THE ROYAL SWEDISH ACADEMY OF SCIENCES

INSTITUT MITTAG-LEFFLER

Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

Borderline gradient continuity for nonlinear parabolic systems

T. Kuusi and G. Mingione

REPORT No. 11, 2013/2014, fall

ISSN 1103-467X
ISRN IML-R- -11-13/14- -SE+fall

BORDERLINE GRADIENT CONTINUITY FOR NONLINEAR PARABOLIC SYSTEMS

TUOMO KUUSI AND GIUSEPPE MINGIONE

ABSTRACT. We consider the evolutionary p-Laplacean system

$$\partial_t u - \triangle_p u = F, \qquad p > \frac{2n}{n+2}$$

in cylindrical domains of $\mathbb{R}^n \times \mathbb{R}$, and prove the continuity of the spatial gradient Du under the Lorentz space assumption $F \in L(n+2,1)$. When F is time independent the condition improves in $F \in L(n,1)$. This is the limiting case of a result of DiBenedetto claiming that Du is Hölder continuous when $F \in L^q$ for $q > n+2$. At the same time, this is the natural nonlinear parabolic analog of a linear result of Stein, claiming the gradient continuity of solutions to the linear elliptic system $\triangle u \in L(n,1)$ is continuous. New potential estimates are derived and moreover suitable nonlinear potentials are used to describe fine properties of solutions.

CONTENTS

1. GRADIENT CONTINUITY AND POTENTIALS

In this paper we are initially interested in the problem of finding sharp conditions ensuring the local (spatial) gradient continuity of solutions $u\colon \Omega_T \to \mathbb{R}^m$ to the evolutionary p-Laplacean system

$$\partial_t u - \operatorname{div}(|Du|^{p-2}Du) = F\,, \qquad \Omega_T = \Omega \times (-T,0)\,, \tag{1.1}$$

in terms of the regularity of the assigned vector field $F\colon \Omega_T \to \mathbb{R}^m$. Here $\Omega \subset \mathbb{R}^n$ denotes an open subset, while $n \geq 2$ and $m \geq 1$. By solutions to (1.1), we mean the usual distributional energy solutions, that is, maps

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

solving (1.1) in the sense of distributions. In the case $p=2$, and already in the scalar case $m=1$, i.e., in the case of the heat equation

$$\partial_t u - \triangle u = F\,, \tag{1.2}$$

solutions can be obtained via convolution with the fundamental solution. This allows to deduce that the Lorentz space regularity condition

$$F \in L(N,1) \Longleftrightarrow \int_0^\infty |\{(x,t) \in \Omega_T \,:\, |F(x,t)| > \lambda\}|^{1/N} d\lambda < \infty \tag{1.3}$$

is sufficient in order to prove the L^∞-gradient regularity. As usual, the number

$$N := n+2$$

denotes the usual parabolic dimension. On the other hand a basic result of DiBenedetto [7] asserts that, for solutions to (1.1), it holds that

$$F \in L^{N+\varepsilon} \Longrightarrow Du \text{ is locally Hölder continuous}\,. \tag{1.4}$$

Note indeed that $L^{N+\varepsilon} \subset L(N,1)$ whenever $\varepsilon > 0$, such inclusion being actually strict. It is therefore natural to ask for a borderline version of (1.4) leading to gradient continuity, or, in other words, for a an extension of the classical gradient continuity criterion (1.3) known for solutions to (1.2). Due to the structure properties of (1.1), one expects that, no matter the degeneracy of the system in question, condition (1.3) still guarantees the continuity of Du, being eventually optimal for dimensional reasons. This is intuitively clear when looking at the asymptotic/steady solutions to (1.1), that is when considering $\triangle_p u = 0$. In fact, first formally decouple this system into in the "linearised" one given by

$$\operatorname{div} W = 0\,, \qquad W := |Du|^{p-2}Du\,.$$

Then, observing that the continuity of W implies that of Du and vice-versa, one is eventually led to think that a borderline condition of the continuity of Du should be independent of p. Ultimately the regularity conditions of F should coincide with the ones that hold for the heat system (1.2). This is on the other hand already transparent from (1.4). A first step towards answering this question has been given in [25], where we were able to prove the local boundedness of Du under the assumption $F \in L(N,1)$ for solutions to the p-Laplacean system; unfortunately, the techniques [25] are not suitable to get continuity results. A further step has been made in [21], where the continuity of Du, again under the assumption $F \in L(N,1)$, has been proved in the scalar case $m = 1$ as a consequence of potential estimates that are again insufficient to cover the vectorial case. In this paper we finally obtain the following conclusive result, that in fact also extends to systems with coefficients satisfying minimal regularity assumptions:

Theorem 1.1. *Let u be a solution to the system*

$$\partial_t u - \operatorname{div}\left(b(x,t)|Du|^{p-2}Du\right) = F\,, \qquad \textit{with } p > \frac{2n}{n+2} \tag{1.5}$$

and assume that

- *the vector field $F\colon \Omega \to \mathbb{R}^m$ satisfies $F \in L(N,1)$ locally in Ω_T*
- *the measurable function $b\colon \Omega \to [\nu, L]$ is Dini-continuous with respect to the space variable x, where $0 < \nu \leq L < \infty$.*

Then Du is continuous in Ω_T.

A few comments on the previous assumptions are now in order. The most relevant feature of the previous theorem is that the condition $F \in L(N,1)$ ensuring the continuity of Du is, as already emphasized above, independent of p. Finding such a condition is indeed the main technical challenge since it implies that the rate of degeneracy of the system plays no role in the gradient continuity analysis. We also recall that the lower bound $p > 2n/(n+2)$ is again essential, as already in the case of one scalar equations with $F \equiv 0$; solutions are in general unbounded when such an assumption is not fulfilled, see [7].

As for the coefficients $b(\cdot)$, we recall that saying that the function $b(\cdot)$ is Dini-continuous with respect to x means that there exits a concave, non-decreasing modulus of continuity $\omega\colon [0,\infty) \to [0,1]$ with $\omega(0) = 0$, satisfying

$$|b(x,t) - b(y,t)| \leq L\omega(|x-y|) \tag{1.6}$$

for every $x, y \in \Omega$ and $t \in (-T, 0)$ and such that

$$\int_0 \omega(\varrho)\, \frac{d\varrho}{\varrho} < \infty\,. \tag{1.7}$$

The Dini-continuity with respect to the space variable is known to be sharp starting from the simplest possible example of linear elliptic equations of the form $\operatorname{div}(b(x)Du) = 0$. In this case, in [19] it has been proved that the gradient of solutions is in general unbounded when coefficients are merely continuous but not Dini-continuous. It is worth remarking that very recently there has been a renewed interested in the study of how Dini-continuity of coefficients affects regularity of solutions, see for instance in [32].

When the right hand side datum F is independent of time, we expect that the condition ensuring the continuity of the gradient coincides with the one known in the elliptic stationary case $-\triangle_p u = F$. In that case the condition $F \in L(n,1)$ turns out to be sharp already for the boundedness of Du. See [4, 5, 6, 12] for comments on such issues and the proof of Theorem 1.2 in Section 6 below for more on Lorentz spaces. As a matter of fact, the proof of Theorem 1.1 also leads to establish the following:

Theorem 1.2. *Let u and $b(\cdot)$ be as in Theorem 1.1 and assume further that the vector field $F(\cdot)$ satisfies*

$$|F(x,t)| \leq g(x)\,, \tag{1.8}$$

where the function $g\colon \Omega \to [0,\infty]$ belongs to the Lorentz space $L(n,1)$. Then Du is continuous in Ω_T. In particular, Du is continuous when $F(\cdot)$ is time independent and belongs to $L(n,1)$.

Theorem 1.2 considerably improves the results in [25], where in the case $\partial_t u - \triangle_p u = F(x) \in L(n,1)$ we were only able to prove the local boudnedness of Du; moreover, in [25], and for purely technical reasons, the case $n = 2$ was still left embarrassingly open while now we indeed cover it.

Theorems 1.1-1.2 naturally relate to a classical result of Stein [36], claiming that Sobolev functions whose gradient lies in $L(n,1)$ are actually continuous. As a corollary, it follows that if $u \in W^{1,2}$ solves the linear system $\triangle u \in L(n,1)$, then Du is continuous. Theorems 1.1-1.2 can be then considered as the natural and nonlinear parabolic analog of this last result.

The ones presented above are not the only results of this paper. They are indeed consequences of more general potential estimates and criteria that are going to be described in the next section, and that we think are of independent interest.

Remark 1.1. Before going on let us observe that since all the regularity results we are going to obtain are local in nature and only depend on the integrability properties of F, whenever we shall consider solutions to (1.5) we shall always assume that $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^m$ is indeed defined on the whole space $\mathbb{R}^{n+1}$. Moreover, whenever we shall consider an assumption such as for instance $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$, we shall always consider F to be defined on the whole $\mathbb{R}^{n+1}$ by setting $F \equiv 0$ outside Ω_T. This turns out top be useful when considering certain nonlinear potentials.

1.1. **Intrinsic geometries and modified potentials.** The approach we are going to adopt in this paper, in order to achieve the proof of Theorems 1.1-1.2, passes through new potentials estimates. They are natural extensions to the vectorial case of the scalar ones previously introduced in the papers [26, 27, 28]. There the boundedness and continuity properties of Du have been obtained through the use of *intrinsic potential estimates*. We remark that, recently, there has been a great interest in nonlinear potential type estimates in the context of degenerate equations of p-Laplacean type: see for instance [9, 10, 11]. The technique introduced here produces new type of estimates that reveal a rather interesting linkage between the type of potentials used and the so-called intrinsic local geometry of the system. Let us recall that in the case of the usual heat equation $\partial_t u - \triangle u = \mu$ - with μ being a general Radon measure - both solutions and their gradients can be pointwise estimated by mean of caloric Riesz potentials

$$\mathbf{I}^{\mu}_{\beta}(x_0,t_0;r) := \int_0^r \frac{|\mu|(Q_\varrho(x_0,t_0))}{\varrho^{N-\beta}}\,\frac{d\varrho}{\varrho}\,, \qquad 0<\beta\leq N\,. \tag{1.9}$$

Here $Q_r(x_0,t_0) := B(x_0,r)\times(t_0-r^2,t_0)$ denotes the standard parabolic cylinders with vertex at (x_0,t_0); see next section for more notation. The use of such potentials eventually allows to get sharp regularity criteria both for solutions and their gradients. When treating the evolutionary p-Laplacean operator for $p\neq 2$, the analysis unavoidably involves the concept of *intrinsic geometry*, introduced and widely employed by DiBenedetto [7, 38, 2]. This means that, in order to rebalance the lack of homogeneity of an operator as $u\to\partial_t u - \operatorname{div}(|Du|^{p-2}Du)$, one is led to carry out the analysis on cylinders whose size depends on the solution itself. More precisely, one considers “intrinsic cylinders” of the type

$$Q_r^\lambda(x_0,t_0) := B(x_0,r)\times(t_0-\lambda^{2-p}r^2,t_0)\,, \qquad \lambda>0\,, \tag{1.10}$$

where a condition as

$$\fint_{Q_r^\lambda(x_0,t_0)} |Du|\,dx\,dt \approx \lambda$$

is simultaneously satisfied; the term “intrinsic” stems from the very fact that the parameter λ appears both in the definition of the cylinder and in the bound to be satisfied on it. This, in some sense, locally rebalances the anisotropicity of the operator that now looks as $u\to\partial_t u - \operatorname{div}(\lambda^{p-2}Du)$ on the cylinder Q_r^λ and therefore ultimately as the heat operator after a change of variables. The final outcome is that homogeneous local estimates, suitable for the relevant regularity iteration schemes, then hold for solutions on such cylinders. According to this viewpoint, in [27, 28] we have introduced a family of intrinsic linear potentials working for the p-Laplacean operator as usual linear potentials do in the case of the Laplacean. More precisely, we have introduced *intrinsic Riesz potentials* of the type

$$\mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r) := \int_0^r \frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\varrho^{N}}\,d\varrho\,,$$

which are indeed built using intrinsic cylinders rather than standard ones. The number $\lambda>0$ is still a free parameter, while the linkage with the solution making the previous potential “intrinsic” comes later, in the formulation of the estimates. The outcome is that the gradient of solutions to the evolutionary p-Laplacean equation $\partial_t u - \operatorname{div}(|Du|^{p-2}Du) = \mu$ can be pointwise estimated, in an optimal way, by $\mathbf{I}^{\mu}_{1,\lambda}$ on intrinsic cylinders. In this paper we give a vectorial analog of these potential bounds, in fact introducing a new family of this time nonlinear potentials, aimed at matching in a peculiar way the intrinsic geometry of the problem with a new family of estimates. Specifically, we shall consider the following modified intrinsic potentials:

$$\mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r) := \int_0^r \left(\frac{|F|^\gamma(Q^\lambda_\varrho(x_0,t_0))}{\varrho^N}\right)^{1/\gamma} d\varrho \quad \text{and} \quad \mathbf{P}^F_\gamma := \mathbf{P}^F_{\gamma,1}\,, \tag{1.11}$$

where $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ in the case $\gamma > 1$ and where, slightly abusing the notation, F is a vector valued signed Borel measure when $\gamma = 1$. Here we are denoting, in a standard way,

$$|F|^\gamma(Q^\lambda_r(x_0,t_0)) := \int_{Q^\lambda_r(x_0,t_0)} |F|^\gamma \, dx\, dt$$

whenever $\gamma > 1$. The potentials $\mathbf{P}^F_{\gamma,\lambda}$ appearing in (1.11) will play in the vectorial context the same role of the Riesz caloric potentials displayed in (1.9), while pointwise potential estimates will be obtained via their use. As a matter of fact, we notice that $\mathbf{P}^F_{1,\lambda} = \mathbf{I}^F_{1,\lambda}$. For instance, potentials $\mathbf{P}^F_\gamma$ *are sufficient to determine the fine properties of the gradient* Du, exactly as the classical Riesz potentials do for linear problems, see Theorem 1.7 below. Surprisingly enough, the new phenomenon emerging here is that a new family of gradient pointwise potential estimates can be obtained via potentials $\mathbf{P}^F_{\gamma,\lambda}$. Such estimates improve on when γ increases although the homogeneity and regularising properties of the considered potentials remain γ-invariant. Interestingly enough, the improvement happens when looking for criteria implying low degree integrability of Du. On the other hand all such estimates uniformly yield the same optimal criteria when looking for gradient boundedness and continuity. Specifically, when considering the potential $\mathbf{P}^F_{\gamma,\lambda}$, we shall always consider the following *rebalancing exponent*:

$$a \equiv a(\gamma) := p - 1 - \frac{p-2}{\gamma}\,. \tag{1.12}$$

This number appears in the estimates and links the potential $\mathbf{P}^F_{\gamma,\lambda}$ to the gradient via the local intrinsic geometry. Let us now come to a detailed description of the results. As it is common, we distinguish the so-called degenerate case $p \geq 2$ from the singular one $2n/(n+2) < p < 2$. This is necessary because the shape of the estimates change due to the fact that in the case $p < 2$ the results must match with those available in the context of equations with measure data.

Theorem 1.3. *Let u be a solution to the system* (1.5) *for $p \geq 2$, where $b(\cdot)$ is as in Theorem 1.1 and $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ with $\gamma \geq n/2$. Then the pointwise potential estimate*

$$|Du(x_0,t_0)| \leq c\left[\mathbf{P}^F_\gamma(x_0,t_0;r)\right]^{1/a} + c\left[\fint_{Q_r(x_0,t_0)} (|Du|^p + 1)\, dx\, dt\right]^{1/2} \tag{1.13}$$

holds for every standard parabolic cylinder $Q_r(x_0,t_0) \subset \Omega_T$ such that (x_0,t_0) is Lebesgue point of Du; the constant c depends only on $n, m, p, \nu, L, \omega(\cdot)$.

Notice that the improving character of (1.13) as γ increases relies on the fact that, while the homogeneity properties of $\mathbf{P}^F_\gamma$, which dictates the behaviour of the operator $F \to \mathbf{P}^F_\gamma$ in various function spaces, are independent of γ, the function $\gamma \to a(\gamma)$ is then increasing. In particular, we notice that $a(\infty) = p - 1$.

For the subquadratic case we instead consider a different range of parameters for γ, limited from below by the duality exponent $(p^*)'$. This assumption guarantees that the right hand side vector field $F(\cdot)$ belongs to the dual of the natural parabolic energy space $C^0(-T,0;L^2) \cap L^p(-T,0;W^{1,p})$. We observe that this is necessary in this context since going below this exponent means falling into the realm of problems with measure data. In this case so-called very weak solutions appear, i.e., solutions that do not belong to the natural energy space. For such solutions methods and results are different and the kind of gradient regularity available completely changes. Specifically, we shall consider the following condition:

$$\gamma \geq (p^*)' = \frac{p(n+2)}{p(n+2)-n}\,, \qquad p^* = \frac{p(n+2)}{n}\,. \tag{1.14}$$

Here p^* is the standard parabolic Sobolev conjugate exponent, see Proposition 3.1 below (together with related, preceding remarks) for the motivation of this terminology. We also refer to Remark 5.1 for more on this aspect.

Theorem 1.4. *Let u be a solution to the system* (1.5) *for* $2n/(n+2) < p < 2$, *where* $b(\cdot)$ *is as in Theorem 1.1 and* $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ *as in* (1.14)*. Then the pointwise potential estimate*

$$\begin{aligned} |Du(x_0,t_0)| \ &\leq\ c\left[\mathbf{P}^F_\gamma(x_0,t_0;r)\right]^{2\gamma/[\gamma p+n(p-2)]} \\ &\quad + c\left[\fint_{Q_r(x_0,t_0)} (|Du|^p+1)\,dx\,dt\right]^{2/[p(n+2)-2n]} \end{aligned}$$

holds provided $\gamma < N$ *and for every standard parabolic cylinder* $Q_r(x_0,t_0) \subset \Omega_T$ *such that* (x_0,t_0) *is Lebesgue point of* Du*; the constant* c *depends only on* $n, m, p, \nu, L, \omega(\cdot)$*. In the case* $\gamma \geq N$ *we instead have again*

$$|Du(x_0,t_0)| \leq c\left[\mathbf{P}^F_\gamma(x_0,t_0;r)\right]^{1/a} + c\left[\fint_{Q_r(x_0,t_0)} (|Du|^p+1)\,dx\,dt\right]^{2/[p(n+2)-2n]}\,.$$

We finally observe that estimates of the last two theorems recover and extend the gradient bounds available in the literature for homogeneous systems of evolutionary p-Laplacean type. Indeed, when $F \equiv 0$, then we recover the classical L^∞-bounds obtained by DiBenedetto for the case $\partial_t u - \triangle_p u = 0$ in [7] and the ones in [29] for the evolutionary p-Laplacean system with coefficients $\partial_t u - \operatorname{div}(b(x,t)|Du|^{p-2}Du) = 0$. These estimates read as

$$\sup_{\frac12 Q_r} |Du| \lesssim \left[\fint_{Q_r} (|Du|^p+1)\,dx\,dt\right]^{\max\{1/2,2/[p(n+2)-2n]\}}$$

and follow by Theorems 1.3-1.4 via a standard covering argument.

Remark 1.2. In both Theorems 1.3-1.4 the main inequalities actually hold *for every* point (x_0,t_0). For the sake of clarity we prefer to give a separate presentation of this, see Theorem 1.7 below.

1.2. **Intrinsic formulations.** The estimates featured in Theorems 1.3-1.4 exhibit a natural inhomogeneity with respect to the gradient, which is the obvious consequence of the inhomogeneity of the evolutionary p-Laplacean system. When instead passing to intrinsic cylinders estimates become homogeneous and this also happens when considering potentials, as shown by the following:

Theorem 1.5. *Let u be a solution to the system* (1.5) *for* $p > 2n/(n+2)$, *where* $b(\cdot)$ *is as in Theorem 1.1 and* $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ *with* γ *as Theorem 1.3-1.4. There exist a constant* $c > 1$ *and a radius* $R_0 > 0$, *both depending only on* $n, m, p, \nu, L, \omega(\cdot)$, *such that the following implication holds:*

$$\begin{aligned} c\left[\mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r)\right]^{1/a} + c\left(\fint_{Q_r^\lambda(x_0,t_0)} |Du|^p\,dx\,dt\right)^{1/p} &\leq \lambda \\ \Longrightarrow |Du(x_0,t_0)| &\leq \lambda \end{aligned} \tag{1.15}$$

whenever $Q_r^\lambda(x_0,t_0) \subset \Omega_T$, (x_0,t_0) *is a Lebesgue point of* Du, *and* $r \leq R_0$*. When the vector field* $a(\cdot)$ *is independent of* x, *no restriction occurs on* r, *i.e.,* $R_0 = \infty$.

The previous theorem indeed features a homogenous estimate that allows to recover the classical L^∞ conditional estimate of DiBenedetto for solutions to the system $\partial_t u - \operatorname{div}(|Du|^{p-2}Du) = 0$. This is the following implication:

$$c\left(\fint_{Q_r^\lambda(x_0,t_0)} |Du|\,dx\,dt\right)^{1/p} \le \lambda \Longrightarrow |Du(x_0,t_0)| \le \lambda$$

for a constant c depending only on n, m, p. It is important to remark that in fact Theorems 1.3-1.4 are both corollaries of Theorem 1.5, which features the most general potential estimate possible.

The methods developed to get (1.15) can be used to get the precise continuity criterion displayed in the next Theorem, from which Theorems 1.1-1.2 actually follow.

Theorem 1.6. *Let u be a solution to the system* (1.5) *for $p > 2n/(n+2)$, where $b(\cdot)$ is as in Theorem 1.1 and $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ with γ as in Theorems 1.3-1.4. Assume that*

$$\lim_{\varrho\to 0} \mathbf{P}_\gamma^F(x,t;\varrho) = 0 \qquad \text{uniformly with respect to } (x,t)\,. \tag{1.16}$$

Then Du is continuous.

1.3. **Fine properties of solutions.** The potentials $\mathbf{P}_\gamma^F$ efficiently replace the usual Riesz potentials in their classical role of describing the fine pointwise properties of solutions. In particular, they allow to determine at which points a so-called precise representative of the gradient can be determined. In the nonlinear case, for solutions to elliptic equations, this can be done via the use of certain nonlinear potentials called Wolff potentials; see [30] for precise results and [20, 21] for the use of Wolff potentials. The following result states that it is possible to achieve the same conclusions in the parabolic vectorial case:

Theorem 1.7 (Fine gradient analysis)**.** *Let u be a solution to the system* (1.5) *for $p > 2n/(n+2)$, where $b(\cdot)$ is as in Theorem 1.1 and $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ with γ as Theorem 1.3-1.4. Assume that for $(x_0,t_0) \in \Omega_T$ it holds that*

$$\mathbf{P}_\gamma^F(x_0,t_0;\varrho) < \infty\,. \tag{1.17}$$

Then (x_0,t_0) is a Lebesgue point of Du, that is, the limit

$$\lim_{\varrho\to 0} \frac{1}{|Q_\varrho(x_0,t_0)|} \int_{Q_\varrho(x_0,t_0)} Du(x,t)\,dx\,dt$$

exists. As a consequence, the estimates of Theorems 1.3-1.5 hold for every point (x_0,t_0) once passing to the precise representative of the spatial gradient Du.

1.4. **A few technical points.** The starting point for the approach of this paper is obviously the one introduced in [25] for the scalar case $m = 1$, in the sense that we will retrieve from that paper a few iteration schemes that will eventually serve to recover potentials in certain summation processes. Unfortunately, the largest part of the a priori estimates valid when dealing with a single equation do not any longer apply here. This implies that the same iteration methods have to be deeply modified to match the current form of the estimates. The main tool is given at this stage by modified potentials defined in (1.11); in turn their definition perfectly matches the shape of the relevant a priori estimates once using the correct intrinsic geometry. This is exactly the point where the rebalancing exponent a defined in (1.12) enters to describe the interplay between the adopted potentials and the local geometry of the system. In order to provide the reader with a brief outline of the main arguments, let us describe a heuristic sketch of the proof of Theorem 1.5 for the simplest model problem (1.1), and in the degenerate case $p \ge 2$. Adapting the arguments for Theorem 1.5 and using different values of λ will eventually lead to continuity assertions for Du. To prove (1.15) we consider a shrinking sequence of dyadic intrinsic cylinders $Q_j \equiv Q^\lambda_{r_j}(x_0,t_0)$, with a geometrically decaying sequence r_i, for which

$$\lambda \lesssim \left(\fint_{Q_j} |Du|^p \, dx \, dt \right)^{1/p} \tag{1.18}$$

holds. Notice that, thanks to a suitable exit time argument, we can always reduce to such a case since otherwise (1.15) follows trivially. At this point one can consider the lifted maps solving $\partial_t w_j - \triangle_p w_j = 0$ in Q_j and coinciding with u on the parabolic boundary on Q_j. Using a localisation argument based on the regularity theory available for w_j together with the information contained in (1.18), we can conclude with the *density information* $\lambda \lesssim |Dw_j|$ on Q_{j+1}. In other words, we are dealing with systems that, up to multiplicative factors in the diffusion part, are non-degenerate. We can therefore perform a comparison argument between u and w_j as, also thanks to (1.18), $u - w_j$ would solve the perturbed heat system on Q_j

$$\partial_t(u - w_j) - \lambda^{p-2} \triangle (u - w_j) = F\,. \tag{1.19}$$

This point needs anyway to be framed in the context of an iteration scheme where, at each stage, the intrinsic geometric condition

$$\left(\fint_{Q_j} |Du|^p \, dx \, dt \right)^{1/p} \lesssim \lambda \tag{1.20}$$

is satisfied, and this is again part of the iteration scheme. Equation (1.19) now allows to perform a careful comparison argument whose ultimate outcome is the following inequality:

$$\begin{aligned} \left(\fint_{Q_{j+1}} |Du - Dw_j|^{p'} \, dx \, dt \right)^{1/p'} &\lesssim \lambda^{1-a} \left(\frac{|F|^\gamma(Q_j(x_0,t_0))}{r_j^{N-\gamma}} \right)^{1/\gamma} \\ &=: \lambda^{1-a} \mu_j\,, \end{aligned} \tag{1.21}$$

where $p' = p/(p-1)$ denotes the standard conjugate exponent of p. Standard telescoping summation gives

$$|Du(x_0,t_0)| \leq \sum_j |(Du)_{Q_{j+1}} - (Du)_{Q_j}| + \left(\fint_{Q^\lambda_r(x_0,t_0)} |Du|^p \, dx \, dt \right)^{1/p}\,.$$

Then, using (1.21) and a rather involved induction procedure, we can estimate the intermediate term by the sum of the terms μ_j and in turn by the potential

$$\sum_j |(Du)_{Q_{j+1}} - (Du)_{Q_j}| \lesssim \lambda^{1-a} \sum_j \mu_j \lesssim \lambda^{1-a} \mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r)$$

so that (1.15) follows by the last two displays. Notice that, in the framework of the iteration scheme, the previous estimate also serves to establish (1.20) at the next step, that is on the cylinder Q_{j+1}. The previous argument is completely heuristic and in fact needs a large amount of technical tools to become rigorous. In particular, essentially all the aspects of regularity theory of solutions to evolutionary p-Laplacian system must be used and recombined in a rather delicate way in order to achieve the final estimates.

2. Notation

We shall denote by c a general positive constant, possibly varying from line to line; special occurrences will be denoted by $c_1, c_2, \bar{c}_1, \bar{c}_2$ or the like. All these constants will always be larger or equal than one; moreover relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n, m, p, \nu, L)$ means that c_1 depends only on n, m, p, ν, L. We 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. Intrinsic parabolic cylinders with vertex (x_0, t_0) and width $r > 0$ have been defined in (1.10), while we recall that standard parabolic cylinders are defined by $Q_r(x_0, t_0) := Q_r^1(x_0, t_0)$. These are the balls with respect to the canonical parabolic metric in $\mathbb{R}^{n+1}$, that is

$$d_{\mathrm{par}}((x,t),(x_0,t_0)) := \max\left\{|x - x_0|, \sqrt{|t - t_0|}\right\} .$$

Very often when dealing with cylinders sharing the same vertex we shall omit to indicate it, simply denoting $Q_r \equiv Q_r(x_0, t_0)$ and $Q_r^\lambda \equiv Q_r^\lambda(x_0, t_0)$. In this paper, the letter λ will always denote a positive number. Accordingly, we consider their dyadic, parabolic dilations

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

for $\tau > 0$. We recall that if $Q = \mathcal{A} \times (t_1, t_2)$ is a cylindrical domain, the usual parabolic boundary of Q is $\partial_{\mathrm{par}} Q := (\mathcal{A} \times \{t_1\}) \cup (\partial \mathcal{A} \times [t_1, t_2))$, and this is nothing else but the standard topological boundary without the upper cap $\bar{\mathcal{A}} \times \{t_2\}$.

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

$$(g)_{\mathcal{O}} \equiv \fint_{\mathcal{O}} g(x,t)\, dx\, dt := \frac{1}{|\mathcal{O}|} \int_{\mathcal{O}} g(x,t)\, dx\, dt$$

its integral average; here $|\mathcal{O}|$ denotes the Lebesgue measure of $\mathcal{O}$. A similar notation is adopted if the integral is only in space or time. Given a vector field $g(x,t)$ with values in $\mathbb{R}^{mn}$, typically a gradient, we shall use the standard L^s-excess functional for $s \geq 1$ defined as

$$E_s(g, Q) := \left(\fint_Q |g - (g)_Q|^s \, dx\, dt\right)^{1/s} , \tag{2.1}$$

where Q is a cylinder. We shall several times use the following elementary property of the excess:

$$\left(\fint_Q |g - (g)_Q|^s \, dx\, dt\right)^{1/s} \leq 2 \left(\fint_Q |g - \tilde{g}|^s \, dx\, dt\right)^{1/s} , \tag{2.2}$$

whenever $\tilde{g} \in \mathbb{R}^{mn}$ and $s \geq 1$. The oscillation of g on any subset $\mathcal{O} \subset \mathbb{R}^{n+1}$ is instead defined as

$$\operatorname*{osc}_{\mathcal{O}} g := \sup_{(x,t),(\tilde{x},\tilde{t}) \in \mathcal{O}} |g(x,t) - g(\tilde{x},\tilde{t})| .$$

3. Preparations for the proofs

3.1. **Basic regularity theory for homogeneous systems.** In this section we are going to recall a few basic results from the regularity theory of solutions to degenerate parabolic systems of p-Laplacean type. In particular, we are interested in solutions $w \in C^0(T; L^2(B, \mathbb{R}^m)) \cap L^p(T; W^{1,p}(B, \mathbb{R}^m))$ to systems with coefficients of the type

$$\partial_t w - \operatorname{div}(b(x,t)|Dw|^{p-2}Dw) = 0 \qquad \text{in } \mathcal{C} := B \times T\,, \tag{3.1}$$

where $B \subset \mathbb{R}^n$ is an open subset and $T \subset \mathbb{R}$ is an interval; the function $b(\cdot)$ is as in Theorem 1.1. Next we recall a few results for solutions $v \in C^0(T; L^2(B, \mathbb{R}^m)) \cap L^p(T; W^{1,p}(B, \mathbb{R}^m))$ to homogeneous systems with space variable independent coefficients

$$\partial_t v - \operatorname{div}(\tilde{b}(t)|Dv|^{p-2}Dv) = 0 \qquad \text{in } \mathcal{C} := B \times T\,. \tag{3.2}$$

In this last case we assume that the function $\tilde{b}\colon T \to \mathbb{R}$ is measurable and again satisfies $\nu \leq \tilde{b}(t) \leq L$ as in Theorem 1.1. In what follows, the main references we are going to adapt the results from are [7, 28, 29].

We start with an L^∞- gradient bound taken from [29]. The novelty relies on the fact that it is stated for systems with Dini-continuous coefficients. In the case when coefficients are not depending on x, the next theorem is a classical result of DiBenedetto [7] (see also [29, Theorem 4.2] for a similar statement).

Theorem 3.1 (Intrinsic $L^p - L^\infty$ bound)**.** *Let w be as in* (3.1)*. There exists a radius R_1, depending only on $n, m, p, \nu, L, \omega(\cdot)$ such that if $Q_r^\lambda \subset B \times T$ is an intrinsic cylinder such that $r \leq R_1$ then the following implication:*

$$\left(\fint_{Q_r^\lambda} |Dw|^p \, dx \, dt \right)^{1/p} \leq c_1 \lambda \quad \Longrightarrow \quad \sup_{\frac{1}{2}Q_r^\lambda} |Dw| \leq c_2 \lambda$$

holds for a function

$$n, m, p, \nu, L, c_1 \to c_2(n, m, p, \nu, L, c_1)\,, \tag{3.3}$$

which is increasing with respect to c_1. The same conclusion obviously holds for the map v defined in (3.2) *and in this case no restriction on r is needed.*

We now proceed with oscillation estimates for Dw. The following result can be obtained using those stated in [29] along the lines described in [28].

Theorem 3.2 (Gradient continuity)**.** *Let w be as in* (3.1)*. Then Dw has a continuous representative in $\mathcal{C}$. Moreover, let R_1 be as in Theorem 3.1, ad let $Q_r^\lambda \subset \mathcal{C}$ be a cylinder with $r \leq R_1$ and such that*

$$\sup_{\frac{1}{2}Q_r^\lambda} |Dw| \leq A\lambda$$

holds for some $A \geq 1$. Then, for any $\delta \in (0,1)$ there exists a positive constant $\sigma_1 \equiv \sigma_1(n, m, p, \nu, L, A, \delta, \omega(\cdot)) \in (0, 1/4)$ such that

$$\operatorname*{osc}_{\sigma_1 Q_r^\lambda} Dw \leq \delta\lambda\,.$$

We next give a suitable extension of a result proved in [29, Theorem 3.1].

Theorem 3.3 (Conditional decay estimate)**.** *Let v be as in* (3.2)*, let $Q_r^\lambda \subset \mathcal{C}$ be a cylinder and $s \geq 1$. Consider numbers $A, B \geq 1$ and $\bar{\varepsilon} \in (0,1)$. Then there exists a constant $\sigma_2 \in (0, 1/4)$ depending only on $n, m, p, \nu, L, A, B, \bar{\varepsilon}$, but not on s, such that if*

$$\frac{\lambda}{B} \leq \sup_{Q_{\sigma_2 r}^\lambda} |Dv| \leq \sup_{\frac{1}{4}Q_r^\lambda} |Dv| \leq A\lambda \tag{3.4}$$

holds, then

$$\left(\fint_{\tau Q_r^\lambda} |Dv-(Dv)_{\tau Q_r^\lambda}|^s\,dx\,dt\right)^{1/s} \le \bar\varepsilon \left(\fint_{\frac14 Q_r^\lambda} |Dv-(Dv)_{\frac14 Q_r^\lambda}|^s\,dx\,dt\right)^{1/s} \tag{3.5}$$

holds too whenever $\tau \in (0,\sigma_2]$.

Proof. This theorem has been essentially proved in [29, Theorem 3.1] using a different excess quantity, namely replacing Dv by $V(Dv)$ in (3.5), and then taking $s=2$. The modifications to get the statement here are minimal, and we ask the reader to consult [29]. Referring [29, Proposition 3.2] we note that the last inequality at page 750 - that is standard Campanato's theory for linear parabolic systems with continuous coefficients - is actually valid with every exponent s replacing 2, and this means that at the end [29, Proposition 3.2] and in particular [29, (3.14)] holds with the choice $E_s(Dw,Q_r^\lambda)$ instead of $E_2(Dw,Q_r^\lambda)$. In turn this yields [29, Proposition 3.2] with the same choice of the excess functional appearing in the last display contained there. From this point on the proof of Theorem 3.3 follows with minor modifications from the one of [29, Theorem 3.1]. □

The next theorem encodes the basic Hölder continuity properties of the spatial gradient of solutions to (3.2); the result is retrievable from [7, 8]. Another proof can be also inferred starting from Theorem 3.3, as shown in [27].

Theorem 3.4 (Gradient Hölder continuity)**.** *Let v be as in* (3.2)*, let $Q_r^\lambda \subset \mathcal{C}$ be a cylinder and $s\ge 1$. Then Dv has a Hölder continuous representative locally in Q_r^λ. Moreover if*

$$\sup_{\frac12 Q_r^\lambda} |Dv| \le A\lambda$$

holds for a certain constant $A\ge 1$, then

$$\operatorname*{osc}_{Q_\varrho^\lambda} Dv \le c_3\lambda\left(\frac{\varrho}{r}\right)^\alpha$$

holds whenever $\varrho \le r/2$, for constants $c_3 \equiv c_3(n,m,p,\nu,L,A)\ge 1$ and $\alpha \equiv \alpha(n,N,p,\nu,L,A)\in(0,1)$. Here $Q_\varrho^\lambda \subset Q_r^\lambda$ are intrinsic cylinders sharing the same vertex.

3.2. **Basic setup for the proofs.** Given an intrinsic cylinder $Q_{2r}^\lambda(x_0,t_0)\subset \Omega_T$ of the type in (1.10), we consider a family of nested parabolic cylinders

$$Q_j \equiv B_j\times T_j \equiv B(x_0,r_j)\times(t_0-\lambda^{2-p}r_j^2,t_0)\subseteq \Omega_T\,, \qquad r_j := \sigma^j r\,, \tag{3.6}$$

for a fixed decay parameter $\sigma\in(0,1/4)$. With the chain of shirking cylinders $\{Q_j\}$ being given, we canonically define two sequences of related comparison maps, namely $\{w_j\}$ and $\{v_j\}$ as follows; we let

$$w_j \in C^0(T_j;L^2(B_j))\cap L^p(T_j;W^{1,p}(B_j))$$

be the unique solution to the Cauchy-Dirichlet problem

$$\begin{cases} \partial_t w_j - \operatorname{div}(b(x,t)|Dw_j|^{p-2}Dw_j) = 0 & \text{in } Q_j \\ w_j = u & \text{on } \partial_{\mathrm{par}} Q_j\,. \end{cases} \tag{3.7}$$

After having defined w_j, we also define

$$v_j \in C^0(t_0-\lambda^{2-p}r_j^2/4,t_0;L^2(\tfrac12 B_j))\cap L^p(t_0-\lambda^{2-p}r_j^2/4,t_0;W^{1,p}(\tfrac12 B_j))$$

as the unique solution to the frozen Cauchy-Dirichlet problem

$$\begin{cases} \partial_t v_j - \operatorname{div}(b(x_0,t)|Dv_j|^{p-2}Dv_j) = 0 & \text{in } \frac12 Q_j \\ v_j = w_j & \text{on } \partial_{\mathrm{par}}\left(\frac12 Q_j\right). \end{cases} \tag{3.8}$$

The monotonicity properties of the p-Laplacean operator can be summarised by the following basic inequality

$$(|z_1|+|z_2|)^{p-2}|z_1-z_2|^2 \leq c(n,m,p)\langle |z_1|^{p-2}z_1-|z_2|^{p-2}z_2, z_1-z_2\rangle\,, \tag{3.9}$$

which holds for all matrixes $z_1, z_2 \in \mathbb{R}^{mn}$ and $p>1$. In order to use (3.9) more efficiently it is convenient to use the auxiliary vector field $V\colon \mathbb{R}^{mn}\to\mathbb{R}^{mn}$ defined as

$$V(z) := |z|^{(p-2)/2}z\,, \tag{3.10}$$

together with the related inequality (see [15]):

$$\frac{|z_1-z_2|}{c} \leq \frac{|V(z_1)-V(z_2)|}{(|z_1|+|z_2|)^{(p-2)/2}} \leq c|z_1-z_2|\,, \tag{3.11}$$

which is valid for all matrixes $z_1, z_2 \in \mathbb{R}^{mn}$ that are not simultaneously null and for $p>1$. The constant c depends only on n, m, p. A basic consequence of (3.11) and Young's inequality is the following estimate, that holds in the range $p\in(1,2]$ and for a constant c, again depending only on n,m,p:

$$|z_1-z_2| \leq c|V(z_1)-V(z_2)|^{2/p} + c|V(z_1)-V(z_2)||z_2|^{(2-p)/2}\,. \tag{3.12}$$

We are ready to get a few preliminary energy estimates for the maps u, w_j, v_j. Subtracting the equations for u and w_j, we have

$$(u-w_j)_t - \operatorname{div}\,(b(x,t)|Du|^{p-2}Du - b(x,t)|Dw_j|^{p-2}Dw_j) = F\,.$$

By multiplying the previous equation by $u-w_j$, a standard argument that can be made rigorous via Steklov averages (see for instance [1]) allows to gain the following standard energy estimate:

$$\begin{aligned} &\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0} \int_{B_j} |u-w_j|^2(x,\tau)\,dx \\ &\qquad + \int_{Q_j} b(x,t)\langle |Du|^{p-2}Du - |Dw_j|^{p-2}Dw_j, Du-Dw_j\rangle\,dx\,dt \\ &\qquad\qquad \leq c\int_{Q_j} |F||u-w_j|\,dx\,dt \end{aligned}$$

so that using (3.11) and passing to averages we get

$$\begin{aligned} &\lambda^{p-2} \sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0} -\!\!\!\!\!\!\int_{B_j} \left|\frac{u-w_j}{r_j}\right|^2 (x,\tau)\,dx \\ &\qquad + -\!\!\!\!\!\!\int_{Q_j} |V(Du)-V(Dw_j)|^2\,dx\,dt \leq c -\!\!\!\!\!\!\int_{Q_j} |F||u-w_j|\,dx\,dt\,, \end{aligned} \tag{3.13}$$

which is an inequality that holds in the full range $p>2n/(n+2)$. In a similar way as in [1] and [29, Lemma 4.3], the inequality in (1.6) implies

$$-\!\!\!\!\!\!\int_{\frac{1}{2}Q_j} |V(Dv_j)-V(Dw_j)|^2\,dx\,dt \leq c[\omega(r_j)]^2 -\!\!\!\!\!\!\int_{\frac{1}{2}Q_j} |Dw_j|^p\,dx\,dt\,, \tag{3.14}$$

for a constant $c\equiv c(n,m,p,\nu,L)$, which is again an inequality that holds in the whole range $p>2n/(n+2)$. By triangle inequality, and recalling the definition of the vector field $V(\cdot)$ given in (3.10), we have $|Dv_j| = |V(Dv_j)|^{2/p} \leq |V(Dw_j)|^{2/p} + c|V(Dv_j)-V(Dw_j)|^{2/p}$, and therefore we also have

$$-\!\!\!\!\!\!\int_{\frac{1}{2}Q_j} |Dv_j|^p\,dx\,dt \leq c -\!\!\!\!\!\!\int_{\frac{1}{2}Q_j} |Dw_j|^p\,dx\,dt\,. \tag{3.15}$$

Moreover, in the case $p \geq 2$ (3.14) implies

$$\fint_{\frac{1}{2}Q_j} |Dv_j - Dw_j|^p \, dx \, dt \leq \bar{c}_0 [\omega(r_j)]^2 \fint_{\frac{1}{2}Q_j} |Dw_j|^p \, dx \, dt \tag{3.16}$$

again for a constant $\bar{c}_0 \equiv \bar{c}_0(n, m, p, \nu, L)$. In a similar manner it is convenient to introduce the quantities

$$S_j := \lambda^{p-2} \sup_{t_0 - \lambda^{2-p} r_j^2 < \tau < t_0} \fint_{B_j} \left| \frac{u - w_j}{r_j} \right|^2 (x, \tau) \, dx + \fint_{Q_j} |Du - Dw_j|^p \, dx \, dt \,. \tag{3.17}$$

In view of (3.13) and of (3.11), when $p \geq 2$ we deduce that for a constant c depending only on n, m, p, ν, L it holds that

$$S_j \leq c r_j \fint_{Q_j} |F| \left| \frac{u - w_j}{r_j} \right| \, dx \, dt \,. \tag{3.18}$$

Once the parameter σ has been fixed and the shrinking chain of cylinders $\{Q_j\}$ has been built as in (3.6), the potential $\mathbf{P}^F_{\lambda,\gamma}(x_0, t_0; 2r)$ admits a decomposition in terms of a related series. Indeed, we shall prove that

$$\sum_{j=0}^{\infty} \mu_j \leq \sigma^{-N/\gamma} \mathbf{P}^F_{\gamma,\lambda}(x_0, t_0; 2r) \tag{3.19}$$

holds, where

$$\mu_j \equiv \mu_j(x_0, t_0) := \left(\frac{|F|^\gamma (Q^\lambda_{r_j}(x_0, t_0))}{r_j^{N-\gamma}} \right)^{1/\gamma} = \left(r_j^{\gamma - N} \int_{Q_j} |F|^\gamma \, dx \, dt \right)^{1/\gamma} . \tag{3.20}$$

Notice, by the way, that

$$\mu_j \leq \sigma^{1-N/\gamma} \mu_{j-1} \tag{3.21}$$

holds whenever $j \geq 1$. Going back to the potentials decomposition, we indeed have

$$
\begin{aligned}
\mathbf{P}^F_{\gamma,\lambda}(x_0, t_0; 2r) &= \int_0^{2r} \left(\frac{|F|^\gamma (Q^\lambda_\varrho(x_0, t_0))}{\varrho^N} \right)^{1/\gamma} d\varrho \\
&= \sum_{j=0}^{\infty} \int_{r_{j+1}}^{r_j} \left(\frac{|F|^\gamma (Q^\lambda_\varrho(x_0, t_0))}{\varrho^{N-\gamma}} \right)^{1/\gamma} \frac{d\varrho}{\varrho} + \int_r^{2r} \left(\frac{|F|^\gamma (Q^\lambda_\varrho(x_0, t_0))}{\varrho^{N-\gamma}} \right)^{1/\gamma} \frac{d\varrho}{\varrho} \\
&\geq \sum_{j=0}^{\infty} \left(\frac{|F|^\gamma (Q_{j+1})}{r_j^{N-\gamma}} \right)^{1/\gamma} \int_{r_{j+1}}^{r_j} \frac{d\varrho}{\varrho} + \left[\frac{|F|^\gamma (Q_0)}{(2r_0)^{N-\gamma}} \right]^{1/\gamma} \int_{r_0}^{2r_0} \frac{d\varrho}{\varrho} \\
&= \sigma^{N/\gamma - 1} \log \left(\frac{1}{\sigma} \right) \sum_{i=0}^{\infty} \mu_{j+1} + \frac{\log 2}{2^{N/\gamma - 1}} \mu_0 \geq \sigma^{N/\gamma} \sum_{j=0}^{\infty} \mu_j \,.
\end{aligned} \tag{3.22}
$$

Finally, we fix a few exponents that will of be of repeated use in the rest of the paper. The Sobolev parabolic conjugate exponent p^* has already been introduced in (1.14). This terminology is motivated by the next standard inequality (see [7]), where, upon choosing $q_2 = p$ and $q_3 = 2$, it follows $q_1 = p^*$.

Proposition 3.1. *Let $v \in L^\infty(T_j; L^{q_3}(B_j)) \cap L^{q_2}(T_j; W^{1,q_2}_0(B_j))$ for $q_2, q_3 \geq 1$. There exists a constant c depending only on n, q_2, q_3 such that the following inequality holds for $q_1 = q_2(n + q_3)/n$:*

$$\fint_{Q_j} \left| \frac{v}{r_j} \right|^{q_1} dx \, dt$$

$$\leq c\left(\fint_{Q_j}|Dv|^{q_2}\,dx\,dt\right)\left(\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{v}{r_j}\right|^{q_3}(x,\tau)\,dx\right)^{q_2/n}.$$

The last relevant exponent we are going to use throughout the paper is q, which is defined as follows:

$$q := \begin{cases} \dfrac{p(n+2)}{n+2p} & \text{if} \qquad p \geq 2 \\ \dfrac{p(n+2)}{p(n+2)-n} & \text{if} \quad \dfrac{2n}{n+2} < p < 2\,. \end{cases} \tag{3.23}$$

We notice that the previous definition implies

$$q := \begin{cases} \left[\dfrac{p'(n+2)}{n}\right]' & \text{if} \qquad p \geq 2 \\ \left[\dfrac{p(n+2)}{n}\right]' & \text{if} \quad \dfrac{2n}{n+2} < p < 2 \end{cases} \tag{3.24}$$

and in particular we have $(p^*)' \leq q$ with $q = (p^*)'$ when $2n/(n+2) < p < 2$. We finally recall that the exponent $a(\gamma)$ remains fixed as in (1.12).

3.3. **Preliminary estimates in the case $p \geq 2$.** In this section we establish a few basic estimates, which will play a key role in the rest of the paper. Such estimates apply to the case $p \geq 2$, while the corresponding ones for the case $2n/(n+2) < p < 2$ will be derived in the subsequent Section 3.4.

Lemma 3.1. *Let u be a solution to the system in* (1.5) *with $p \geq 2$ and let w_j be defined as in* (3.7) *with $j \geq 0$. Assume that $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ for some $\gamma \geq (p^*)'$. There exist constants $\bar{c}_1 \equiv \bar{c}_1(n,m,p,\nu)$ and $\bar{c}_2 \equiv \bar{c}_2(n,m,p,\nu,\sigma)$ such that*

$$\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{1/p} \leq \bar{c}_1\lambda\left(\frac{\mu_j}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}} \tag{3.25}$$

holds for $j \geq 0$, and, when $j \geq 1$, the following inequality holds too:

$$\left(\fint_{Q_j}|Dw_j-Dw_{j-1}|^p\,dx\,dt\right)^{1/p} \leq \bar{c}_2\lambda\left(\frac{\mu_{j-1}}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}}. \tag{3.26}$$

Proof. In order to use (3.18) we apply Hölder's inequality to get

$$\begin{aligned} r_j\fint_{Q_j}|F|\left|\frac{u-w_j}{r_j}\right|dx\,dt \\ &\leq r_j\left(\fint_{Q_j}|F|^{(p^*)'}\,dx\,dt\right)^{1/(p^*)'}\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*} \\ &\leq c\left(\frac{r_j^{\gamma-N}}{\lambda^{2-p}}\int_{Q_j}|F|^\gamma\,dx\,dt\right)^{1/\gamma}\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*}, \end{aligned} \tag{3.27}$$

where c only depends on n . To estimate the last integral we use Proposition 3.1 (with the choice of the exponents described before its statement), thereby getting

$$\begin{aligned} \left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*} &\leq c\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{1/p^*} \\ &\quad\cdot\left(\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx\right)^{p/(np^*)} \end{aligned}$$

and therefore we have, recalling the definition of S_j in (3.17), that

$$\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*}\le \lambda^{\frac{(2-p)p}{np^*}}S_j^{\frac{n+p}{np^*}}=\lambda^{\frac{2-p}{n+2}}S_j^{\frac{n+p}{p(n+2)}}\,. \tag{3.28}$$

We remark that estimates in displays (3.27)-(3.28) are valid in the case $2n/(n+2)<p<2$ as well. Combining the content of the last display with (3.18) and (3.27) yields

$$S_j\le \lambda^{\frac{2-p}{n+2}}\left(\frac{r_j^{\gamma-N}}{\lambda^{2-p}}\int_{Q_j}|F|^\gamma\,dx\,dt\right)^{1/\gamma}S_j^{\frac{n+p}{p(n+2)}}\,,$$

so that, dividing by $S_j^{(n+p)/(np+2p)}$ when $S_j>0$ (observe that when $S_j=0$ there is nothing to prove), gives

$$\left(\int_{Q_1}|Du-Dw_j|^p\,dx\,dt\right)^{1/p}\le c\lambda^{\frac{2-p}{n(p-1)+p}}\left[\left(\frac{r_j^{\gamma-N}}{\lambda^{2-p}}\int_{Q_j}|F|^\gamma\,dx\,dt\right)^{1/\gamma}\right]^{\frac{n+2}{n(p-1)+p}}\,.$$

Now (3.25) follows by the definition of the number $a(\gamma)$ in (1.12) and the definition of μ_j in (3.20). Finally, notice that

$$\begin{aligned}\left(\fint_{Q_j}|Dw_j-Dw_{j-1}|^p\,dx\,dt\right)^{1/p}&\le\left(\frac{|Q_{j-1}|}{|Q_j|}\right)^{1/p}\left(\fint_{Q_{j-1}}|Du-Dw_{j-1}|^p\,dx\,dt\right)^{1/p}\\&\qquad+\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{1/p}\\&\le \bar c_1\left(\frac{|Q_{j-1}|}{|Q_j|}\right)^{1/p}\lambda\left(\frac{\mu_{j-1}}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}}+\bar c_1\lambda\left(\frac{\mu_j}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}}\end{aligned}$$

so that (3.26) follows recalling the definition of r_j, r_{j-1} and (3.21). □

Lemma 3.2. *Let u be a solution to the system in* (1.5) *with $p\ge 2$ and let w_j, v_j be defined as in* (3.7)-(3.8)*, respectively, with $j\ge 1$. Assume that $F\in L^\gamma(\Omega_T,\mathbb{R}^m)$ for some $\gamma\ge q$, where q is defined in* (3.23)*. Suppose further that*

$$\mu_{j-1}\le\lambda^a\,, \tag{3.29}$$

and that the bounds

$$\sup_{\frac12 Q_j}|Dw_j|\le A\lambda \qquad\text{and}\qquad \frac{\lambda}{A}\le|Dw_{j-1}|\le A\lambda\qquad\text{in } Q_j \tag{3.30}$$

hold for some constant $A\ge 1$. Then there exists a constant $\bar c_3$ depending only on n,m,p,ν,L,σ and A such that

$$\left(\fint_{\frac12 Q_j}|Du-Dv_j|^{p'}\,dx\,dt\right)^{1/p'}\le \bar c_3\left[\lambda\omega(r_j)+\lambda^{1-a}\mu_{j-1}\right]\,. \tag{3.31}$$

Proof. We start proving there exists a constant $c\equiv c(n,m,p,\nu,\sigma,A)$ such that

$$\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{1/p'}\le c\lambda^{1-a}\mu_{j-1}\,. \tag{3.32}$$

To this aim we start defining the rescaled maps $\bar{w}_{j-1} := w_{j-1}/\lambda$ and $\bar{w}_j := w_j/\lambda$ (recall that throughtout the paper we are assuming $\lambda > 0$). Then we estimate, by mean of the second inequality in (3.30), as

$$\begin{aligned}
&\left(\fint_{Q_j} |Du - Dw_j|^{p'}\,dx\,dt\right)^{1/p'} \\
&\leq A^{p-2}\left(\fint_{Q_j} |D\bar{w}_{j-1}|^{(p-2)p'}|Du - Dw_j|^{p'}\,dx\,dt\right)^{1/p'} \\
&\leq c\left(\fint_{Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^{(p-2)p'}|Du - Dw_j|^{p'}\,dx\,dt\right)^{1/p'} \\
&\quad +c\left(\fint_{Q_j} |D\bar{w}_j|^{(p-2)p'}|Du - Dw_j|^{p'}\,dx\,dt\right)^{1/p'} =: cI_1 + cI_2\,, \qquad (3.33)
\end{aligned}$$

where the constant c depends on p, A. Hölder's inequality and Lemma 3.1 yield

$$\begin{aligned}
I_1 &\leq \lambda^{2-p}\left(\fint_{Q_j} |Dw_j - Dw_{j-1}|^{p}\,dx\,dt\right)^{(p-2)/p} \\
&\qquad\qquad \cdot\left(\fint_{Q_j} |Du - Dw_j|^{p}\,dx\,dt\right)^{1/p} \\
&\leq c\lambda\left(\frac{\mu_{j-1}}{\lambda^a}\right)^{\frac{(n+2)(p-1)}{n(p-1)+p}} \leq c\lambda^{1-a}\mu_{j-1}\,, \qquad (3.34)
\end{aligned}$$

where $c \equiv c(n, m, p, \nu, \sigma)$ and where in the last line we have used (3.29) and that

$$\frac{(n+2)(p-1)}{n(p-1)+p} \geq 1 \Longleftrightarrow p \geq 2\,. \qquad (3.35)$$

We proceed with the estimation of I_2, initially via Hölder's inequality

$$\begin{aligned}
I_2 &= \left(\fint_{Q_j} |D\bar{w}_j|^{(p-2)p'/2}|Du - Dw_j|^{p'}|D\bar{w}_j|^{(p-2)p'/2}\,dx\,dt\right)^{1/p'} \\
&\leq \left(\fint_{Q_j} |D\bar{w}_j|^{p-2}|Du - Dw_j|^{2}\,dx\,dt\right)^{1/2}\left(\fint_{Q_j} |D\bar{w}_j|^{p}\,dx\,dt\right)^{(p-2)/(2p)}. \qquad (3.36)
\end{aligned}$$

Triangle inequality together with (3.30), that in fact implies that $1/A \leq |D\bar{w}_{j-1}| \leq A$ in Q_j, and yet using (3.26), gives

$$\fint_{Q_j} |D\bar{w}_j|^{p}\,dx\,dt \leq c\left(\frac{\mu_{j-1}}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}} + c \leq c\,.$$

Here the constant c depends on n, m, p, ν, A, σ. This piece of information together with (3.36) allows us to conclude with

$$I_2 \leq c\lambda^{(2-p)/2}\left(\fint_{Q_j} (|Dw_j| + |Du|)^{p-2}|Du - Dw_j|^{2}\,dx\,dt\right)^{1/2}, \qquad (3.37)$$

again for $c \equiv c(n, m, p, \nu, A, \sigma)$. Next, Proposition 3.1 and Young's inequality yield

$$\left(\fint_{Q_j} \left|\frac{u - w_j}{r_j}\right|^{q_1}\,dx\,dt\right)^{1/q_1} \leq c\left(\fint_{Q_j} |Du - Dw_j|^{p'}\,dx\,dt\right)^{1/p'}$$

$$+c\left(\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx\right)^{1/2},$$

where now $q_1 = p'(n+2)/n$. By noticing that $q_1' = q$ when $p \geq 2$ (recall (3.24)) and that $q \leq n/2 \leq \gamma$, and using also (3.13) and (3.11), we have

$$\begin{aligned}
&\lambda^{p-2}\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx \\
&\qquad+\fint_{Q_j}(|Dw_j|+|Du|)^{p-2}|Du-Dw_j|^2\,dx\,dt \\
&\leq cr_j\fint_{Q_j}|F|\left|\frac{u-w_j}{r_j}\right|\,dx\,dt \\
&\leq cr_j\left(\fint_{Q_j}|F|^q\,dx\,dt\right)^{1/q}\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{q_1}\,dx\,dt\right)^{1/q_1} \\
&\leq c(\varepsilon)r_j^2\lambda^{2-p}\left(\fint_{Q_j}|F|^\gamma\,dx\,dt\right)^{2/\gamma} \\
&\qquad+\varepsilon\lambda^{p-2}\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{2/p'} \\
&\qquad\qquad+\varepsilon\lambda^{p-2}\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx\,.
\end{aligned}$$

Combining the content of the last two displays, choosing ε small enough and reabsorbing terms yields

$$\begin{aligned}
&\lambda^{2-p}\fint_{Q_j}(|Dw_j|+|Du|)^{p-2}|Du-Dw_j|^2\,dx\,dt \\
&\qquad\leq c\varepsilon\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{2/p'}+c(\varepsilon)\lambda^{2(1-a)}\mu_j^2\,,
\end{aligned}$$

where we have used the definition of $a(\gamma)$ given in (1.12). The last inequality and (3.37) give

$$I_2 \leq c\sqrt{\varepsilon}\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{1/p'}+c(\varepsilon)\lambda^{1-a}\mu_{j-1}\,,$$

where the constants also depend on n, m, p, ν, L, A and σ. We have also appealed to (3.21). Using the last inequality together with (3.33)-(3.34) we conclude with

$$\begin{aligned}
\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{1/p'} &\leq c\sqrt{\varepsilon}\left(\fint_{Q_j}|Du-Dw_j|^{p'}\,dx\,dt\right)^{1/p'} \\
&\quad+c(\varepsilon)\lambda^{1-a}\mu_{j-1}
\end{aligned}$$

so that (3.32) follows by choosing ε such that $c\sqrt{\varepsilon} = 1/2$ and reabsorbing terms. Next we prove that, for a constant $c \equiv c(n, m, p, \nu, L, A, \sigma)$, it holds that

$$\left(\fint_{\frac{1}{2}Q_j}|Dv_j-Dw_j|^{p'}\,dx\,dt\right)^{1/p'} \leq c\lambda\omega(r_j)+c\lambda^{1-a}\mu_{j-1} \tag{3.38}$$

and this, together with (3.32) and an elementary manipulation, allows to conclude with (3.31); we are therefore left with the proof of (3.38). Using (3.16) and the first inequality in (3.30) gives

$$\left(\fint_{\frac{1}{2}Q_j} |Dv_j - Dw_j|^p \, dx \, dt\right)^{1/p} \leq c[\omega(r_j)]^{2/p} A\lambda \,. \tag{3.39}$$

Similarly to (3.33) we write

$$\begin{aligned}
&\left(\fint_{\frac{1}{2}Q_j} |Dv_j - Dw_j|^{p'} \, dx \, dt\right)^{1/p'} \\
&\leq c\left(\fint_{\frac{1}{2}Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^{(p-2)p'} |Dv_j - Dw_j|^{p'} \, dx \, dt\right)^{1/p'} \\
&\qquad + c\left(\fint_{\frac{1}{2}Q_j} |D\bar{w}_j|^{(p-2)p'} |Dv_j - Dw_j|^{p'} \, dx \, dt\right)^{1/p'} =: cI_3 + cI_4
\end{aligned} \tag{3.40}$$

for a constant c depending on p, A. We estimate I_3 using first Hölder's and Young's inequality and eventually (3.39)

$$\begin{aligned}
I_3 &\leq c\left(\fint_{\frac{1}{2}Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^p \, dx \, dt\right)^{(p-2)/p} [\omega(r_j)]^{2/p}\lambda \\
&\leq c\omega(r_j)\lambda + c\lambda^{1-p} \fint_{\frac{1}{2}Q_j} |Dw_j - Dw_{j-1}|^p \, dx \\
&\leq c\omega(r_j)\lambda + c\lambda\left(\frac{\mu_{j-1}}{\lambda^a}\right)^{\frac{p(n+2)}{n(p-1)+p}} \leq c\omega(r_j)\lambda + c\lambda^{1-a}\mu_{j-1} \,,
\end{aligned}$$

where, in the last line, we have used (3.26), (3.29) and (3.35). To estimate I_4, observe that, since $|D\bar{w}_j| \leq A$ in $\frac{1}{2}Q_j$, then, using (3.11) and

$$\begin{aligned}
I_4 &\leq A^{(p-2)/2}\left(\fint_{\frac{1}{2}Q_j} |D\bar{w}_j|^{(p-2)p'/2} |Dv_j - Dw_j|^{p'} \, dx \, dt\right)^{1/p'} \\
&\leq c\lambda^{(2-p)/2}\left(\fint_{\frac{1}{2}Q_j} (|Dv_j| + |Dw_j|)^{p-2} |Dv_j - Dw_j|^2 \, dx \, dt\right)^{1/2} \\
&\leq c\lambda^{(2-p)/2}\left(\fint_{\frac{1}{2}Q_j} |V(Dv_j) - V(Dw_j)|^2 \, dx \, dt\right)^{1/2} \\
&\leq c\lambda^{(2-p)/2}\omega(r_j)\left(\fint_{\frac{1}{2}Q_j} |Dw_j|^p \, dx \, dt\right)^{1/2} \leq c\omega(r_j)\lambda \,.
\end{aligned}$$

Merging the estimates found for I_3 and I_4 with (3.40) finally yields (3.38) and the proof of the lemma is complete. □

3.4. **Preliminary estimates in the case $2n/(n+2) < p < 2$.** In this section we give the sub quadratic version of the lemmas presented in the previous one. Although the outcomes are similar, the nature of arguments is different.

Lemma 3.3. *Let u be a solution to the system in* (1.5) *with $2n/(n+2) < p < 2$ and let w_j, v_j be defined as in* (3.7)-(3.8)*, respectively, with $j \geq 1$. Assume that $F \in$*

$L^\gamma(\Omega_T,\mathbb{R}^m)$ for some $\gamma \geq q$, where q is defined in (3.23), and that the inequalities

$$\mu_j \leq \lambda^a \qquad \text{and} \qquad \left(\fint_{Q_j} |Du|^p \, dx \, dt \right)^{1/p} \leq H_3 \lambda \tag{3.41}$$

hold for a constant $H_3 \geq 1$. Then there exists a constant $\bar{c}_4 \equiv \bar{c}_4(n,m,p,\nu,L,\sigma,H_3)$ such that

$$\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^p \, dx \, dt \right)^{1/p} \leq \bar{c}_4 \left[\lambda\omega(r_j) + \lambda^{1-a}\mu_{j-1} \right] \tag{3.42}$$

and moreover there exists a constant $\bar{c}_1 \equiv \bar{c}_1(n,m,p,\nu,\sigma,H_3)$ such that

$$\left(\fint_{Q_j} |Du - Dw_j|^p \, dx \, dt \right)^{1/p} \leq \bar{c}_1 \lambda^{1-a} \mu_j \, . \tag{3.43}$$

Proof. We first prove (3.43). We recall that the vector field $V(\cdot)$ has been defined in (3.10). We start applying (3.12) in order to have

$$|Du - Dw_j|^p \leq c|V(Du) - V(Dw_j)|^2 + c|V(Du) - V(Dw_j)|^p |Du|^{(2-p)p/2}$$

for $c \equiv c(n,m,p)$, so that, by means of Hölder's inequality we have

$$\begin{aligned} \left(\fint_{Q_j} |Du - Dw_j|^p \, dx \, dt \right)^{1/p} &\leq c \left(\fint_{Q_j} |V(Du) - V(Dw_j)|^2 \, dx \, dt \right)^{1/p} \\ &+ c \left(\fint_{Q_j} |V(Du) - V(Dw_j)|^2 \, dx \, dt \right)^{1/2} \left(\fint_{Q_j} |Du|^p \, dx \, dt \right)^{(2-p)/2p} . \end{aligned}$$

By using (3.13) and keeping the definition of S_j given in (3.17) we have

$$S_j \leq c r_j \fint_{Q_j} |F| \left| \frac{u - w_j}{r_j} \right| dx \, dt + c\lambda^{\frac{(2-p)p}{2}} \left(r_j \fint_{Q_j} |F| \left| \frac{u - w_j}{r_j} \right| dx \, dt \right)^{p/2} ,$$

where we have also used (3.41) and $c \equiv c(n,m,p,\nu,H_3)$. Estimating as in (3.27) and (3.28) we get

$$r_j \fint_{Q_j} |F| \left| \frac{u - w_j}{r_j} \right| dx \, dt \leq c\lambda^{\frac{2-p}{n+2}} \left(\frac{r_j^{\gamma - N}}{\lambda^{2-p}} \int_{Q_j} |F|^\gamma \, dx \, dt \right)^{1/\gamma} S_j^{\frac{n+p}{p(n+2)}} ,$$

so that

$$\begin{aligned} S_j \quad &\leq \quad c\lambda^{\frac{2-p}{n+2}} \left(\frac{r_j^{\gamma - N}}{\lambda^{2-p}} \int_{Q_j} |F|^\gamma \, dx \, dt \right)^{1/\gamma} S_j^{\frac{n+p}{p(n+2)}} \\ &\quad + c\lambda^{\frac{(2-p)p}{2}} \left(r_j \fint_{Q_j} |F| \left| \frac{u - w_j}{r_j} \right| dx \, dt \right)^{p/2} . \end{aligned}$$

Applying Young's inequality with exponents

$$\left(\frac{p(n+2)}{n+p}, \frac{p(n+2)}{n(p-1)+p} \right)$$

and reabsorbing terms in a standard way we conclude with

$$\left(\fint_{Q_j} |Du - Dw_j|^p \, dx \, dt \right)^{1/p}$$

$$\leq c\lambda^{\frac{2-p}{n(p-1)+p}}\left[\left(\frac{r_j^{\gamma-N}}{\lambda^{2-p}}\int_{Q_j}|F|^\gamma\,dx\,dt\right)^{1/\gamma}\right]^{\frac{n+2}{n(p-1)+p}}$$

$$+c\lambda^{\frac{2-p}{2}}\left(r_j \fint_{Q_j}|F|\left|\frac{u-w_j}{r_j}\right|\,dx\,dt\right)^{1/2} =: I_5+I_6\,. \tag{3.44}$$

We now estimate the terms I_5 and I_6. Since

$$\frac{n+2}{n(p-1)+p}>1 \Longleftrightarrow p<2$$

using the first inequality in (3.41) we can estimate as follows:

$$I_5 = c\lambda\left(\frac{\mu_j}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}} \leq c\lambda^{1-a}\mu_j\,.$$

The estimate of I_6 is more involved; we start applying Proposition 3.1 to get, eventually using Young's inequality, that

$$\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*} \leq c\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{1/p}$$

$$+c\left(\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx\right)^{1/2}$$

and, by using (3.13) and recalling that we are assuming that $\gamma>(p^*)'$, we have

$$\lambda^{p-2}\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx$$

$$\leq cr_j\fint_{Q_j}|F|\left|\frac{u-w_j}{r_j}\right|\,dx\,dt$$

$$\leq cr_j\left(\fint_{Q_j}|F|^{(p^*)'}\,dx\,dt\right)^{1/(p^*)'}\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{1/p^*}$$

$$\leq c(\varepsilon)\lambda^{2-p}\left(r_j^\gamma\fint_{Q_j}|F|^\gamma\,dx\,dt\right)^{2/\gamma}+\varepsilon\lambda^{p-2}\left(\fint_{Q_j}\left|\frac{u-w_j}{r_j}\right|^{p^*}dx\,dt\right)^{2/p^*}$$

$$\leq c\varepsilon\lambda^{p-2}\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{2/p}$$

$$+c\varepsilon\lambda^{p-2}\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx$$

$$+c(\varepsilon)\lambda^{2-p}\left(r_j^\gamma\fint_{Q_j}|F|^\gamma\,dx\,dt\right)^{2/\gamma}\,. \tag{3.45}$$

Choosing ε small enough (and in particular smaller than $1/2$) in the previous inequality and reabsorbing terms in a standard way yields

$$\lambda^{p-2}\sup_{t_0-\lambda^{2-p}r_j^2<\tau<t_0}\fint_{B_j}\left|\frac{u-w_j}{r_j}\right|^2(x,\tau)\,dx$$

$$\leq c\varepsilon\lambda^{p-2}\left(\fint_{Q_j}|Du-Dw_j|^p\,dx\,dt\right)^{2/p}+c(\varepsilon)\lambda^{2-p}\left(r_j^\gamma\fint_{Q_j}|F|^\gamma\,dx\,dt\right)^{2/\gamma}\,.$$

The last estimate and again (3.45) give

$$
\begin{aligned}
r_j \fint_{Q_j} |F| \left| \frac{u-w_j}{r_j} \right| dx\, dt \quad \leq \quad & c\varepsilon\lambda^{p-2} \left(\fint_{Q_j} |Du - Dw_j|^p \, dx\, dt \right)^{2/p} \\
& + c(\varepsilon)\lambda^{2-p} \left(r_j^\gamma \fint_{Q_j} |F|^\gamma \, dx\, dt \right)^{2/\gamma}
\end{aligned}
$$

so that recalling the definition of I_6 in (3.44) we have

$$
\begin{aligned}
I_6 \quad & \leq \quad c\sqrt{\varepsilon} \left(\fint_{Q_j} |Du - Dw_j|^p \, dx\, dt \right)^{1/p} + c(\varepsilon)\lambda^{2-p} \left(r_j^\gamma \fint_{Q_j} |F|^\gamma \, dx\, dt \right)^{1/\gamma} \\
& = \quad c\sqrt{\varepsilon} \left(\fint_{Q_j} |Du - Dw_j|^p \, dx\, dt \right)^{1/p} + c(\varepsilon)\lambda^{1-a}\mu_j \, .
\end{aligned}
$$

Connecting the estimates found for I_5 and I_6 to the one in (3.44) and again choosing (for the second time) ε small enough in order to reabsorb terms, and finally recalling (3.21), we complete the proof of (3.43). In order to conclude the proof of the Lemma we now derive the following inequality:

$$
\left(\fint_{\frac{1}{2}Q_j} |Dv_j - Dw_j|^p \, dx\, dt \right)^{1/p} \leq c[\lambda\omega(r_j) + \lambda^{1-a}\mu_j] \tag{3.46}
$$

for a constant $c \equiv c(n, m, p, \nu, L, \sigma, H_3)$, that indeed gives (3.42) when combined with (3.43) and (3.21). Then we use (3.11) and Hölder's inequality

$$
\begin{aligned}
& \left(\fint_{\frac{1}{2}Q_j} |Dv_j - Dw_j|^p \, dx\, dt \right)^{1/p} \\
& \quad \leq c \left(\fint_{\frac{1}{2}Q_j} |V(Dv_j) - V(Dw_j)|^2 \, dx\, dt \right)^{1/2} \\
& \qquad \cdot \left(\fint_{\frac{1}{2}Q_j} (|Dv_j| + |Dw_j|)^p \, dx\, dt \right)^{(2-p)/2p} \\
& \quad \overset{(3.14),(3.15)}{\leq} c\omega(r_j) \left(\fint_{\frac{1}{2}Q_j} |Dw_j|^p \, dx\, dt \right)^{1/p} \\
& \quad \overset{(3.41),(3.43)}{\leq} c\omega(r_j)\lambda + c\lambda^{1-a}\mu_j \, ,
\end{aligned}
$$

so that (3.46) follows and the proof of the lemma can be completed taking (3.21) into account. □

3.5. **Two purely technical lemmas in the full range** $p > 2n/(n+2)$**.** The following two purely technical lemmas will be instrumental to the proof of the main potential and continuity estimates. We shall analyse the impact of certain density conditions such as

$$
\left(\fint_{Q_j} |Du|^p \, dx\, dt \right)^{1/p} \leq H_3\lambda \quad \text{and} \quad \sum_{j=0}^{\infty} \frac{\mu_j}{\lambda^a} \leq \left(\frac{1}{H_2} \right)^{\max\{1,[n(p-1)+p]/(n+2)\}} \tag{3.47}
$$

for constants $H_2, H_3 \geq 1$, and like

$$\max\left\{\left(\fint_{Q_{j-1+k}} |Du|^p \,dx\,dt\right)^{1/p}, \left(\fint_{Q_{j+k}} |Du|^p \,dx\,dt\right)^{1/p}\right\} \geq \frac{\varepsilon\lambda}{10^4} \tag{3.48}$$

for some fixed integer $k \geq 2$ and some $\varepsilon \in (0,1)$. In the following the constants $\bar{c}_1, \ldots, \bar{c}_4$ are those introduced through Lemmas 3.1-3.3, with the present choice of the constant H_3 introduced in (3.47).

Lemma 3.4. *Let u be a solution to the system in* (1.5) *with $p > 2n/(n+2)$ and let w_j, v_j be defined as in* (3.7),(3.8)*, respectively, with $j \geq 0$. Assume also that $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ for some $\gamma \geq q$ and q is defined in* (3.23)*. Assume that* (3.47) *holds together with*

$$\bar{c}_1\bar{c}_4 \leq H_2 \qquad \text{and} \qquad [\omega(r_0)]^{\min\{1,2/p\}} \leq \frac{1}{\bar{c}_0\bar{c}_4} \tag{3.49}$$

and consider a constant $A \geq 1$ such that $A \geq c_2(n,m,p,\nu,L,8^N H_3)$, where the function c_2 has been defined in (3.3) *(where we are indeed taking $c_1 = 8^N H_3$). Then*

$$\sup_{\frac{1}{2}Q_j} |Dw_j| \leq A\lambda \qquad \text{and} \qquad \sup_{\frac{1}{4}Q_j} |Dv_j| \leq A\lambda \tag{3.50}$$

hold. Moreover, we have that

$$\left(\fint_{Q_{j+l}} |Du - Dw_j|^p \,dx\,dt\right)^{1/p} \leq \frac{\bar{c}_1\lambda}{\sigma^{Nl} H_2} \tag{3.51}$$

holds for every integer $l \geq 0$.

Lemma 3.5. *Let u be a solution to the system in* (1.5) *with $p \geq 2$ and let w_j, v_j be defined as in* (3.7)-(3.8)*, respectively, with $j \geq 1$. Assume that $F \in L^\gamma(\Omega_T, \mathbb{R}^m)$ for some $\gamma \geq q$ and q is defined in* (3.23)*. Assume also that* (3.47) *holds together with*

$$\frac{10^7\bar{c}_1\bar{c}_4}{\varepsilon\sigma^{N(k+10)}} \leq H_2 \qquad \text{and} \qquad [\omega(r_0)]^{\min\{1,2/p\}} \leq \frac{\varepsilon\sigma^{N(k+10)}}{10^7\bar{c}_0\bar{c}_4 H_3}\,. \tag{3.52}$$

Then it follows:

$$\left(\fint_{Q_{j+k}} |Du - Dv_j|^p \,dx\,dt\right)^{1/p} \leq \frac{\varepsilon\sigma^N\lambda}{10^7} \tag{3.53}$$

and

$$\left(\fint_{Q_{j+k}} |Du - Dw_{j-1}|^p \,dx\,dt\right)^{1/p} \leq \frac{\varepsilon\sigma^N\lambda}{10^7}\,. \tag{3.54}$$

If in addition we also assume (3.48)*, then we have*

$$\frac{\varepsilon\lambda}{10^5} \leq \sup_{Q_j} |Dw_{j-1}| \qquad \text{and} \qquad \frac{\varepsilon\lambda}{10^5} \leq \sup_{Q_{j+1}} |Dv_j|\,. \tag{3.55}$$

Finally, the inequalities in the last line follow as well if we just assume the following inequality:

$$\left(\fint_{Q_{j+1}} |Du|^p \,dx\,dt\right)^{1/p} \geq \frac{\varepsilon\lambda}{10^4}\,. \tag{3.56}$$

Proof of Lemma 3.4. Let us first consider the case $p \geq 2$; Lemma 3.1 and the second inequality in (3.47) imply that whenever $l \geq 0$ is an integer, the following holds:

$$
\begin{aligned}
\left(\fint_{Q_{j+l}} |Du - Dw_j|^p \, dx\, dt\right)^{1/p} &\leq \sigma^{-\frac{Nl}{p}} \left(\fint_{Q_j} |Du - Dw_j|^p \, dx\, dt\right)^{1/p} \\
&\leq \bar{c}_1 \sigma^{-\frac{Nl}{p}} \lambda \left(\frac{\mu_j}{\lambda^a}\right)^{\frac{n+2}{n(p-1)+p}} \leq \frac{\bar{c}_1 \lambda}{\sigma^{Nl} H_2}\,. \qquad (3.57)
\end{aligned}
$$

Observe that we have used the fact that $[n(p-1)+p]/(n+2) \geq 1$ for $p \geq 2$. In the case $2n/(n+2) < p < 2$, we directly use Lemma 3.3 to have

$$
\left(\fint_{Q_{j+l}} |Du - Dw_j|^p \, dx\, dt\right)^{1/p} \leq \bar{c}_1 \sigma^{-Nl} \lambda^{1-a} \mu_j \leq \frac{\bar{c}_1 \lambda}{\sigma^{Nl} H_2} \qquad (3.58)
$$

so that (3.51) is proved. Using the last two inequalities for $l = 0$ together with (3.47) yields and using the first inequality in (3.49) we have

$$
\left(\fint_{Q_j} |Dw_j|^p \, dx\, dt\right)^{1/p} \leq 2H_3 \lambda \qquad (3.59)
$$

so that Theorem 3.1 and the assumed bound lower bound on A yield the first inequality in (3.50). In order to prove the second inequality in (3.50), we start by the case $p \geq 2$. Using triangle inequality, (3.51) and (3.16) we estimate as follows:

$$
\begin{aligned}
&\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^p \, dx\, dt\right)^{1/p} \\
&\leq \frac{2^{N/p} \bar{c}_1 \lambda}{H_2} + \bar{c}_0 2^{N/p} [\omega(r_0)]^{2/p} \left(\fint_{Q_j} |Dw_j|^p \, dx\, dt\right)^{1/p} \\
&\leq 2^{N+1} \left[\frac{\bar{c}_1}{H_2} + H_3\right] \lambda \\
&\leq 2^{N+2} H_3 \lambda\,,
\end{aligned}
$$

where we have used also (3.49). In the case $2n/(n+2) < p < 2$ we instead employ Lemma 3.3 and (3.47) as follows:

$$
\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^p \, dx\, dt\right)^{1/p} \leq \bar{c}_4 \left[\lambda \omega(r_0) + \lambda^{1-a} \mu_{j-1}\right] \leq \left[\frac{\bar{c}_4}{H_2} + 1\right] \lambda \leq 2\lambda\,.
$$

Using triangle inequality and (3.47) yields

$$
\left(\fint_{\frac{1}{2}Q_j} |Dv_j|^p \, dx\, dt\right)^{1/p} \leq 8^N H_3 \lambda
$$

so that the second inequality in (3.50) again follows by Theorem 3.1. □

Proof of Lemma 3.5. We first prove (3.53), therefore only using (3.47) and (3.51). Let us consider the case $p \geq 2$; triangle inequality, (3.16) and (3.57) give

$$
\left(\fint_{Q_{j+k}} |Du - Dv_j|^p \, dx\, dt\right)^{1/p} \leq \frac{\bar{c}_1 \lambda}{\sigma^{Nk} H_2} + \frac{\bar{c}_0 [\omega(r_0)]^{2/p}}{\sigma^{N(k+1)}} \left(\fint_{Q_j} |Dw_j|^p \, dx\, dt\right)^{1/p}.
$$

Then (3.52) and (3.59) yield

$$\left(\fint_{Q_{j+k}} |Du - Dv_j|^p \, dx\, dt\right)^{1/p} \leq \frac{\varepsilon\sigma^N\lambda}{10^7}\,.$$

In the case $2n/(n+2) < p < 2$ we instead employ Lemma 3.3 as follows:

$$\begin{aligned}\left(\fint_{Q_{j+k}} |Du - Dv_j|^p \, dx\, dt\right)^{1/p} &\leq \bar{c}_4\sigma^{-Nk}\left[\lambda\omega(r_0) + \lambda^{1-a}\mu_j\right]\\ &\leq \frac{\varepsilon\sigma^{2N}\lambda}{10^7} + \frac{\bar{c}_4\sigma^{-Nk}\lambda}{H_2} \leq \frac{\varepsilon\sigma^N\lambda}{10^7}\end{aligned}$$

so that (3.53) is proved. We observe that (3.57)-(3.58) still hold with j replaced by $j-1$ as not it is $j \geq 1$; eventually taking $l = k+1$ we get

$$\left(\fint_{Q_{j+k}} |Du - Dw_{j-1}|^p \, dx\, dt\right)^{1/p} \leq \frac{\bar{c}_1\lambda^{1-a}\mu_{j-1}}{\sigma^{N(k+1)}} \leq \frac{\bar{c}_1\lambda}{\sigma^{N(k+1)}H_2}$$

so that (3.54) follows by using the first inequality in (3.52). We finally proceed to the proof of (3.55), thereby also assuming (3.48);. Moreover, we assume that the maximum in (3.48) is attained by the second integral; the case it is attained by the first can be treated in the same way once we notice that since we are assuming $k \geq 2$, then $Q_{j-1+k} \subset Q_{j+1}$. Now, triangle inequality and (3.54) imply

$$\begin{aligned}\frac{\varepsilon\lambda}{10^4} &\leq \left(\fint_{Q_{j+k}} |Du|^p \, dx\, dt\right)^{1/p}\\ &\leq \frac{\varepsilon\lambda}{10^7} + \left(\fint_{Q_{j+k}} |Dw_{j-1}|^p \, dx\, dt\right)^{1/p} \leq \frac{\varepsilon\lambda}{10^7} + \sup_{Q_j} |Dw_{j-1}|\end{aligned}$$

from which the first inequality in (3.55) follows. Similarly, by (3.53) we have

$$\begin{aligned}\frac{\varepsilon\lambda}{10^4} &\leq \left(\fint_{Q_{j+k}} |Du|^p \, dx\, dt\right)^{1/p}\\ &\leq \frac{\varepsilon\lambda}{10^7} + \left(\fint_{Q_{j+k}} |Dv_j|^p \, dx\, dt\right)^{1/p} \leq \frac{\varepsilon\lambda}{10^7} + \sup_{Q_{j+1}} |Dv_j|\end{aligned}$$

and also the second inequality in (3.55) follows. Finally, the validity of (3.55) under assumption (3.56) is obviously implicitly contained above. □

4. Proof of Theorem 1.5

The proof goes in several steps and involves an induction argument. In the following we select a Lebesgue point of the spatial gradient $(x_0, t_0) \in \Omega_T$, i.e.,

$$\lim_{\varrho\to 0} (Du)_{Q_\varrho(x_0,t_0)} = Du(x_0, t_0)\,. \tag{4.1}$$

Note that this definition of Lebesgue point is slightly different from the one usually reported in the literature, in the sense that the cylinders used here are not centred on the point (x_0, t_0), that is in fact the vertex. Moreover, we are going to use a similar definition with the intrinsic cylinders. To clarify these issues, let us define

$$\mathcal{L}_\lambda := \left\{(x,t) \in \Omega_T \ :\ \lim_{\varrho\to 0} (Du)_{Q_\varrho^\lambda(x,t)} = Du(x,t)\right\}$$

for $\lambda > 0$. Using [37, Chapter 1, Page 8] we now have $\mathcal{L}_\lambda = \mathcal{L}_1 =: \mathcal{L}$ for all $0 < \lambda < \infty$ and therefore in (4.1) we can replace $Q_\varrho(x_0, t_0)$ by $Q_\varrho^\lambda(x_0, t_0)$ in (4.1).

Needless to say, almost every point in Ω_T, with respect to the Lebesgue measure in $\mathbb{R}^{n+1}$, satisfies such a property, and for this see again [37, Chapter 1, Page 8]. In the following s is defined by

$$s := \begin{cases} p' & \text{if} \quad p \geq 2 \\ p & \text{if} \quad 1 < p < 2\,. \end{cases} \tag{4.2}$$

Accordingly, we shall use the excess functional $E_s(\cdot)$ introduced in (2.1). With $\sigma \in (0, 1/4)$ to be determined in a few lines, we adopt the set-up of Section 3.2 with the shrinking intrinsic cylinders $\{Q_i\}$ defined in (3.6). We start recalling that the quantities μ_i have already been defined in (3.20) whenever $i \geq 0$. Moreover, keeping in mind the definition of excess functional given in (2.1), we define

$$E_i := E_s(Du, Q_i) = \left(\fint_{Q_i} |Du - (Du)_{Q_i}|^s \, dx\, dt\right)^{1/s}, \quad a_i := |(Du)_{Q_i}| \tag{4.3}$$

again for $i \geq 0$. Finally, for indexes $i \geq 1$ we define

$$M_i := \max\left\{ \left(\fint_{Q_{i-1}} |Du|^p \, dx\, dt\right)^{1/p}, \left(\fint_{Q_i} |Du|^p \, dx\, dt\right)^{1/p} \right\} \tag{4.4}$$

and, again for $i \geq 1$

$$C_i := M_i + 2\sigma^{-N} E_i\,.$$

Note that C_i contains two different Lebesgue norms: the term M_i contains the L^p-norm of the gradient, while the remaining one contains an L^s-based functional. Indeed, by (4.2) we have $s < p$ for $p > 2$.

Step 1: Choice of the constants. Now, to proceed with the proof, with no loss of generality we shall verify (1.15) with $Q_{2r}^\lambda(x_0, t_0)$ instead of $Q_r^\lambda(x_0, t_0)$. We start choosing numbers $H_1, H_2 \geq 1$ appearing in the following lower bound for λ:

$$\begin{aligned} \lambda \quad > \quad & H_1 \left(\fint_{Q_{2r}^\lambda(x_0,t_0)} |Du|^p \, dx\, dt\right)^{1/p} \\ & + H_2^{\max\{1,[n(p-1)+p]/(n+2)\}/a} \left[\mathbf{P}^F_{\lambda,\gamma}(x_0, t_0; 2r)\right]^{1/a}, \end{aligned}$$

where $r \in (0, R_0/2)$ and the radius R_0 is going to be determined in a few lines, according the constant dependence stated in Theorem 1.5, which will be then proved with the choice $c := 2\max\{H_1, H_2^{\max\{1,[n(p-1)+p]/(n+2)\}/a}\}$. In the rest of the proof, when using Lemmas 3.4-3.5 we shall always do that with the choice $H_3 = 1$, and this, in the rest of the proof, fixes the value of the constant H_3. We immediately notice that, for every choice of $\sigma \in (0, 1/4)$, by (3.19) we also have

$$\begin{aligned} \lambda \quad > \quad & H_1 \left(\fint_{Q_{2r}^\lambda(x_0,t_0)} |Du|^p \, dx\, dt\right)^{1/p} \\ & + H_2^{\max\{1,[n(p-1)+p]/(n+2)\}/a} \sigma^{N/(a\gamma)} \left[\sum_{j=0}^{\infty} \mu_j\right]^{1/a}. \end{aligned} \tag{4.5}$$

By looking at Theorem 3.1, we take $c_1 = 8^N$ and this determines the corresponding constant c_2 introduced in (3.3) as a function depending only on n, m, p, ν, L. Next, keeping Theorems 3.2 and 3.3 in mind, we set

$$A := 10^{5N} \max\{c_2, 200\}, \quad B := 10^5, \quad \delta := 10^{-7}, \quad \bar{\varepsilon} := 4^{-(N+4)}\,. \tag{4.6}$$

With the choices in (4.6), we determine $\sigma_1 \equiv \sigma_1(n,m,p,\nu,L,\omega(\cdot))$ and $\sigma_2 \equiv \sigma_2(n,m,p,\nu,L)$ from Theorems 3.2 and 3.3, respectively. Set

$$\sigma := \min\{\sigma_1, \sigma_2, 1000^{-p/N}p^{-1/N}A^{-(p-1)/N}, 16^{-Np}\} \in (0,1/4)\,. \tag{4.7}$$

The choices made above guarantee that all fixed parameters $A, B, \bar{\varepsilon}, \delta, \sigma$ depend only on $n,m,p,\nu,L,\omega(\cdot)$. Furthermore, let c_3 and α be as in Theorem 3.4, corresponding to the choice of A in (4.6). In this way they both depend only on n,m,p,ν,L. Next, let k be the smallest integer satisfying

$$c_3\sigma^{k\alpha} \leq \frac{\sigma^N}{10^6} \quad \text{and} \quad k \geq 4 \tag{4.8}$$

so that $k \equiv k(n,m,p,\nu,L,\omega(\cdot))$. Observe that now we have fixed the quantity σ and therefore the setting described in Section 3.2 becomes effective: the cylinders $\{Q_j\}$ in (3.6) and the maps w_j, v_j defined in (3.7),(3.8), respectively, are determined. In particular, all the constants $\bar{c}_1, \ldots, \bar{c}_4$ defined through Lemmas 3.1-3.3 and after Section 3.2 are now fixed as quantities depending only on $n,m,p,\nu,L,\omega(\cdot)$. In order to get a constant larger than all the previous ones (which, we remind the reader, are all larger than one) we define

$$c_f = \prod_{i=0}^{4} \bar{c}_i\,. \tag{4.9}$$

Now we set

$$H_1 := \frac{100^N 10^6}{\sigma^{4N}}\,, \qquad H_2 := \frac{2^N A c_f}{\sigma^{N(k+10)}} \tag{4.10}$$

and as such they ultimately depend on $n,m,p,\nu,L,\omega(\cdot)$, as required in the statement of Theorem 1.5. We then pass to the choice of R_0; we first take $R_2 \equiv R_2(n,m,p,\nu,L,\omega(\cdot))$ to be the largest positive number such that

$$c_f\sigma^{-N(k+10)}[\omega(R_2)]^{\min\{1,2/p\}} + c_f\sigma^{-(2N+1)}\int_0^{R_2} \omega(\varrho)\,\frac{d\varrho}{\varrho} \leq \frac{1}{2^N 10^7} \tag{4.11}$$

is satisfied; needless to say we are using the finiteness condition in (1.7). Finally, with $R_1 \equiv R_1(n,m,p,\nu,L,\omega(\cdot))$ being introduced in Theorem 3.1, we let

$$R_0 := \min\{R_1, R_2\}/4\,. \tag{4.12}$$

In the following we shall always take $2r \leq R_0$ so that, with $r \equiv r_0 \in (0, R_2/8]$, then

$$\begin{aligned} \int_0^{R_2} \omega(\varrho)\,\frac{d\varrho}{\varrho} &= \sum_{i=0}^{\infty} \int_{r_{i+1}}^{r_i} \omega(\varrho)\,\frac{d\varrho}{\varrho} + \int_{r_0}^{R_2} \omega(\varrho)\,\frac{d\varrho}{\varrho} \\ &\geq \log\left(\frac{1}{\sigma}\right)\sum_{i=0}^{\infty} \omega(r_{i+1}) + \log 4\omega(r_0) \geq \sigma \sum_{i=0}^{\infty} \omega(r_i)\,. \end{aligned} \tag{4.13}$$

Matching (4.11) and (4.13) yields

$$c_f\sigma^{-N(k+4)}[\omega(r_0)]^{\min\{1,2/p\}} + c_f\sum_{i=0}^{\infty} \omega(r_i) \leq \frac{\sigma^{2N}}{2^N 10^6}\,. \tag{4.14}$$

Before going on we would like to emphasise that with the choice of the various constants made above the assumptions (3.49) and (3.52) (with $\varepsilon = 1$) from Lemmas 3.4 and 3.5, respectively, are satisfied.

Step 2: Induction. By (4.10) and (4.5) it follows that

$$C_1 \leq 5\sigma^{-2N}\left(\fint_{Q_0} |Du|^p\,dx\,dt\right)^{1/p} \leq 5\sigma^{-3N}\left(\fint_{2Q_0} |Du|^p\,dx\,dt\right)^{1/p} \leq \frac{\lambda}{10^3}\,.$$

Furthermore, without loss of generality, we may assume that there is an exit time index $i_e \geq 1$ such that $C_i > \lambda/10^3$ whenever $i > i_e$ and $C_{i_e} \leq \lambda/10^3$. Indeed, if such an index does not exist, we have a subsequence $(i_j)_j$ of indexes such that $C_{i_j} \leq \lambda/10^3$ as $j \to \infty$. But, as we assume that (x_0, t_0) is a Lebesgue point of Du, then also

$$|Du(x_0,t_0)| = \lim_{j\to\infty} a_{i_j} \leq \limsup_{j\to\infty} C_{i_j} \leq \frac{\lambda}{10^3}$$

would hold and the proof would be complete. Thus, from now on, we shall work under the assumptions

$$C_i > \frac{\lambda}{10^3} \quad \text{for} \quad i > i_e\,, \quad \text{and} \quad C_{i_e} \leq \frac{\lambda}{10^3}\,. \tag{4.15}$$

With c_f being the constant in (4.9), we now introduce the following conditions:

$$j \geq i_e, \quad \mathrm{F}(j) \quad : \quad M_j \leq \lambda$$

$$j > i_e, \quad \mathrm{S}(j) \quad : \quad \sum_{i=i_e+1}^{j} E_i \leq \frac{1}{2}\sum_{i=i_e}^{j-1} E_i + \frac{2c_f\lambda}{\sigma^N}\sum_{i=i_e}^{j-1}\omega(r_i) + \frac{2c_f\lambda^{1-a}}{\sigma^N}\sum_{i=i_e-1}^{j-2}\mu_i$$

that will serve in the next iteration scheme:

$$\mathrm{F}(i_e) \implies \mathrm{S}(i_e+1) \tag{4.16}$$

$$\mathrm{F}(j) \text{ and } \mathrm{S}(j) \implies \mathrm{S}(j+1) \qquad \forall\, j > i_e \tag{4.17}$$

$$\mathrm{F}(j) \text{ and } \mathrm{S}(j+1) \implies \mathrm{F}(j+1) \qquad \forall\, j \geq i_e\,. \tag{4.18}$$

The validity of $\mathrm{F}(j)$ and $\mathrm{S}(j)$ for all the corresponding indexes j will imply the one of Theorem 1.5 as

$$|Du(x_0,t_0)| = \lim_{j\to\infty} a_j \leq \lim_{j\to\infty} M_j \leq \lambda.$$

Let us anyway remark that $\mathrm{F}(i_e)$, which is actually the starting point of the induction scheme, is automatically satisfied since $M_{i_e} \leq C_{i_e} \leq \lambda/10^3$.

Step 4: Consequences of F(j). Here we derive a few inequalities, which are a consequence of $\mathrm{F}(j)$ for $j \geq i_e$. The first ones we record are the following:

$$\sup_{\frac{1}{2}Q_j} |Dw_j|,\ \sup_{\frac{1}{2}Q_{j-1}} |Dw_{j-1}| \leq A\lambda \quad \text{and} \quad \sup_{\frac{1}{4}Q_j} |Dv_j| \leq A\lambda\,. \tag{4.19}$$

These follow directly from Lemma 3.4 as the choices made for A, H_2 and R_0 on one hand (and $H_3 = 1$), and the induction assumption $\mathrm{F}(j)$ on the other hand, allow to verify the assumptions (3.47) and (3.49) for indexes j and $j-1$. We proceed by deriving a decay estimate for the quantity E_{j+1}. We are going to apply Lemma 3.5 with the choice $\varepsilon = H_3 = 1$ and with the integer k being the one defined in (4.8). By (4.19) and the choice of δ in (4.6), Theorem 3.2 applied to w_{j-1} gives

$$\operatorname*{osc}_{Q_j} Dw_{j-1} \leq \frac{\lambda}{10^7}\,. \tag{4.20}$$

In a similar way (4.19) allows to apply Theorem 3.4 that, keeping (4.8) in mind, allows to conclude

$$\operatorname*{osc}_{Q_{j+k}} Dv_j \leq c_3\sigma^{k\alpha}\lambda \leq \frac{\sigma^N\lambda}{10^6}\,.$$

Using this last inequality together with (2.2) and (3.53) (with $\varepsilon = 1$) gives

$$
\begin{aligned}
2\sigma^{-N}E_{j+k} &\leq 4\sigma^{-N}E_s(Dv_j, Q_{j+k}) + 4\sigma^{-N}\left(\fint_{Q_{j+k}} |Du - Dv_j|^s\,dx\,dt\right)^{1/s} \\
&\leq 4\sigma^{-N}\operatorname*{osc}_{Q_{j+k}} Dv_j + 4\sigma^{-N}\left(\fint_{Q_{j+k}} |Du - Dv_j|^p\,dx\,dt\right)^{1/p} \leq \frac{\lambda}{10^5}\,.
\end{aligned}
$$

But, as $C_{j+k} > \lambda/10^3$ for $j \geq i_e$, this last fact and the inequality in the previous display rephrase as

$$
\frac{\lambda}{10^3} \leq C_{j+k} = M_{j+k} + 2\sigma^{-N}E_{j+k} \leq M_{j+k} + \frac{\lambda}{10^5}\,.
$$

The inequality in the above display and the definition of M_{j+k} given in (4.4) at this point imply

$$
\max\left\{\left(\fint_{Q_{j-1+k}} |Du|^p\,dx\,dt\right)^{1/p}, \left(\fint_{Q_{j+k}} |Du|^p\,dx\,dt\right)^{1/p}\right\} \geq \frac{\varepsilon\lambda}{10^4} \geq \frac{\lambda}{2000}\,.
$$

In turn this means that (3.48) is satisfied with $\varepsilon = 1$ and we can use the remaining part of Lemma 3.5, that is (3.55). The first lower bound in (3.55) - with $\varepsilon = 1$ - now yields the existence of a point $(\tilde{x}, \tilde{t}) \in Q_j$ such that $2|Dw_{j-1}(\tilde{x}, \tilde{t})| \geq \lambda/10^5$ and therefore the oscillation control in (4.20) gives

$$
\frac{\lambda}{A} \leq \inf_{Q_j} |Dw_{j-1}|\,. \tag{4.21}
$$

By this last inequality and (4.19) we are able to apply Lemma 3.2 for the case $p \geq 2$, while in the case $2n/(n+2)$ we can anyway apply Lemma 3.3. Notice that the application of Lemma 3.2 is allowed since we are assuming $\gamma \geq n/2$ and in turn $n/2 \geq q$ for $p \geq 2$. In any case we come up with

$$
\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^s\,dx\,dt\right)^{1/s} \leq c_f[\lambda\omega(r_j) + \lambda^{1-a}\mu_{j-1}]\,. \tag{4.22}
$$

Next, by the last inequality in (4.19) and again the second one in (3.55), conditions (3.4) for Theorem 3.3 are satisfied and therefore an application of (3.5) with $\bar{\varepsilon} = 4^{-(N+4)}$ yields

$$
E_s(Dv_j, Q_{j+1}) \leq 4^{-(N+4)}E_s(Dv_j, (1/4)Q_j)\,. \tag{4.23}
$$

Notice that we have used the definition of (4.7) that gives $Q_{j+1} = \sigma Q_j \subset \sigma_2 Q_j$. Now, by (4.22), recalling the definitions in (4.3) and using (2.2) repeatedly, both

$$
\begin{aligned}
E_s(Dv_j, (1/4)Q_j) &\leq 2^{2N+1}E_j + 2^{N+1}\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^s\,dx\,dt\right)^{1/s} \\
&\leq 2^{2N+1}E_j + 2^{N+1}c_f[\lambda\omega(r_j) + \lambda^{1-a}\mu_{j-1}]
\end{aligned}
$$

and

$$
\begin{aligned}
2E_s(Dv_j, Q_{j+1}) &\geq E_{j+1} - 2\sigma^{-N}\left(\fint_{\frac{1}{2}Q_j} |Du - Dv_j|^s\,dx\,dt\right)^{1/s} \\
&\geq E_{j+1} - 2\sigma^{-N}c_f[\lambda\omega(r_j) + \lambda^{1-a}\mu_{j-1}]
\end{aligned}
$$

hold. Combining the inequalities in the last three displays and summarising the content of this Step finally gives the following implication:

$$\mathrm{F}(j) \Longrightarrow E_{j+1} \leq \frac{1}{2}E_j + \frac{2c_f\lambda}{\sigma^N}\omega(r_j) + \frac{2c_f\lambda^{1-a}}{\sigma^N}\mu_{j-1}\,. \tag{4.24}$$

Step 5: Proof of (4.16)-(4.18). The verification of (4.16) is a straightforward consequence of (4.24) as we have already seen that $\mathrm{F}(i_e)$ holds. As for (4.17), that is the validity of $\mathrm{S}(j+1)$, this simply follows by summing the inequality appearing in display (4.24) to the one yielded by the definition of $\mathrm{S}(j)$. It remains to prove (4.18); let us assume that $\mathrm{F}(j)$ and $\mathrm{S}(j+1)$ hold for some $j \geq i_e$. An easy manipulation of $\mathrm{S}(j+1)$ gives

$$\sum_{i=i_e}^{j+1} E_i \leq 2E_{i_e} + \frac{4c_f\lambda}{\sigma^N}\sum_{i=0}^{\infty}\omega(r_i) + \frac{4c_f\lambda^{1-a}}{\sigma^N}\sum_{i=0}^{\infty}\mu_i$$

and the right hand side can be estimated recalling that $C_{i_e} \leq \lambda/10^3$ and using (4.5), (4.14), to obtain

$$\sum_{i=i_e}^{j+1} E_i \leq \frac{\sigma^N\lambda}{500}\,. \tag{4.25}$$

On the other hand, Jensen's inequality implies

$$a_{i+1} - a_i \leq \left(\fint_{Q_{i+1}} |Du - (Du)_{Q_i}|^s\,dx\,dt\right)^{1/s} \leq \sigma^{-N}E_i \tag{4.26}$$

so that, using the inequalities in the last two displays and telescoping yields

$$a_{l+1} \leq a_{i_e} + \sigma^{-N}\sum_{i=i_e}^{l} E_i \leq \frac{\lambda}{10^3} + \frac{\lambda}{500} \leq \frac{\lambda}{200} \tag{4.27}$$

for all $l \in \{i_e, \ldots, j\}$. Here we also used that, thanks to (4.15), we have

$$a_{i_e} \leq \left(\fint_{Q_{i_e}} |Du|^p\,dx\,dt\right)^{1/p} \leq M_{i_e} \leq C_{i_e} \leq \frac{\lambda}{10^3}\,.$$

We now proceed to the proof of (4.18); the easiest case is when $2n/(n+2) < p < 2$, when $s = p$ and therefore

$$\left(\fint_{Q_{j+1}} |Du|^p\,dx\,dt\right)^{1/p} \leq E_{j+1} + a_{j+1} \leq \frac{\sigma^N\lambda}{500} + \frac{\lambda}{200} \leq \lambda$$

so that $M_{j+1} \leq \lambda$ (also recall the definition of M_j). The case $p \geq 2$ requires slightly more care, and we start observing that, since $\mathrm{F}(j)$ is in force, as in Step 4 we can use Lemma 3.4 and in particular (3.51). Therefore, by triangle inequality and (3.51), and yet the choice of H_2 in (4.10), we gain

$$\left(\fint_{Q_{j+1}} |Du|^p\,dx\,dt\right)^{1/p} \leq \frac{\sigma^N\lambda}{2^N 10^5} + \left(\fint_{Q_{j+1}} |Dw_j|^p\,dx\,dt\right)^{1/p} \tag{4.28}$$

and in turn

$$\begin{aligned}\fint_{Q_{j+1}} |Dw_j|^p\,dx\,dt &\leq a_{j+1}\fint_{Q_{j+1}} |Dw_j|^{p-1}\,dx\,dt \\ &\quad + \fint_{Q_{j+1}} |Dw_j|^{p-1}\left(|Du - (Du)_{Q_{j+1}}| + |Du - Dw_j|\right)\,dx\,dt\,.\end{aligned}$$

Using Young's inequality we get

$$a_{j+1} \fint_{Q_{j+1}} |Dw_j|^{p-1}\, dx\, dt \le \frac{p-1}{p} \fint_{Q_{j+1}} |Dw_j|^p\, dx\, dt + \frac{a_{j+1}^p}{p}\,.$$

Matching the inequalities in the last two displays, reabsorbing terms, and using (4.19), yields

$$\begin{aligned}\fint_{Q_{j+1}} |Dw_j|^p\, dx\, dt &\le a_{j+1}^p \\ &\quad + p(A\lambda)^{p-1} \fint_{Q_{j+1}} \left(|Du-(Du)_{Q_{j+1}}| + |Du - Dw_j|\right)\, dx\, dt \\ &\le p\left(\frac{\lambda}{200}\right)^p + p(A\lambda)^{p-1}E_{j+1} + p(A\lambda)^{p-1}\left(\fint_{Q_{j+1}} |Du-Dw_j|^p\, dx\, dt\right)^{1/p},\end{aligned}$$

so that, using (4.25), (3.51) and the choice of σ in (4.7), yields

$$\fint_{Q_{j+1}} |Dw_j|^p\, dx\, dt \le \left(\frac{\lambda}{8}\right)^p.$$

Combining the last inequality with (4.28) allows to conclude with the proof of the inequality $M_{j+1} \le \lambda$. This ends the verification of the induction step and thereby concludes the whole proof. We just confine ourselves to remark that the radius R_0 appearing in the statement of Theorem 1.5 is the one that has been determined in (4.12); now, notice that in the case no coefficients are present, i.e., the function $b(\cdot)$ in (1.5) is constant, then we automatically have $\omega(\cdot) = 0$. At this stage any choice of R_1 and R_2 is sufficient in Theorem 3.1 and (4.11), respectively, and therefore we can take $R_0 = \infty$ in Theorem 1.5.

5. Proof of Theorems 1.3-1.4

Clearly, it is sufficient to prove the estimates under the additional condition that $2r < R_0$, where R_0 is the radius appearing in the statement of Theorem 1.5. The general case follows in a standard way enlarging the constant appearing in estimates of a factor that has a power like dependence on $\max\{1, R_0^{-n}\}$; note indeed that R_0 in turn depends only on n, m, p, ν, L and $\omega(\cdot)$. With $\tilde c \equiv \tilde c(n,m,p,\nu,L,\omega(\cdot)) > 1$ being the constant appearing in Theorem 1.5, we start setting

$$h(\lambda) \equiv h_r(\lambda) := \lambda - \tilde c \left(\fint_{Q_r^\lambda} (|Du|+1)^p\, dx\, dt\right)^{1/p} - \tilde c\left[\mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r)\right]^{1/a} \tag{5.1}$$

whenever $Q_r^\lambda \equiv Q_r^\lambda(x_0,t_0) \subset \Omega_T$. The function $h(\cdot)$ is continuous at any point, is finite and moreover $h(1) < 0$ as $\tilde c > 1$ and we are assuming that $Q_r \equiv Q_r^1 \subset \Omega_T$. We now start with the proof of Theorem 1.3 and therefore we consider the case $p \ge 2$. Since in this case $Q_r^\lambda \subset Q_r$ for all $\lambda \ge 1$, then we have

$$h(\lambda) \ge \lambda - \tilde c \lambda^{\frac{p-2}{p}} \left(\fint_{Q_r} (|Du|+1)^p\, dx\, dt\right)^{1/p} - \tilde c\left[\mathbf{P}^F_{\gamma}(x_0,t_0;r)\right]^{1/a}$$

so that $h(\lambda) \to \infty$ when $\lambda \to \infty$. It follows that there exists a finite number $\lambda > 1$ such that $h(\lambda) = 0$, that is λ satisfies (1.15). Therefore we can apply Theorem 1.5 that together with Young's inequality gives

$$\lambda + |Du(x_0,t_0)| \le 2c\lambda^{\frac{p-2}{p}} \left(\fint_{Q_r} (|Du|+1)^p\, dx\, dt\right)^{1/p} + 2\tilde c\left[\mathbf{P}^F_{\gamma}(x_0,t_0;r)\right]^{1/a}$$

$$\leq \frac{\lambda}{2} + c \left(\fint_{Q_r} (|Du|+1)^p \, dx\, dt \right)^{1/2} + c \left[\mathbf{P}^F_\gamma (x_0,t_0;r) \right]^{1/a}$$

with $c \equiv c(n,m,p,\nu,L,\omega(\cdot))$, and Theorem 1.3 is proved. In the case $2n/(n+2) < p < 2$ we do consider cylinders of the type

$$Q^\lambda_{r_\lambda}(x_0,t_0) := B(x_0,\lambda^{(p-2)/2}r) \times (t_0 - r^2, t_0)\,, \quad \text{where} \quad r_\lambda = \lambda^{(p-2)/2} r \tag{5.2}$$

and accordingly, we consider $h(\lambda) \equiv h_{r_\lambda}(\lambda)$ in (5.1). For this type of cylinders notice that since now $p < 2$, we have $Q^\lambda_{r_\lambda}(x_0,t_0) \subset Q_r(x_0,t_0)$ for $\lambda \geq 1$. Therefore, by changing variables $\varrho_\lambda := \lambda^{(p-2)/2}\varrho$ we then have, again for $\lambda \geq 1$, that

$$\begin{aligned} \mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r_\lambda) &= \int_0^r \left(\frac{|F|^\gamma (Q^\lambda_{\varrho_\lambda}(x_0,t_0))}{\varrho_\lambda^N} \right)^{1/\gamma} d\varrho_\lambda \\ &\leq \lambda^{\frac{2-p}{2}\left(\frac{N}{\gamma}-1\right)} \mathbf{P}^F_\gamma (x_0,t_0;r)\,. \end{aligned} \tag{5.3}$$

Therefore the inequality

$$\begin{aligned} h(\lambda) \;\geq\; & \lambda - \tilde{c}\lambda^{\frac{(2-p)n}{2p}} \left(\fint_{Q_r} (|Du|+1)^p \, dx\, dt \right)^{1/p} \\ & - \tilde{c}\lambda^{\frac{2-p}{2a}\left(\frac{N}{\gamma}-1\right)} \left[\mathbf{P}^F_\gamma (x_0,t_0;r) \right]^{1/a} \end{aligned}$$

holds whenever $\lambda \geq 1$. Recall now the lower bound in (1.5), that is

$$p > \frac{2n}{n+2} \Longleftrightarrow \frac{(2-p)n}{2p} < 1\,. \tag{5.4}$$

Moreover, remembering the definition of $a(\gamma)$ given in (1.12), an elementary but lengthy computation shows that the main conditions assumed in Theorem 1.4, that is $p > 2n/(n+2)$ and $\gamma \geq (p^*)'$, imply

$$\frac{2-p}{2a(\gamma)} \left(\frac{N}{\gamma} - 1 \right) < \frac{2}{a((p^*)')(n+2)} \left(\frac{N}{(p^*)'} - 1 \right) = 1\,. \tag{5.5}$$

This, together with (5.4), allows to conclude that $h(\lambda) \to \infty$ when $\lambda \to \infty$, and therefore we find $\lambda > 1$ such that $h(\lambda) = 0$. Applying Theorem 1.5 on the cylinder $Q^\lambda_{r_\lambda}(x_0,t_0) \subset Q_r$ for such λ yields

$$\begin{aligned} \lambda + |Du(x_0,t_0)| \;\leq\; & 2c\lambda^{\frac{(p-2)n}{2p}} \left(\fint_{Q_r} (|Du|+1)^p \, dx\, dt \right)^{1/p} \\ & + 2\tilde{c}\lambda^{\frac{2-p}{2a}\left(\frac{N}{\gamma}-1\right)} \left[\mathbf{P}^F_\gamma (x_0,t_0;r) \right]^{1/a}\,. \end{aligned}$$

Thanks to (5.4) we can apply Young's inequality with conjugate exponents

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

to deal with the intermediate term in the second-last display and, in the case $\gamma < N$, with exponents

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

to deal with the last one. Summarising, we get

$$\begin{aligned} \lambda + |Du(x_0,t_0)| \;\leq\; & \frac{\lambda}{2} + c \left(\frac{1}{|Q_r|} \int_{Q^\lambda_r} (|Du|+1)^p \, dx\, dt \right)^{2/[p(n+2)-2n]} \\ & + c \left[\mathbf{P}^F_\gamma (x_0,t_0;r) \right]^{\max\{\frac{2\gamma}{\gamma p + n(p-2)}, \frac{1}{a}\}}\,, \end{aligned}$$

from which Theorem 1.4 follows noticing that

$$\max\left\{\frac{2\gamma}{\gamma p+n(p-2)},\frac{1}{a}\right\}=\frac{1}{a}\Longleftrightarrow \gamma\geq N\,.$$

Remark 5.1. The range considered for the proof of Theorem 1.4, that is

$$\gamma \geq (p^*)' \qquad \text{and} \qquad p>\frac{2n}{n+2}\,,$$

is used to get (5.5), that eventually allows to derive the final potential a priori estimate. The balance between the two conditions in (5.5) precisely reflects in the equality in the right hand side of (5.5). As a matter of fact, assuming a lower integrability on F, that is considering smaller values of γ, forces to take larger values of p. In particular, in the scalar case $m=1$ one considers the range

$$\gamma = 1 \qquad \text{and} \qquad p>2-\frac{1}{n+1}\,,$$

which is in fact typical of measure data problems [3].

6. Proof of Theorems 1.1-1.2 and 1.6

We first prove Theorem 1.6 and then we infer Theorems 1.1-1.2 as corollaries.

Proof of Theorem 1.6. Let us observe that by Theorems 1.3 and 1.4, and eventually passing to open subsets compactly contained in Ω_T, we can assume w.l.o.g. that the gradient is globally bounded, and therefore for the rest of the proof we let, with $H_3\geq 1$

$$\lambda:=\frac{\|Du\|_{L^\infty(\Omega_T)}+1}{H_3}<\infty\,. \tag{6.1}$$

The main objective of the proof consists of showing that

$$\lim_{\varrho\to 0}(Du)_{Q_\varrho^\lambda(x_0,t_0)} \tag{6.2}$$

exists locally uniformly in Ω_T, thereby defining a continuous precise representative of Du. To this aim, we consider a cylinder $Q_0\Subset\Omega_T$ and prove that for every $\varepsilon>0$ there exists a radius $r_\varepsilon\leq d_{\rm par}(Q_0,\partial\Omega_T)/100$, depending only on $n,m,p,\nu,L,\omega(\cdot),\varepsilon,\lambda$, such that

$$|(Du)_{Q_\varrho^\lambda(x_0,t_0)}-(Du)_{Q_\rho^\lambda(x_0,t_0)}|\leq\lambda\varepsilon \qquad \text{holds for every } \varrho,\rho\in(0,r_\varepsilon] \tag{6.3}$$

whenever $(x_0,t_0)\in Q_0$. Before going on, we yet consider a few preliminary facts; starting by a cylinder $Q_r(x_0,t_0)\subset\Omega_T$ with vertex in Q_0, it is not guaranteed that the cylinders of the type $Q_r^\lambda(x_0,t_0)$ are still contained in Ω_T when $p<2$ (recall that $\lambda\geq 1$ here). In this case we shall then consider cylinders of the type $Q_\varrho^\lambda(x_0,t_0)$ with $\varrho\leq r_\lambda=\lambda^{(p-2)/2}r$; compare with the definition in (5.2). Ultimately, it will be sufficient to consider the radius r_ε with the initial restriction

$$r_\varepsilon\leq r_m:=\min\{d_{\rm par}(Q_0,\partial\Omega_T)/100,\lambda^{(p-2)/2}d_{\rm par}(Q_0,\partial\Omega_T)/100\}\,.$$

As for the potentials, let us observe that in any case we still have that

$$\lim_{\varrho\to 0}\mathbf{P}^F_{\gamma,\lambda}(x,t;\varrho)=0 \qquad \text{uniformly with respect to } (x,t)\,. \tag{6.4}$$

Indeed, when $p\geq 2$ this follows by the trivial inequality $\mathbf{P}^F_{\gamma,\lambda}\leq\mathbf{P}^F_{\gamma}$ while in the other case it is sufficient to use (5.3). The rest proof goes now in three steps.

Remark 6.1. The reason for introducing the constant H_3 in (6.1) is instrumental and in this proof we can take $H_3 = 1$. The role of the constant H_3 will become apparent in the proof of Theorem 1.7 in the last section of this paper, where the proof of Theorem 1.6 will be revisited and where a larger value of H_3 will be considered.

Step 1: A preliminary smallness condition. As remarked a few lines above we shall consider cylinders of the type $Q_r^\lambda \subset \Omega_T$ with vertex contained in Ω_0 and $r \leq r_m$; various additional restrictions on the size of r will be imposed in due the course of the proof, as for Theorem 1.5. We start proving that for every $\varepsilon \in (0,1)$, there exists a radius $r_\varepsilon \equiv r_\varepsilon(n,m,p,\nu,L,\omega(\cdot),F(\cdot),\varepsilon,H_3) \in (0,r_m)$ such that

$$E_s(Du, Q_\varrho^\lambda) < \lambda\varepsilon \tag{6.5}$$

holds whenever $\varrho \in (0, r_\varepsilon]$ and $Q_\varrho^\lambda \subset \Omega_T$; we recall that the number s has been defined in (4.2). Repeating the scheme of the proof of Theorem 1.5, we start taking $c_1 = 8^N H_3$ in Theorem 3.1 and this determines the constant c_2 in (3.3) as a function of n,m,p,ν,L,H_3. In view of the application of Theorems 3.2 and 3.3, we this time let

$$A := \frac{10^{12pN}\max\{c_2, 200\}}{\varepsilon}\,, \quad B := \frac{10^5}{\varepsilon}\,, \quad \delta := \frac{\varepsilon}{10^7}\,, \quad \bar\varepsilon := \frac{\varepsilon}{10^{5N}}\,. \tag{6.6}$$

With the choice in (6.6) we determine $\sigma_1 \equiv \sigma_1(n,m,p,\nu,L,\omega(\cdot),\varepsilon,H_3)$ and $\sigma_2 \equiv \sigma_2(n,m,p,\nu,L,\varepsilon,H_3)$ from Theorems 3.2 and 3.3, respectively and finally define

$$\sigma := \min\{\sigma_1, \sigma_2, 16^{-Np}\} \in (0, 1/4)$$

that as such only depends on $n,m,p,\nu,L,\omega(\cdot),\varepsilon,H_3$. Again, all the constants $\bar c_1,\ldots,\bar c_4$ defined through Lemmas 3.1-3.3 and after Section 3.2 are now fixed as quantities depending only on $n,m,p,\nu,L,\omega(\cdot),\varepsilon,H_3$, and accordingly, the constant c_f is defined as in (4.9). With σ being now chosen, we can once again consider the chain of shrinking cylinders introduced in Section 3.2

$$Q_j \equiv Q_{r_j}^\lambda(x_0,t_0)\,, \qquad r_j = \sigma^j r\,, \qquad \text{where } \ r \in (\sigma\bar R_0, \bar R_0]\,, \tag{6.7}$$

for every integer $j \geq 0$; the maps w_j and v_j are then defined as in (3.7) and (3.8), respectively, while the starting radius $\bar R_0$ will be defined as follows. To start with, we recall that $R_1 \equiv R_1(n,m,p,\nu,L,\omega(\cdot),H_3)$ still denotes the radius considered in Theorems 3.1 and 3.2 while we take $R_2 \equiv R_2(n,m,p,\nu,L,\omega(\cdot),F(\cdot),\varepsilon,H_3)$ such that the condition

$$\begin{aligned}&\sup_{(x_0,t_0)\in\Omega_T}\left[\mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;4R_2)\right]^{1/a}\\ &\qquad +\lambda[\omega(R_2)]^{\min\{1,2/p\}} + \lambda\int_0^{4R_2}\omega(\varrho)\,\frac{d\varrho}{\varrho} \leq \frac{\sigma^{30N}\lambda\varepsilon}{10^8 c_f H_3}\end{aligned} \tag{6.8}$$

holds; this is allowed by (6.4). We now set $\bar R_0 := \min\{R_1, R_2, r_m\}/4$ so that $\bar R_0 \equiv \bar R_0(n,m,p,\nu,L,\omega(\cdot),F(\cdot),\varepsilon,H_3)$ and the definition in (6.7) becomes effective. In turn, computations as in (3.22) and (4.13) give

$$\begin{aligned}&\sup_{0<\varrho\leq R_2}\ \sup_{(x_0,t_0)\in\Omega_T}\left(\varrho^{\gamma-N}\int_{Q_\varrho^\lambda(x_0,t_0)}|F|^\gamma\,dx\,dt\right)^{1/(a\gamma)}\\ &+\sup_{(x_0,t_0)\in\Omega_T}\left[\sum_{j=0}^\infty\mu_j(x_0,t_0)\right]^{1/a} + \lambda[\omega(R_2)]^{\min\{1,2/p\}} + \lambda\sum_{i=0}^\infty\omega(r_i) \leq \frac{\lambda}{H_2H_3}\,,\end{aligned} \tag{6.9}$$

where we have set

$$H_2 := \frac{10^8 c_f}{\sigma^{20N}\varepsilon}\,. \tag{6.10}$$

The numbers $\mu_j \equiv \mu_j(x_0,t_0)$, with respect to the chain of cylinders in (6.7), are then defined as in (3.20). The key to the proof of (6.5) relies on showing that for every integer $j \geq 1$ the following inequality:

$$E_{j+1} = \left(\fint_{Q_{j+1}} |Du-(Du)_{Q_{j+1}}|^s\,dx\,dt\right)^{1/s} < \lambda\varepsilon \tag{6.11}$$

is true. Indeed, let us assume for a moment that (6.11) holds with the choice $r_\varepsilon := \sigma^2\bar{R}_0$. Consider $\varrho \leq \sigma^2\bar{R}_0$; this means there exists an integer $k \geq 2$ such that $\sigma^{k+1}\bar{R}_0 < \varrho \leq \sigma^k\bar{R}_0$. Therefore we have $\varrho = \sigma^k r$ for some $r \in (\sigma\bar{R}_0, \bar{R}_0]$ and (6.5) follows from (6.11) with this particular choice of r. We now prove (6.11); with $j \geq 1$ it will be sufficient to prove the following implication:

$$\begin{aligned} &\left(\fint_{Q_{j+1}} |Du|^p\,dx\,dt\right)^{1/p} \geq \frac{\lambda\varepsilon}{50} \\ &\Longrightarrow E_{j+1} \leq \frac{\varepsilon E_j}{300} + \frac{2c_f\lambda}{\sigma^N}\omega(r_j) + \frac{2c_f\lambda^{1-a}}{\sigma^N}\mu_{j-1}\,. \end{aligned} \tag{6.12}$$

Indeed, if the first inequality in the previous display does not hold then (6.11) follows trivially. In turn, observe that if (6.12) holds, as in any case we have $E_j \leq 2\lambda$ by (6.1) and (6.9)-(6.10) are in force, it follows that $E_{j+1} \leq \varepsilon\lambda/50$ as a direct consequence of (6.12). All in all, we are reduced to check the validity of (6.12). To this aim, note that (6.12) is essentially similar to (4.24); therefore we may proceed as for (4.24) by taking into account the different choice of the constants A, σ and H_2 made here. We apply Lemmas 3.4 and 3.5 with the number ε considered in (6.5) and $k = 1$ as far as Lemma 3.5 is concerned. In particular, in this last lemma we use assumption (3.56) rather than (3.48), and this is possible because now the first inequality in (6.12) is in force. Note also that by the definition in (6.1) we also have that

$$\left(\fint_{Q_j} |Du|^p\,dx\,dt\right)^{1/p} \leq H_3\lambda \tag{6.13}$$

so that (3.47) is in force. The final outcome is again that the following inequalities hold:

$$\sup_{\frac12 Q_j}|Dw_j|,\ \sup_{\frac12 Q_{j-1}}|Dw_{j-1}| \leq A\lambda \quad \text{and} \quad \sup_{\frac14 Q_j}|Dv_j| \leq A\lambda\,.$$

Next, Theorem 3.2 applied to w_{j-1} gives this time

$$\operatorname*{osc}_{Q_j} Dw_{j-1} \leq \frac{\lambda\varepsilon}{10^5}\,,$$

which is the proper analog of (4.20) in this context. Proceeding as after (4.20) we can then prove (4.21) with the current choice of A, made in (6.6). This allows to repeat the calculations done in (4.23)-(4.24) thereby concluding with the inequality in the second line of (6.12) and therefore with (6.5).

Step 2: Conclusion. We conclude with yet another version of the previous argument and we start taking $R_3 \leq r_m$ such that

$$\sup_{0<\varrho\leq R_3}\ \sup_{(x,t)\in Q_0} E_s(Du, Q_\varrho^\lambda(x,t)) \leq \frac{\sigma^{20N}\lambda\varepsilon}{10^5} \tag{6.14}$$

and then we set $\tilde{R}_0 := \min\{R_1, R_2, R_3, r_m\}/2$, so that $\tilde{R}_0$ ultimately depends again on $n, m, p, \nu, L, \omega(\cdot), F(\cdot), \varepsilon, H_3$ only. The sequence of shrinking cylinders $\{Q_j\}$ is now defined as $Q_j \equiv Q^\lambda_{r_j}(x_0,t_0)$ with $r_j = \sigma^j \tilde{R}_0$ for $j \geq 0$. We will prove that

$$
|(Du)_{Q_h} - (Du)_{Q_k}| \leq \frac{\lambda\varepsilon}{12} \qquad \text{holds whenever } 2 \leq k \leq h\,. \tag{6.15}
$$

Here we show how to use (6.15) to finish the proof and to verify (6.3) with the choice $r_\varepsilon := \sigma^2 \tilde{R}_0$. Indeed, let us fix $0 < \rho < \varrho \leq r_\varepsilon$. This means that there exist two integers, $2 \leq k \leq h$, such that $\sigma^{k+1}\tilde{R}_0 < \varrho \leq \sigma^k \tilde{R}_0$ and $\sigma^{h+1}\tilde{R}_0 < \rho \leq \sigma^h \tilde{R}_0$. Applying (6.14) and Hölder's inequality we get

$$
\begin{aligned}
|(Du)_{Q^\lambda_\varrho(x_0,t_0)} - (Du)_{Q_{k+1}}| &\leq \fint_{Q_{k+1}} |Du - (Du)_{Q^\lambda_\varrho(x_0,t_0)}|\,dx\,dt \\
&\leq \sigma^{-N} E_s(Du, Q^\lambda_\varrho(x_0,t_0)) \leq \frac{\lambda\varepsilon}{10}\,,
\end{aligned}
$$

and, similarly,

$$
|(Du)_{Q^\lambda_\varrho(x_0,t_0)} - (Du)_{Q_{h+1}}| \leq \frac{\lambda\varepsilon}{10}\,.
$$

At this stage, the inequalities appearing in the last two displays and (6.15) allows to conclude with (6.3). It remains to prove (6.15); to this aim we introduce the sets

$$
\mathcal{L} := \left\{ j \in \mathbb{N} \ : \ \left(\fint_{Q_j} |Du|^p\,dx\,dt \right)^{1/p} < \frac{\lambda\varepsilon}{50} \right\}\,,
$$

and $\mathcal{C}^s_i = \{j \in \mathbb{N} : i \leq j \leq i+s,\ i \in \mathcal{L},\ j \notin \mathcal{L} \text{ if } j > i\}$ for $s \in \mathbb{N}$ and, finally, the number $j_e := \min \mathcal{L}$. When $j_e = \infty$ this means $\mathcal{L}$ is empty and the first inequality in (6.12) holds for every $j \geq 1$. We are now ready for the proof of (6.15), obviously assuming $k < h$ and distinguishing three different cases, i.e.: $k < h \leq j_e$, $j_e \leq k < h$ and $k < j_e < h$. If $k < h \leq j_e$; we use (6.12) and the definition of j_e to infer that the inequality

$$
E_{j+1} \leq \frac{1}{2}E_j + \frac{2c_f\lambda}{\sigma^N}\omega(r_j) + \frac{2c_f}{\sigma^N}\mu_{j-1} \tag{6.16}
$$

holds for every $j \in \{k-1, \ldots, h-2\}$ so that, summing up and using (6.9) and (6.14) yields

$$
\sum_{i=k}^{h-1} E_i \leq E_{k-1} + \frac{4c_f\lambda}{\sigma^N}\sum_{j=0}^{\infty}\omega(r_j) + \frac{4c_f\lambda^{1-a}}{\sigma^N}\sum_{j=0}^{\infty}\mu_j \leq \frac{\sigma^{2N}\lambda\varepsilon}{50}
$$

and (6.15) follows since, as in (4.26), we have

$$
|(Du)_{Q_h} - (Du)_{Q_k}| \leq \sum_{i=k}^{h-1} |(Du)_{Q_{i+1}} - (Du)_{Q_i}| \leq \sigma^{-N}\sum_{i=k}^{h-1} E_i \leq \frac{\lambda\varepsilon}{50}\,.
$$

Next, if $j_e \leq k < h$, we first prove the inequalities

$$
|(Du)_{Q_h}| \leq \frac{\lambda\varepsilon}{25} \qquad \text{and} \qquad |(Du)_{Q_k}| \leq \frac{\lambda\varepsilon}{25} \tag{6.17}
$$

and then conclude with (6.15). In turn, we confine ourselves to prove the first inequality in (6.17), the proof of the second one being completely similar when $k > j_e$ (note that $|(Du)_{Q_k}| \leq \lambda\varepsilon/25$ is trivial if $k = j_e \in \mathcal{L}$). If $h \in \mathcal{L}$, the first inequality in (6.17) follows by definition of $\mathcal{L}$. If $h \notin \mathcal{L}$, then, as $h > j_e$, we take $\mathcal{C}^{s_h}_{i_h}$ with $s_h > 0$, such that $h \in \mathcal{C}^{s_h}_{i_h}$; notice that $h > i_h$ as $h \notin \mathcal{L} \ni i_h$. Then (6.12) implies (6.16) for $j \in \{i_h, \ldots, i_h + s_h - 1\}$. Summing up and using (6.9) and (6.14) then gives

$$\sum_{i=i_h}^{i_h+s_h} E_i \le 2E_{i_h} + \frac{4c_f\lambda}{\sigma^N}\sum_{j=0}^{\infty}\omega(r_j) + \frac{4c_f\lambda^{1-a}}{\sigma^N}\sum_{j=0}^{\infty}\mu_{j-1} \le \frac{\sigma^{2N}\lambda\varepsilon}{50}\,.$$

As in (4.26), we have

$$|(Du)_{Q_h} - (Du)_{Q_{i_h}}| \le \sigma^{-N}\sum_{i=i_h}^{h-1} E_i \le \sigma^{-N}\sum_{i=i_h}^{i_h+s_h} E_i \le \frac{\lambda\varepsilon}{50}$$

and then, using that $|(Du)_{Q_{i_h}}| \le \lambda\varepsilon/50$ as $i_h \in \mathcal{L}$, we have

$$|(Du)_{Q_h}| \le |(Du)_{Q_{i_h}}| + |(Du)_{Q_h} - (Du)_{Q_{i_h}}| \le \frac{\lambda\varepsilon}{25}\,,$$

that is (6.17). The last remaining case is when $k < j_e < h$. In fact, this case can be treated by a combination of the arguments used for the preceding ones: we prove that the inequalities in display (6.17) still hold. Indeed, the first inequality in (6.17) follows exactly as in the second case. As for the second estimate in (6.17), let us remark that, as $j_e \in \mathcal{L}$, we have that $|(Du)_{Q_{j_e}}| \le \lambda\varepsilon/50$. On the other hand, we can use the first case $k < h \le j_e$ with $h = j_e$, thereby obtaining $|(Du)_{Q_{j_e}} - (Du)_{Q_k}| \le \lambda\varepsilon/50$ and therefore the second inequality in (6.17) follows via triangle inequality. The proof is complete. □

Proof of Theorems 1.1-1.2. We first recall a few basic properties and definitions concerning Lorentz spaces, we refer for instance to [13, 35] for more details. Let $G\colon \Omega_T \to \mathbb{R}^{mn}$ be a measurable function; we say that G belongs to the Lorentz space $G \in L(\xi, q)$ for $\xi \ge 1$ and $q \in (0,\infty)$ iff

$$\int_0^\infty \left(\lambda^\xi |\{(x,t) \in \Omega_T \,:\, |G(x,t)| > \lambda\}|\right)^{q/\xi} \frac{d\lambda}{\lambda} < \infty\,.$$

Denote by $|G|^*\colon [0,\infty] \to [0,\infty]$ the decreasing rearrangement of $|G|$ (see [14]); a standard characterisation of $L(\xi,q)$ for $\xi > 1$ prescribes that $G \in L(\xi,q)$ iff

$$\int_0^\infty \left(|G|^{**}(\varrho)\varrho^{1/\xi}\right)^q \frac{d\varrho}{\varrho} < \infty\,, \tag{6.18}$$

where

$$|G|^{**}(\tau) := \frac{1}{\tau}\int_0^\tau |G|^*(\varrho)\,d\varrho\,, \qquad \tau > 0\,, \tag{6.19}$$

denotes the usual maximal type operator related to $|G|^*$. We recall that the rearrangement of G is characterised by the following standard Hardy's inequality [14]:

$$\int_{\mathcal{O}} |G|(x,t)\,dx\,dt \le \int_0^{|\mathcal{O}|} |G|^*(\varrho)\,d\varrho\,,$$

which holds whenever $\mathcal{O} \subset \Omega_T$ is a measurable subset. Now, take $\gamma < N$ such that $N/2 \le \gamma$ if $p \ge 2$ and $(p^*)' \le \gamma$ if $2n/(n+2) < p < 2$, and observe that Hardy's inequality and the definition in (6.19) imply

$$\fint_{Q_\varrho(x_0,t_0)} |F|^\gamma\,dx\,dt \le \frac{1}{\omega_n\varrho^N}\int_0^{\omega_n\varrho^N} (|F|^\gamma)^*(s)\,ds \le (|F|^\gamma)^{**}(\omega_n\varrho^N) \tag{6.20}$$

and, after integration, we get

$$\sup_{(x_0,t_0)} \mathbf{P}^F_\gamma(x_0,t_0;r) \le c(n)\int_0^r \left[(|F|^\gamma)^{**}(\omega_n\varrho^N)\right]^{1/\gamma} d\varrho\,. \tag{6.21}$$

The last integral is finite since after changing variables we have

$$\int_0^r \left[(|F|^\gamma)^{**}(\omega_n \varrho^N)\right]^{1/\gamma} d\varrho \le c \int_0^\infty \left[\varrho^{\gamma/N}(|F|^\gamma)^{**}(\varrho)\right]^{1/\gamma} \frac{d\varrho}{\varrho} < \infty$$

and in turn the finiteness of the last quantity follows from the fact that $F \in L(N,1)$ implies $|F|^\gamma \in L(N/\gamma, 1/\gamma)$. Here we are using the characterisation of Lorentz spaces described in (6.18) and the fact that $\gamma < N$ so that $N/\gamma > 1$. From (6.21) and the absolute continuity of integrals we deduce that (1.16) is satisfied and the continuity of the gradient stated in Theorem 1.1 now follows from Theorem 1.6. As for Theorem 1.2, we take $\gamma < n$, still satisfying the lower bounds considered above, and notice that, as for (6.20), we have

$$\fint_{Q_\varrho(x_0,t_0)} |F|^\gamma \, dx\, dt \le \fint_{B_\varrho(x_0)} g^\gamma \, dx \le (g^\gamma)^{**}(\omega_n \varrho^n)\,,$$

where we have used (1.8). Integration of the previous inequality yields

$$\sup_{(x_0,t_0)} \mathbf{P}^F_\gamma(x_0,t_0;r) \le \int_0^r \left[(g^\gamma)^{**}(\omega_n\varrho^n)\right]^{1/\gamma} d\varrho$$

and, proceeding as after (6.21), we find that (1.16) follows from the finiteness of the last integral in the previous display, which is in turn implied by $g \in L(n,1)$ as in the case of Theorem 1.1. Again, Theorem 1.2 follows from Theorem 1.6.

7. Proof of Theorem 1.7

The proof is based on a suitable revisitation of the proofs of Theorems 1.5-1.6. Essentially, from the proof of Theorem 1.5 we retain that the local averages of Du are bounded, and from the one of Theorem 1.6 we shall infer that they converge. Specifically, we shall prove that there exists a number λ, obviously depending on the point (x_0,t_0), such that the limit in (6.2) exists at the point (x_0,t_0). This is sufficient in view of the invariance of the Lebesgue points of Du with respect to the intrisic cylinders; a fact we have already observed in the beginning of the proof of Theorem 1.5. Notice also that the difference with the proof of Theorem 1.6 lies in the fact that there we proved the existence of (6.2) uniformly with respect to the considered point, while here we only prove it at (x_0,t_0); this is the natural outcome of assuming that (1.17) holds, and that therefore the limit in (1.16) is zero at $(x,t) \equiv (x_0,t_0)$, rather than uniformly. To proceed, first of all, with $\bar{c} \equiv \bar{c}(n,m,p,\nu,L,\omega(\cdot))$ denoting the constant appearing in (1.15), we observe that we can find a number $\lambda \ge 1$ such that

$$\lambda = \tilde{c}\left(\fint_{Q_r^\lambda(x_0,t_0)} (|Du|+1)^p \, dx\, dt\right)^{1/p} + \tilde{c}\left[\mathbf{P}^F_{\gamma,\lambda}(x_0,t_0;r)\right]^{1/a}\,,$$

where $a \equiv a(\gamma)$ is the exponent defined in (1.12). This simply follows using the argument developed at the beginning of the proof of Theorems 1.3-1.4 in Section 5 that in turn employs the function $h(\lambda)$ defined in (5.1). Notice that here we are already using the finiteness assumption (1.17). This λ obviously satisfies the first inequality in (1.15) and therefore we can go back to the proof of Theorem 1.5. Revisiting the arguments there we see that we can repeat everything but the passage to the limit made in Step 2, since now we do not know that (x_0,t_0) is a Lebesgue point of Du. What we instead retain from the proof is the uniform bound on the shrinking chain of cylinders

$$\left(\fint_{Q_{r_j}^\lambda(x_0,t_0)} |Du|^p \, dx\, dt\right)^{1/p} \le \lambda\,, \qquad r_j = \sigma^j r\,,$$

where $\sigma \equiv \sigma(n, m, p, \nu, L, \omega(\cdot))$ has been defined in (4.7). From this it is then immediate to see that new uniform bound

$$\left(\fint_{Q_\varrho^\lambda(x_0,t_0)} |Du|^p \, dx \, dt \right)^{1/p} \leq \sigma^{N/p}\lambda =: H_3\lambda \tag{7.1}$$

holds whenever $\varrho \leq r$. Notice that in this way we have that H_3 only depends on $n, m, p, \nu, L, \omega(\cdot)$ and this follows from the choice made in (4.7). We can now conclude the proof appealing to the proof of Theorem 1.6 with the choice of H_3 made in (7.1). Notice that now the number σ chosen in the proof of Theorem 1.6 is different from the one introduced in Theorem 1.5 and displayed in (7.1). Let us now briefly report the necessary explanations on how to modify the proof of Theorem 1.6. We first remark that in order to carry out all the arguments there, it is not necessary to have that the gradient is bounded, but is it is sufficient to have (7.1) at our disposal. This serves to verify (6.13), which is essentially the only point where the gradient boundedness is used. Moreover, and more importantly, the condition (1.16) assumed for Theorem 1.6 is there used to infer that the limit in (6.2) exists and is uniform with respect to the choice of (x_0, t_0). On the other hand, if we only want to prove the existence of the limit in (6.2) at the point (x_0, t_0), the finiteness condition in (1.17) suffices. This can be seen observing that, in order to treat the averages $(Du)_{Q_\varrho^\lambda(x_0,t_0)}$ instead of choosing (6.8), we can now choose R_2 in order to satisfy the smallness condition at (x_0, t_0) only, i.e.:

$$\left[\mathbf{P}^F_{\gamma,\lambda}(x_0, t_0; 4R_2)\right]^{1/a} + \lambda[\omega(R_2)]^{\min\{1,2/p\}} + \lambda \int_0^{4R_2} \omega(\varrho) \, \frac{d\varrho}{\varrho} \leq \frac{\sigma^{30N}\lambda\varepsilon}{10^8 c_f} \, ,$$

where the number $\varepsilon \in (0, 1)$ has been introduced in (6.3). After this choice, and with (7.1) at disposal, it is easy to check that the proof of Theorem 1.6 works as before, while we again replace (6.14) by its non-uniform version, i.e.

$$\sup_{0<\varrho\leq R_3} E_s(Du, Q_\varrho^\lambda(x_0, t_0)) \leq \frac{\sigma^{20N}\lambda\varepsilon}{10^5} \, .$$

Proceeding as after (6.14) we then prove that for every ε there exists a radius r_ε, depending this time also on the point (x_0, t_0) such that (6.3) holds and therefore (6.2) follows. The proof is complete and the paper comes to an end.

Acknowledgements. The authors are supported by the ERC grant 207573 "Vectorial Problems" and by the Academy of Finland project "Regularity theory for nonlinear parabolic partial differential equations". The authors also thank Paolo Baroni for remarks on a preliminary version of the paper and the institute Mittag-Leffer for hospitality in the frame of the program "Evolutionary problems".

□

References

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

[2] Baroni P.: *New contributions to Nonlinear Calderón-Zygmund theory.* Ph. D. Thesis, Scuola Normale Superiore, Pisa 2012.

[3] Boccardo L. & Dall'Aglio A. & Gallouët T. & Orsina L.: Nonlinear parabolic equations with measure data. *J. Funct. Anal.* 147 (1997), 237–258.

[4] Cianchi A.: Maximizing the L^∞ norm of the gradient of solutions to the Poisson equation. *J. Geom. Anal.* 2 (1992), 499–515.

[5] Cianchi A. & Maz'ya V.: Global Lipschitz regularity for a class of quasilinear equations. *Comm. PDE* 36 (2011), 100–133.

[6] Cianchi A. & Maz'ya V.: Global boundedness of the gradient for a class of nonlinear elliptic systems. *Arch. Rat. Mech. Anal.*, to appear.

[7] DiBenedetto E.: *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993.
[8] DiBenedetto E. & Friedman A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. reine ang. Math. (Crelles J.)* 357 (1985), 1–22.
[9] DiBenedetto E. & Gianazza U. & Vespri V.: Harnack estimates for quasi-linear degenerate parabolic differential equations. *Acta Math.* 200 (2008), 181–209.
[10] DiBenedetto E. & Gianazza U. & Vespri V.: Subpotential lower bounds for nonnegative solutions to certain quasi-linear degenerate parabolic equations. *Duke Math. J.* 143 (2008), 1–15.
[11] DiBenedetto E. & Gianazza U. & Vespri V.: Forward, backward and elliptic Harnack inequalities for non-negative solutions to certain singular parabolic partial differential equations. *Ann. Scu. Norm. Sup. Pisa Cl. Sci. (V)* 9 (2010), 385–422.
[12] Duzaar F. & Mingione G.: Local Lipschitz regularity for degenerate elliptic systems. *Ann. Inst. H. Poincaré Anal. Non Linèaire* 27 (2010), 1361-1396.
[13] Grafakos L.: *Classical and modern Fourier analysis.* Pearson Edu. Inc., Upper Saddle River, NJ, 2004.
[14] Hardy G. H. & Littlewood J. E. & Pólya G.: *Inequalities. Reprint of the 1952 edition.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988.
[15] Hamburger C.: Regularity of differential forms minimizing degenerate elliptic functionals. *J. reine angew. Math. (Crelles J.)* 431 (1992), 7–64.
[16] Havin M. & Mazya V. G.: A nonlinear potential theory. *Russ. Math. Surveys* 27 (1972), 71–148.
[17] Hedberg L.I. & Wolff T.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* 33 (1983), 161–187.
[18] Jaye B. & Verbitsky I.: Local and global behaviour of solutions to nonlinear equations with natural growth terms. *Arch. Rat. Mech. Anal.* 204 (2012), 627–681.
[19] Jin T. & Mazya V. & Van Schaftingen J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. *C. R. Acad. Sci. Paris Ser. I* 347 (2009), 773-778.
[20] Kilpeläinen T. & Malý J.: Degenerate elliptic equations with measure data and nonlinear potentials. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)* 19 (1992), 591–613.
[21] Kilpeläinen T. & Malý J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* 172 (1994), 137–161.
[22] Kinnunen J. & Lewis J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* 102 (2000), 253–271.
[23] Kinnunen J. & Lewis J. L.: Very weak solutions of parabolic systems of p-Laplacian type. *Ark. Mat.* 40 (2002), 105–132.
[24] Kinnunen J. & Lukkari T. & Parviainen M.: An existence result for superparabolic functions. *J. Funct. Anal.* 258 (2010), 713–728.
[25] Kuusi T. & Mingione G.: Potential estimates and gradient boundedness for nonlinear parabolic systems. *Rev. Mat. Iberoamericana* 28 (2012), 535–576.
[26] Kuusi T. & Mingione G.: The Wolff gradient bound for degenerate parabolic equations. *J. Europ. Math. Soc.*, to appear.
[27] Kuusi T. & Mingione G.: Gradient regularity for nonlinear parabolic equations. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)* 12 (2013).
[28] Kuusi T. & Mingione G.: Riesz potentials and nonlinear parabolic equations, to appear. *Arch. Rat. Mech. Anal.*, to appear.
[29] Kuusi T. & Mingione G.: New perturbation methods for nonlinear parabolic problems. *J. Math. Pures Appl. (IX)* 98 (2012), 390–427.
[30] Kuusi T. & Mingione G.: Guide to nonlinear potential estimates. *Bull. Math. Sci.*, to appear.
[31] Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Diff. Equ.* 76 (1988), 203–212.
[32] Maz'ya V. & McOwen R.: Differentiability of solutions to second-order elliptic equations via dynamical systems. *J. Differential Equations* 250 (2011), 1137–1168.
[33] Phuc N.C. & Verbitsky I. E.: Quasilinear and Hessian equations of Lane-Emden type. *Ann. of Math. (II)* 168 (2008), 859–914.
[34] Phuc N.C. & Verbitsky I. E.: Singular quasilinear and Hessian equations and inequalities. *J. of Funct. Anal.* 256 (2009), 1875–1906.
[35] Stein E. M. & Weiss G.: *Introduction to Fourier analysis on Euclidean spaces.* Princeton Math. Ser., 32. Princeton Univ. Press, Princeton, N.J. 1971.
[36] Stein E. M.: Editor's note: the differentiability of functions in $\mathbb{R}^n$. *Ann. of Math. (II)* 113 (1981), 383–385.

[37] Stein E. M.: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton Math. Series, 43. Princeton University Press, Princeton, NJ, 1993.

[38] Urbano J.M.: *The method of intrinsic scaling. A systematic approach to regularity for degenerate and singular PDEs.* Lecture Notes in Mathematics, 1930. Springer-Verlag, Berlin, 2008.

Tuomo Kuusi, Aalto University Institute of Mathematics, P.O. Box 111000 FI-00076 Aalto, Finland

E-mail address: tuomo.kuusi@tkk.fi

Dipartimento di Matematica e Informatica, Università di Parma, Parco Area delle Scienze 53/a, Campus, 43124 Parma, Italy

E-mail address: giuseppe.mingione@unipr.it.