M}(x_0,t_0)} - (V_s(Du))_{Q_{h+1}}| \leq \frac{\lambda_M^{p/2}\varepsilon}{10}\,. \tag{5.24}$$

Using (5.23)-(5.24) together with (5.19), we conclude with (5.15), and the proof is complete.

Proof of Theorem 1.4. The proof revisits the one of Theorem 1.3, and makes essential use of Lemma 5.3; and in particular of the explicit dependence of the radius r_ε found in (5.7). For this reason we shall adopt the notation introduced in the proof of Theorem 1.3. Our aim is to show that, for every cylinder $\tilde{Q} \Subset \Omega_T$ as in Section 5.2, there exists a radius $R_1 > 0$, depending on n, p, ν, L, h, R_0, an exponent $h_1 \in (0,1)$, depending only on n, p, ν, L, h, *but independent of* λ_M, and finally a constant c, depending on n, p, ν, L, h, such that the decay estimate

$$E_s\left(Du, Q_\varrho^{\lambda_M}(x_0,t_0)\right) \leq c\lambda_M^{p/2}\varrho^{h_1} \tag{5.25}$$

holds whenever $\varrho \leq R_1$ and $(x_0,t_0) \in \tilde{Q}$, where $R_1 \equiv R_1(n,p,\nu,L,R_0)$. At this point, the local Hölder continuity of Du in Ω_T as described in the statement of Theorem 1.4 follows from a classical Campanato type integral characterization of the Hölder continuity originally observed by Da Prato [8]. In Lemma 5.3 we take $\varepsilon = \varrho$ with $\varrho \leq R_0$, where R_0 has been initially determined in (5.6). By recalling (5.9), verifying (5.10) amounts to take R (which equals $R(\varepsilon)$ in the notation of Lemma 5.3) such that

$$\omega(R) \leq \frac{\varrho^{\frac{n+2+\alpha_1}{\alpha_1}}}{4^p 100 c_V} \qquad \Longleftrightarrow \qquad R \leq \frac{\varrho^{\frac{n+2+\alpha_1}{\alpha_1 h}}}{c_4} =: \frac{\varrho^{\frac{1}{h_1}}}{c_4}\,,$$

for a new constant c_4 depending on n, p, ν, L, h. Using this relation in (5.8), and keeping in mind (5.7), we easily have that

$$E_s\left(Du, Q^{\lambda_M}_{\varrho^{1/h_1}/c_4}(x_0,t_0)\right) \leq \lambda_M^{p/2}\varrho$$

holds whenever $\varrho \leq R_1$, for a suitable R_1, from which (5.25) follows after changing variables. □

Acknowledgements. The authors are supported by the ERC grant 207573 “Vectorial Problems” and by the Academy of Finland project “Potential estimates and applications for nonlinear parabolic partial differential equations”. The authors also thank Paolo Baroni for remarks on a preliminary version of the paper and the referee for his/her careful reading of the manuscript.

References

[1] Acerbi E. & Mingione G.: Gradient estimates for a class of parabolic systems. *Duke Math. J.* 136 (2007), 285–320.

[2] Baroni P.: Regularity in parabolic Dini-continuous systems. *Forum Math.* 23 (2011), 1281–1322.

[3] Byun S.S.: Optimal $W^{1,p}$ regularity theory for parabolic equations in divergence form. *J. Evol. Equ.* 7 (2007), 415–428.

[4] Byun S.S. & Wang L.: Parabolic equations with BMO nonlinearity in Reifenberg domains. *J. Reine Angew. Math. (Crelle J.)* 615 (2008), 1–24.

[5] Byun S.S. & Wang L.: Fourth-order parabolic equations with weak BMO coefficients in Reifenberg domains. *J. Diff. Equ.* 245 (2008), 3217–3252.

[6] Campanato S.: Equazioni paraboliche del secondo ordine e spazi $L^{2,\theta}(\Omega,\delta)$. *Ann. Mat. Pura Appl. (IV)* 73 (1966), 55–102.

[7] Chipot M. & Evans L. C.: Linearisation at infinity and Lipschitz estimates for certain problems in the calculus of variations. *Proc. Roy. Soc. Edinburgh Sect. A* 102 (1986), 291–303.

[8] Da Prato G.: Spazi $\mathcal{L}^{p,\theta}(\Omega,\delta)$ e loro proprietà. *Ann. Mat. Pura Appl. (IV)* 69 (1965), 383–392.

[9] DiBenedetto E.: *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993.
[10] DiBenedetto E. & Friedman A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. reine ang. Math. (Crelles J.)* 357 (1985), 1–22.
[11] Dolzmann G. & Kristensen J.: Higher integrability of minimizing Young measures. *Calc. Var.* 22 (2005), 283–301.
[12] Foss M.: Global regularity for almost minimizers of nonconvex variational problems. *Ann. Mat. Pura e Appl. (IV)* 187 (2008), 263-231.
[13] Hamburger C.: Regularity of differential forms minimizing degenerate elliptic functionals. *J. reine angew. Math. (Crelles J.)* 431 (1992), 7–64.
[14] Kinnunen J. & Parviainen M.: Stability for degenerate parabolic equations. *Adv. Calc. Var.* 3 (2010), 29–48.
[15] Kinnunen J. & Lewis J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* 102 (2000), 253–271.
[16] Kinnunen J. & Lindqvist P.: Summability of semicontinuous supersolutions to a quasilinear parabolic equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (V)* 4 (2005), 59–78.
[17] Kristensen J. & Mingione G.: The singular set of minima of integral functionals. *Arch. Ration. Mech. Anal.* 180 (2006), 331–398.
[18] Kristensen J. & Mingione G.: The singular set of Lipschitzian minima of multiple integrals. *Arch. Ration. Mech. Anal.* 184 (2007), 341–369.
[19] Korte R. & Kuusi T.: A note on the Wolff potential estimate for solutions to elliptic equations involving measures. *Adv. Calc. Var.* 3 (2010), 99–113.
[20] Kuusi T. & Mingione G.: The Wolff gradient bound for degenerate parabolic equations. *Preprint 2010.*
[21] Kuusi T. & Mingione G.: Gradient regularity for nonlinear parabolic equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (V)*, in press.
[22] Kuusi T. & Mingione G.: Nonlinear potential estimates in parabolic problems. *Rend. Lincei - Mat. Appl.* 22 (2011), 161–174.
[23] Leone C. & Passareli di Napoli A. & Verde A.: Lipschitz regularity for some asymptotically subquadratic problems. *Nonlinear Anal. TMA* 67 (2007), 1532–1539.
[24] Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Differential Equations* 76 (1988), 203–212.
[25] Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis.* University of Washington, St. Louis.
[26] Jin T. & Mazya V. & Van Schaftingen J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. *Comptes Rendus Mathematique* 347 (2009), 773-778.
[27] Mingione G.: Bounds for the singular set of solutions to non linear elliptic systems. *Calc. Var.* 18 (2003), 373–400.
[28] Mingione G.: Regularity of minima: an invitation to the dark side of the Calculus of Variations. *Applications of Mathematics* 51 (2006), 355–426.
[29] Misawa M.: Local Hölder regularity of gradients for evolutional p-Laplacian systems. *Ann. Mat. Pura Appl. (IV)* 181 (2002), 389–405.
[30] Misawa M.: Partial regularity results for evolutional p-Laplacian systems with natural growth. *Manuscripta Math.* 109 (2002), 419–454.
[31] Misawa M. & Ogawa T.: Regularity condition by mean oscillation to a weak solution of the 2-dimensional harmonic heat flow into sphere. *Calc. Var.* 33 (2008), 391–415.
[32] Raymond J.P.: Lipschitz regularity of solutions of some asymptotically convex problems. *Proc. Roy. Soc. Edinburgh Sect. A* 117 (1991), 59–73.
[33] Scheven C. & Schmidt T.: Asymptotically regular problems. I. Higher integrability. *J. Diff. Equ.* 248 (2010), 745–791.

Tuomo Kuusi, Aalto University Institute of Mathematics, P.O. Box 11100 FI-00076 Aalto, Finland
E-mail address: `tuomo.kuusi@tkk.fi`

Dipartimento di Matematica, Università di Parma, Parco Area delle Scienze 53/a, Campus, 43124 Parma, Italy
E-mail address: `giuseppe.mingione@unipr.it.`