to appear in the Journal of the European Mathematical Society

THE WOLFF GRADIENT BOUND FOR DEGENERATE PARABOLIC EQUATIONS

TUOMO KUUSI AND GIUSEPPE MINGIONE

Abstract. The spatial gradient of solutions to non-homogeneous and degenerate parabolic equations of p-Laplacean type can be pointwise estimated by natural Wolff potentials of the right hand side measure.

Contents

1. Introduction and results

In this paper we consider non-homogeneous, possibly degenerate parabolic equations in cylindrical domains $\Omega_T = \Omega \times (-T, 0)$, where $\Omega \subset \mathbb{R}^n$ is a bounded domain, $n \geq 2$, and $T > 0$. The equations in question are quasilinear and of the type

$$u_t - \operatorname{div} a(Du) = \mu \, , \tag{1.1}$$

where in the most general case μ is a Borel measure with finite total mass, i.e.

$$|\mu|(\Omega_T) < \infty \, .$$

From now on, without loss of generality, we shall assume that the measure is defined on $\mathbb{R}^{n+1}$ by letting $\mu\lfloor_{\mathbb{R}^{n+1}\setminus\Omega_T}= 0$; therefore we shall assume that

$$|\mu|(\mathbb{R}^{n+1}) < \infty\,.$$

A chief model example for the equations treated here is given by the familiar evolutionary p-Laplacean equation

$$u_t - \operatorname{div}(|Du|^{p-2}Du) = \mu\,, \tag{1.2}$$

and in fact, when considering (1.1), we shall assume the following growth and parabolicity conditions on the C^1-vector field $a\colon \mathbb{R}^n \to \mathbb{R}^n$

$$\begin{cases} |a(z)| + |\partial a(z)|(|z|^2+s^2)^{1/2} \le L(|z|^2+s^2)^{(p-1)/2} \\ \nu(|z|^2+s^2)^{(p-2)/2}|\xi|^2 \le \langle \partial a(z)\xi,\xi\rangle \end{cases} \tag{1.3}$$

whenever $z,\xi \in \mathbb{R}^n$, where $0<\nu\le L$ are positive numbers. For the following we fix $s\ge 0$, which is a parameter that will be used to distinguish the degenerate case ($s=0$), that catches the model equation in (1.2), from the nondegenerate one ($s>0$). In this paper we shall always assume

$$p \ge 2\,.$$

The so called "singular case" $p<2$ can still be treated starting by the techniques introduced in this paper and will be presented elsewhere (see [25]) in order to make the presentation here not too long and since new and nontrivial arguments must be introduced. For further notation and definitions adopted in this paper - and especially for those concerning parabolic cylinders - we immediately refer the reader to Section 2 below; we just remark from the very beginning that in the rest of the paper λ will always denote a positive real number: $\lambda>0$.

The regularity theory for the equations considered in this paper has been established in the fundamental work of DiBenedetto, and we refer the reader to the monograph [10] for a state-of-the-art presentation of the basic aspects of the theory.

1.1. **Elliptic Wolff potential estimates.** The main aim of this paper is to provide pointwise estimates for the spatial gradient Du of solutions to (1.1) in terms of suitable nonlinear potentials of the right hand side measure μ. Our results fill a basic gap between the elliptic theory, where potential estimates are available, and the parabolic one, where this is still an open issue. For this reason, let us briefly summarize the story, that actually starts with the fundamental results of Kilpeläinen & Malý [17], who proved that when considering elliptic equations of the type

$$-\operatorname{div} a(Du) = \mu\,,$$

solutions can be pointwise estimated via Wolff potentials $\mathbf{W}^{\mu}_{\beta,p}(x_0,r)$. These are defined by

$$\mathbf{W}^{\mu}_{\beta,p}(x_0,r) := \int_0^r \left(\frac{|\mu|(B(x_0,\varrho))}{\varrho^{n-\beta p}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\,, \qquad \beta>0\,, \tag{1.4}$$

and reduce to the standard (truncated) Riesz potentials when $p=2$

$$\mathbf{W}^{\mu}_{\beta/2,2}(x_0,r) = \mathbf{I}^{\mu}_{\beta}(x,r) = \int_0^r \frac{\mu(B(x_0,\varrho))}{\varrho^{n-\beta}}\frac{d\varrho}{\varrho}\,, \qquad \beta>0\,, \tag{1.5}$$

with the first equality being true for nonnegative measures. The estimate of Kilpeläinen & Malý [17] is

$$|u(x_0)| \le c \fint_{B(x_0,r)} (|u|+rs)\,dx + c\mathbf{W}^{\mu}_{1,p}(x_0,2r)\,, \tag{1.6}$$

and holds whenever $B(x,2r) \subset \Omega$ is a ball centered at x_0 with radius $2r$, with x_0 being a Lebesgue point of u; here c depends only on n,p,ν,L. Another interesting approach to (1.6) was later given by Trudinger & Wang in [40, 41] and Kuusi & Korte in [22]. This result has been upgraded to the gradient level in [37] for the case $p=2$ and then in [12, 13] for $p \geq 2-1/n$ (see also [26, 27] for relevant developments), where the following estimate is proved:

$$|Du(x_0)| \leq c \fint_{B(x_0,r)} (|Du|+s)\,dx + c\mathbf{W}^{\mu}_{1/p,p}(x_0,2r)\,, \tag{1.7}$$

for $c \equiv c(n,p,\nu,L)$. Estimates (1.6) and (1.7) are the nonlinear counterparts of the well-known estimates valid for solutions to the Poisson equation $-\triangle u = \mu$ in $\mathbb{R}^n$ - here we take $n \geq 3$, μ being a locally integrable function and u being the only solution decaying to zero at infinity. In this case such estimates are an immediate consequence of the representation formula

$$u(x_0) = \frac{1}{n(n-2)|B_1|}\int_{\mathbb{R}^n} \frac{d\mu(x)}{|x-x_0|^{n-2}}\,, \tag{1.8}$$

and on the whole space take the form

$$|u(x_0)| \leq c\mathbf{I}_2^{|\mu|}(x_0,\infty) \qquad \text{and} \qquad |Du(x_0)| \leq c\mathbf{I}_1^{|\mu|}(x_0,\infty)\,. \tag{1.9}$$

The importance of estimates as (1.6) and (1.7) mainly relies in the fact that they allow to deduce several basic properties of solutions to quasilinear equations by simply analyzing the behavior of related Wolff potentials. Indeed, Wolff potentials are an essential tool in order to study the fine properties of Sobolev functions and, more in general, to build a reasonable nonlinear potential theory [14, 15].

In this paper we concentrate on the higher order estimate (1.7) - the most delicate one - and give a natural analog of it in the case of possibly degenerate parabolic equations of p-Laplacean type as those in (1.1) and (1.2). Now, while in the non-degenerate case $p=2$ the proof of the Wolff potential (spatial) gradient estimate is similar to the one for the elliptic case, as shown in [12], the case $p \neq 2$ requires very different means. Indeed, the equations considered become anisotropic (multiple of solutions no longer solve similar equations) and as a consequence all the a priori estimates available for solutions - starting from those concerning the homogeneous case $\mu = 0$ - are not homogeneous. Ultimately, the iteration methods introduced in [17, 40, 41, 36, 37, 12] cannot be any longer applied. As a matter of fact, even the notion of potentials used must be revisited in a way that fits the local structure of the equations considered. This is not only a technical fact but instead is linked to behavior that the p-Laplacean type degeneracy exhibits in the parabolic case. Indeed, as we shall see in the next section, so-called intrinsic geometry of the problem will appear [9, 10].

1.2. **The intrinsic approach, and intrinsic potentials.** Due to the anisotropic structure of the equations considered here, the use - both in formulation of the results, and in the techniques employed - of the concept of *intrinsic geometry*, widely discussed in [10], is needed. This prescribes that, although the equations considered are anisotropic, they behave as isotropic equations when considered in space/time cylinders *whose size depend on the solution itself.* To outline how such an *intrinsic approach* works, let us consider a domain, actually a cylinder Q, where, roughly speaking, the size of the gradient norm is approximately λ – possibly in some integral averaged sense – i.e.

$$|Du| \approx \lambda > 0\,. \tag{1.10}$$

In this case we shall consider cylinders of the type

$$Q = Q_r^{\lambda}(x_0,t_0) \equiv B(x_0,r) \times (t_0 - \lambda^{2-p}r^2, t_0)\,, \tag{1.11}$$

where $B(x_0, r) \subset \mathbb{R}^n$ is the usual Euclidean ball centered at x_0 and with radius $r > 0$. Note that, when $\lambda \equiv 1$ or when $p = 2$, the cylinder in (1.11) reduces to the standard parabolic cylinder given by

$$Q_r(x_0, t_0) \equiv Q_r^1(x_0, t_0) \equiv B(x_0, r) \times (t_0 - r^2, t_0)\,.$$

Indeed, the case $p = 2$ is the only one admitting a non-intrinsic scaling and local estimates have a natural homogeneous character. In this case the equations in question are automatically non-degenerate. The *heuristics of the intrinsic scaling method* can now be easily described as follows: assuming that in a cylinder Q as in (1.11), the size of the gradient is approximately λ as in (1.10). Then we have that the equation

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

looks like

$$u_t = \operatorname{div}(\lambda^{p-2}Du) = \lambda^{p-2}\triangle u$$

which, after a scaling, that is considering $v(x, t) := u(x_0 + \varrho x, t_0 + \lambda^{2-p}\varrho^2 t)$ in $B(0, 1) \times (-1, 0)$, reduces to the heat equation

$$v_t = \triangle v$$

in $B(0, 1) \times (-1, 0)$. This equation, in fact, admits favorable a priori estimates for solutions. The success of this strategy is therefore linked to a rigorous construction of such cylinders in the context of intrinsic definitions. Indeed, the way to express a condition as (1.10) is typically in an averaged sense like for instance

$$\left(\frac{1}{|Q_r^\lambda|}\int_{Q_r^\lambda} |Du|^{p-1}\,dx\,dt\right)^{1/(p-1)} = \left(\fint_{Q_r^\lambda} |Du|^{p-1}\,dx\,dt\right)^{1/(p-1)} \approx \lambda\,. \tag{1.12}$$

A problematic aspect in (1.12) occurs as the value of the integral average must be comparable to a constant which is involved in the construction of its support $Q_r^\lambda \equiv Q_r^\lambda(x_0, t_0)$, exactly according to (1.11). As a consequence of the use of such intrinsic geometry, all the a priori estimates for solutions to evolutionary equations of p-Laplacean type admit a formulation that becomes natural only when expressed in terms of intrinsic parameters and cylinders as Q_r^λ and λ.

The first novelty of this paper is that *we shall adopt the intrinsic geometry approach in the context of nonlinear potential estimates.* This will naturally give raise to a class of *intrinsic Wolff potentials* that reveal to be the natural objects to consider, as their structure allows to recast the behavior of the Barenblatt solution - the so-called nonlinear fundamental solution - for solutions to general equations; see Section 1.5 below. The intrinsic potential estimates will then imply estimates via standard potentials, *in a way that respects the natural scaling of the equations considered*; see Section 1.3 below.

To begin with, in accordance to the standard elliptic definition in (1.4), and with $\lambda > 0$ at the moment being only *an arbitrary free parameter*, we define

$$\mathbf{W}_\lambda^\mu(x_0, t_0; r) := \int_0^r \left(\frac{|\mu|(Q_\varrho^\lambda(x_0, t_0))}{\lambda^{2-p}\varrho^{N-1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\,, \qquad N := n + 2\,. \tag{1.13}$$

In the above construction, we therefore start building the relevant potential by using intrinsic cylinders $Q_\varrho^\lambda(x_0, t_0)$ as in (1.11), while N is the usual parabolic dimension; notice that when $p = 2$ the one in (1.13) becomes a standard caloric Riesz potential, see also Remark 1.2 below. Also, the integral appearing in (1.13) is the natural intrinsic counterpart of the Wolff potential $\mathbf{W}^\mu_{1/p,p}$ intervening in the elliptic gradient estimate (1.7), and it reduces to it when μ is time independent; see also Theorem 1.3 below.

The connection with solutions to (1.1), therefore making $\mathbf{W}^{\mu}_{\lambda}$ *an intrinsic potential* in this context, is then given by the following:

Theorem 1.1 (Intrinsic potential bound). *Let u be a solution to* (1.1) *such that Du is continuous in Ω_T and that $\mu \in L^1$. There exists a constant $c > 1$, depending only on n, p, ν, L, such that if $\lambda > 0$ is a generalized root of*

$$\lambda = c\beta + c\int_0^{2r}\left(\frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho} \quad (= c\beta + c\mathbf{W}^{\mu}_{\lambda}(x_0,t_0;2r)) \tag{1.14}$$

and if

$$\left(\fint_{Q^{\lambda}_r}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} \le \beta\,, \tag{1.15}$$

where $Q^{\lambda}_{2r} \equiv Q^{\lambda}_{2r}(x_0,t_0) \equiv B(x_0,2r)\times(t_0-\lambda^{2-p}4r^2,t_0) \subset \Omega_T$ is an intrinsic cylinder with vertex at (x_0,t_0), then

$$|Du(x_0,t_0)| \le \lambda\,. \tag{1.16}$$

The meaning of generalized root is clarified in Remark 1.1 below. Statements as the one of Theorem 1.1, i.e. involving intrinsic quantities and cylinders, are completely natural when describing the local properties of the evolutionary p-Laplacean equation (see for instance [10]). Indeed, a careful reading of its proof easily shows that if Theorem 1.1 holds for a certain constant c, then it also holds for any larger constant; as a consequence we obtain the following:

Reformulation of Theorem 1.1. *There exists a constant $c \ge 1$, depending only on n, p, ν, L, such that whenever $Q^{\lambda}_r \equiv Q^{\lambda}_r(x_0,t_0) \subset \Omega_T$ then*

$$c\left(\fint_{Q^{\lambda}_r}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} + c\mathbf{W}^{\mu}_{\lambda}(x_0,t_0;2r) \le \lambda \Rightarrow |Du(x_0,t_0)| \le \lambda\,. \tag{1.17}$$

In this way, when $\mu \equiv 0$, the previous reformulation gives back the classical gradient bound of DiBenedetto [10], see Theorem 3.3 below, that is

$$c\left(\fint_{Q^{\lambda}_r}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} \le \lambda \Rightarrow |Du(x_0,t_0)| \le \lambda\,.$$

Remark 1.1 (Generalized roots and their existence). By saying that λ is a generalized root of (1.14), where $\beta > 0$ and $c \ge 1$ are given constants, we mean a (the smallest can be taken) positive solution of the previous equation, with the word generalized referring to the possibility that no root exists in which case we simply set $\lambda = \infty$. The main point is that, given $\beta > 0$, the existence of a finite root is guaranteed when

$$\mathbf{W}^{\mu}_{1}(x_0,t_0;2r) = \int_0^{2r}\left(\frac{|\mu|(Q_{\varrho}(x_0,t_0))}{\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho} < \infty\,. \tag{1.18}$$

Here recall that μ is defined on the whole $\mathbb{R}^{n+1}$. For this, let us consider the function

$$h(\lambda) := \lambda - c\beta - c\lambda^{\frac{p-2}{p-1}}\int_0^{2r}\left(\frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho}$$

defined for $\lambda > 0$. Observe that $h(\cdot)$ is a continuous function and moreover $h(\lambda) < 0$ for $\lambda < c\beta$. On the other hand it holds that

$$\lim_{\lambda\to\infty} h(\lambda) \ge \lim_{\lambda\to\infty}\left[\lambda - c\beta - c\mathbf{W}^{\mu}_{1}(x_0,t_0;2r)\lambda^{\frac{p-2}{p-1}}\right] = \infty\,.$$

Therefore there exists λ solving $h(\lambda) = 0$, that is, a solution to (1.14). Of course the existence of a generalized root does not suffice to apply Theorem 1.1 in that the intrinsic relation (1.15) still has to be satisfied. This problem is linked to the one of finding an intrinsic cylinder $Q_{2r}^{\lambda} \subset \Omega_T$ where (1.15) does hold; this is for instance the case when $Q_{2r} \subset \Omega_T$ and $\lambda \geq 1$. Theorem 1.2 below deals precisely with this situation. Another example of significant situation is given in Section 1.5 below.

Remark 1.2. In the case $p = 2$ it is easy to see that Theorem 1.1 implies the bound

$$|Du(x_0, t_0)| \leq c \fint_{Q_r} |Du| \, dx \, dt + c\mathbf{I}_1^{\mu}(x_0, t_0; 2r) \tag{1.19}$$

whenever $Q_{2r} \equiv Q_{2r}(x_0, t_0) \subset \Omega_T$ is a standard parabolic cylinder, where

$$\mathbf{I}_1^{\mu}(x_0, t_0; 2r) := \int_0^{2r} \frac{|\mu|(Q_\varrho(x_0, t_0))}{\varrho^{N-1}} \frac{d\varrho}{\varrho} \tag{1.20}$$

is the parabolic Riesz potential of μ and $N = n + 2$ is the parabolic dimension. Estimate (1.19) has been originally obtained in [12]. When instead considering the associated elliptic problem and μ is time independent, Theorem 1.1 gives back the elliptic estimate (1.7). For this see also Theorem 1.3 below.

Remark 1.3 (Stability of the constants)**.** We remark that the constant c appearing in Theorem 1.1 is stable when $p \to 2$ (and indeed the estimate (1.19) is covered by the proof). We also give an approach to the gradient Hölder continuity of solutions to degenerate parabolic equations yielding a priori estimates with stable constants when $p \to 2$.

1.3. **Intrinsic estimates yield explicit potential estimates.** The next result tells that Theorem 1.1 always yields a priori estimates on arbitrary standard parabolic cylinders, and we can therefore abandon the intrinsic geometry. As a consequence, standard Wolff potentials, considered with respect to the parabolic metric, appear (recall the definition in (1.13) and compare it with the one in (1.20)).

Theorem 1.2 (Parabolic Wolff potential bound)**.** *Let u be a solution to* (1.1) *such that Du is continuous in Ω_T and* (1.25) *holds. There exists a constant c, depending only on n, p, ν, L, such that*

$$\begin{aligned} |Du(x_0, t_0)| &\leq c \fint_{Q_r} (|Du| + s + 1)^{p-1} \, dx \, dt \\ &\qquad + c \left[\int_0^{2r} \left(\frac{|\mu|(Q_\varrho(x_0, t_0))}{\varrho^{n+1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho} \right]^{p-1} \\ &= c \fint_{Q_r} (|Du| + s + 1)^{p-1} \, dx \, dt + c[\mathbf{W}_1^{\mu}(x_0, t_0; 2r)]^{p-1} \end{aligned} \tag{1.21}$$

holds whenever $Q_{2r} \equiv Q_{2r}(x_0, t_0) \equiv B(x_0, 2r) \times (t_0 - 4r^2, t_0) \subset \Omega_T$ is a standard parabolic cylinder with vertex at (x_0, t_0).

To check the consistency of estimate (1.21) with the ones already present in the literature we observe that when $\mu \equiv 0$, estimate (1.21) reduces the classical L^∞-gradient bound available for solutions to the evolutionary p-Laplacean equation; see [10, Chapter 8, Theorem 5.1']. The importance of estimates as those in Theorems 1.1-1.2 - as well as those of estimates (1.6)-(1.7) - is rather clear: *the growth behavior of solutions can be now completely described via potentials of the right hand side data μ*, completely bypassing the structure of the equation. For instance, all kinds of regularity results for the gradient in rearrangement invariant functions spaces *follow at once* by the properties of Wolff potentials, which are known by other means. For such aspects and applications we refer for instance [17, 38].

Proof of Theorem 1.2. Without loss of generality we may assume that the quantity $\mathbf{W}_1^{\mu}(x_0, t_0; 2r)$ in (1.18) is finite, otherwise there is nothing to prove. Next, let us consider the function

$$h(\lambda) := \lambda - c\lambda^{\frac{p-2}{p-1}} A(\lambda)\,,$$

where

$$A(\lambda) := \left(\frac{1}{|Q_r|}\int_{Q_r^\lambda} (|Du| + s + 1)^{p-1}\, dx\, dt\right)^{1/(p-1)} + \int_0^{2r} \left(\frac{|\mu|(Q_\varrho^\lambda)}{\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}$$

and c is again the constant appearing in Theorem 1.1. We consider the function $h(\cdot)$ defined for all those λ such that $Q_r^\lambda \subset \Omega_T$; observe that the domain of definition of $h(\cdot)$ includes $[1, \infty)$ as $Q_r^\lambda \subset Q_r \subset \Omega_T$ when $\lambda \geq 1$. Again, observe that $h(\cdot)$ is a continuous function and moreover $h(1) < 0$ as $c > 1$. On the other hand, observe that

$$\lim_{\lambda\to\infty} h(\lambda) \geq \lim_{\lambda\to\infty} \lambda - c\lambda^{\frac{p-2}{p-1}} B = \infty\,,$$

where

$$B := \left(\fint_{Q_r} (|Du| + s + 1)^{p-1}\, dx\, dt\right)^{1/(p-1)} + \int_0^{2r} \left(\frac{|\mu|(Q_\varrho)}{\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\,.$$

It follows that there exists a number $\lambda > 1$ such that $h(\lambda) = 0$, that is λ solves (1.14) with

$$\begin{aligned} \beta &= \left(\fint_{Q_r^\lambda} (|Du| + s + 1)^{p-1}\, dx\, dt\right)^{1/(p-1)} \\ &= \lambda^{\frac{p-2}{p-1}} \left(\frac{1}{|Q_r|}\int_{Q_r^\lambda} (|Du| + s + 1)^{p-1}\, dx\, dt\right)^{1/(p-1)}. \end{aligned}$$

Therefore we can apply Theorem 1.1 and (1.16) gives

$$\lambda + |Du(x_0, t_0)| \leq 2c\beta + 2c\int_0^{2r} \left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\,. \tag{1.22}$$

On the other hand, observe that by Young's inequality with conjugate exponents $((p-1)/(p-2), p-1)$ we have

$$\begin{aligned} 2c\beta &\leq \frac{\lambda}{4} + \frac{\tilde{c}}{|Q_r|}\int_{Q_r^\lambda} (|Du| + s + 1)^{p-1}\, dx\, dt \\ &\leq \frac{\lambda}{4} + \tilde{c}\fint_{Q_r} (|Du| + s + 1)^{p-1}\, dx\, dt \end{aligned}$$

where we have also used that $Q_r^\lambda \subset Q_r$ as $\lambda > 1$, and $\tilde{c}$ depends only on n, p, ν, L. Similarly, observe that

$$2c\int_0^{2r} \left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho} \leq \frac{\lambda}{4} + \tilde{c}\left[\int_0^{2r} \left(\frac{|\mu|(Q_\varrho)}{\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\right]^{p-1}, \tag{1.23}$$

where again $\tilde{c} \equiv \tilde{c}(n, p, \nu, L)$. The last two inequalities and (1.22) yield (1.21). □

Finally, when μ is time independent, or admits a favorable decomposition, it is possible to get rid of the intrinsic geometry effect in the potential terms. The main point is that we avoid the loss in the right hand side caused by the rough estimate

$$|\mu|(Q_\varrho^\lambda) \leq |\mu|(Q_\varrho)\,, \qquad \text{for } \lambda \geq 1\,,$$

used in the proof of Theorem 1.2 (which is anyway the best possible in that generality). We indeed go back to the elliptic regime; the result is in the next theorem.

Theorem 1.3 (Elliptic-Parabolic Wolff potential bound)**.** *Let u be a solution to* (1.1) *such that Du is continuous in Ω_T and* (1.25) *holds. Assume that the measure μ satisfies*

$$|\mu| \leq \mu_0 \otimes f\,,$$

where $f \in L^\infty(-T,0)$ and μ_0 is a Borel measure on Ω with finite total mass; here the symbol $\times$ stands for the usual tensor product of measures. Then there exists a constant c, depending only on n,p,ν,L, such that

$$|Du(x_0,t_0)| \leq c \fint_{Q_r} (|Du|+s+1)^{p-1}\,dx\,dt + c\|f\|_{L^\infty}^{1/(p-1)} \mathbf{W}^{\mu_0}_{1/p,p}(x_0,2r) \tag{1.24}$$

whenever $Q_{2r}(x_0,t_0) \equiv B(x_0,2r)\times(t_0-4r^2,t_0) \subset \Omega_T$ is a standard parabolic cylinder having (x_0,t_0) as vertex. The (elliptic) Wolff potential $\mathbf{W}^{\mu_0}_{1/p,p}$ is defined in (1.4).

Proof. Proceed as for Theorem 1.2 until estimate (1.23); this has in turn to be replaced by

$$\begin{aligned}\int_0^{2r}\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho} &\leq \|f\|_{L^\infty}^{1/(p-1)}\int_0^{2r}\left(\frac{|\mu_0|(B_\varrho(x_0))}{\varrho^{n-1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho}\\ &= \|f\|_{L^\infty}^{1/(p-1)}\mathbf{W}^{\mu_0}_{1/p,p}(x_0,2r)\end{aligned}$$

and (1.24) follows. □

1.4. **Approximation, a priori estimates, and regularity assumptions.** Following a traditional custom in regularity theory, Theorems 1.1-1.3 have been given in the form of a priori estimates for more regular solutions and problems. This means that when treating equations as (1.1), we are considering energy solutions i.e. $u \in L^p(-T,0;W^{1,p}(\Omega))$ such that Du is continuous in Ω_T, while the measure μ will be considered as being actually an integrable function:

$$\mu \in L^1(\mathbb{R}^{n+1})\,. \tag{1.25}$$

This is by no means restrictive in view of the available approximation and existence theory. Indeed, as described in the pioneering paper [4] (see also [19, 20]), distributional solutions $u \in L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ to Cauchy-Dirichlet problems involving equations as (1.1) - with μ being now a general Borel measure with finite total mass - are found via approximation as limits of solutions to suitably regularized problems

$$(u_h)_t - \operatorname{div} a(Du_h) = \mu_h \in C^\infty\,. \tag{1.26}$$

Here we have $u_h \in L^p(-T,0;W^{1,p}(\Omega))$, $u_h \to u$ in $L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ and $\mu_h \to \mu$ weakly* in the sense of measures. The approximating measures are canonically obtained by convolution (see for instance [35, Chapter 5]) and in the parabolic case the natural procedure is to take the so called parabolic convolution (using mollifiers backward in time). This motivates the following:

Definition 1 ([4, 19, 20])**.** A SOLA (Solution Obtained as Limits of Approximations) to (1.1) is a distributional solution $u \in L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ to (1.1) in Ω_T, such that u is the limit of solutions $u_h \in L^p(-T,0;W^{1,p}(\Omega))$ of equations as (1.26), in the sense that $u_h \to u$ in $L^{p-1}(-T,0;W^{1,p-1}(\Omega))$, $L^\infty \ni \mu_h \to \mu$ weakly* in the sense of measures and such that

$$\limsup_h |\mu_h|(Q) \leq |\mu|(\lfloor Q \rfloor_{\mathrm{par}}) \tag{1.27}$$

for every cylinder $Q = B \times (t_1, t_2) \subseteq \Omega_T$, where $B \subset \Omega$ is a bounded open subset.

We refer to (2.3) below for the definition of parabolic closure of Q, that is $\lfloor Q \rfloor_{\rm par}$; the property in (1.27) is typically satisfied when approximating, in a standard way, μ via convolution with backward-in-time mollifiers. SOLAs are actually the class of solutions which are commonly employed in the literature, since all general existence theorems are based on approximation methods; we refer to [5, 4, 12, 20, 37] for a comprehensive discussion. We also remark that, in general, distributional solutions to measure data problems do not belong to $L^p(-T,0;W^{1,p}(\Omega))$ and for this reason they are called *very weak solutions*; moreover, the uniqueness problem, i.e. finding a function class where solutions are unique, is still open – already in the elliptic case. Also SOLAs are not known to be unique but in special cases (see the discussion in [2, 4, 6, 18, 38]).

The validity of Theorems 1.1-1.3 for a SOLA now follows applying their "a priori" versions to Du_h in a suitable way, see Section 4.3 below. Summarizing, we have

Theorem 1.4. *The statements of Theorems 1.1-1.3 remain valid for SOLA* $u \in L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ *to* (1.1) *whenever* (x_0,t_0) *is a Lebesgue point of* Du.

We also remark that the previous theorem continues to hold for a local SOLA, in the sense that we can consider local approximations methods, and solutions u which are such that $u \in L^{p-1}_{\rm loc}(-T,0;W^{1,p-1}_{\rm loc}(\Omega))$; see [19, 20].

1.5. **Comparison with the Barenblatt solution.** A standard quality test for regularity estimates in degenerate parabolic problems consists of measuring the extent they allow to recast the behavior of the Barenblatt, fundamental solution; see for instance [10, Chapter 11] ad [20, 42]. Here we show that this is the case for Theorem 1.1 and concentrate on the case $p > 2$. The Barenblatt solution is an explicit very weak solution to

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

in the whole $\mathbb{R}^{n+1}$, the measure δ being the Dirac delta function charging the origin and c_b is a suitable normalizing constant depending only on n, p, and its expression is

$$\mathcal{B}_p(x,t) = \begin{cases} t^{-n/\theta}\left(c_b - \theta^{1/(1-p)}\dfrac{p-2}{p}\left(\dfrac{|x|}{t^{1/\theta}}\right)^{p/(p-1)}\right)_+^{(p-1)/(p-2)} & t > 0 \\ 0 & t \le 0\,. \end{cases}$$

Here $\theta = n(p-2) + p$ and $c_b \equiv c_b(n,p)$ is a renormalizing constant such that

$$\int_{\mathbb{R}^n} \mathcal{B}_p(x,t)\,dx = 1$$

for all $t > 0$. A direct computation reveals that the gradient of $\mathcal{B}_p(x,t)$ satisfies the estimate

$$|D\mathcal{B}_p(x_0,t_0)| \le c t_0^{-(n+1)/\theta} \tag{1.28}$$

whenever $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$; in turn this prescribes the blow-up behavior at the origin of the fundamental solution, which is typical of a situation where a Dirac measure appears. What it matters here is that Theorem 1.1 (used with $s = 0$, of course) allows to recast, quantitatively, the bound in (1.28) for a SOLA to general degenerate nonlinear equations and this tells that the intrinsic formulation given there is the correct one.

Theorem 1.5. *Let* u *be a SOLA to the equation*

$$u_t - \operatorname{div} a(Du) = \delta$$

in $\mathbb{R}^{n+1}$, under the assumptions (1.3) *with $s = 0$ and $p > 2$, and assume that $u \in L^{p-1}(-\infty, T; W^{1,p-1}(\mathbb{R}^n))$, for every $T > 0$. Then there exists a constant $c \equiv c(n, p, \nu, L)$ such that*

$$|Du(x_0, t_0)| \leq ct_0^{-(n+1)/\theta}, \qquad \theta = n(p-2) + p, \tag{1.29}$$

holds for every Lebesgue point $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$ of Du.

Proof. Take $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$ to be a Lebesgue point of Du; notice that

$$A_r^{p-1}(\lambda) := \frac{1}{|Q_r(x_0, t_0)|} \int_{Q_r^\lambda(x_0,t_0)} |Du|^{p-1}\, dx\, dt \to 0 \tag{1.30}$$

uniformly in $\lambda \in (0, \infty)$, as $r \to \infty$ for all $x_0 \in \mathbb{R}^n$, $t_0 > 0$. For $\lambda > 0$ define r_{t_0} via $\lambda^{2-p} r_{t_0}^2 = t_0$ that is $r_{t_0} = \lambda^{(p-2)/2}\sqrt{t_0}$, so that we have

$$\begin{aligned} \int_0^\infty \left(\frac{\delta(Q_\varrho^\lambda(x_0, t_0))}{\lambda^{2-p}\varrho^{n+1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho} &= \int_{\lambda^{(p-2)/2}\sqrt{t_0}}^\infty \left(\frac{1}{\lambda^{2-p}\varrho^{n+1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho} \\ &= c(n,p)\lambda^\gamma t_0^{-(n+1)/[2(p-1)]}, \end{aligned} \tag{1.31}$$

where $\gamma := [1 - (n+1)/2](p-2)/(p-1) < 0$. With c being the constant appearing in Theorem 1.1, now define, for $\lambda > 0$ and $r > 1$ the function $h_r \colon (0, \infty) \to \mathbb{R}$ as

$$\begin{aligned} & h_r(\lambda) \\ & := \lambda - c\lambda^{(p-2)/(p-1)} A_r(\lambda) - c \int_0^r \left(\frac{\delta(Q_\varrho^\lambda(x_0, t_0))}{\lambda^{2-p}\varrho^{n+1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho} \\ & = \lambda - c\lambda^{(p-2)/(p-1)} A_r(\lambda) - c \max \left\{ \int_{\lambda^{(p-2)/2}\sqrt{t_0}}^r \left(\frac{1}{\lambda^{2-p}\varrho^{n+1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho}, 0 \right\} \\ & \geq \lambda - c\lambda^{(p-2)/(p-1)} A_r(\lambda) - \tilde{c}\lambda^\gamma t_0^{-(n+1)/[2(p-1)]}, \end{aligned} \tag{1.32}$$

so that $h_r(\lambda) \to \infty$ as $\lambda \to \infty$ (recall (1.30)). On the other hand, $h_r(\cdot)$ stays negative close to zero and therefore there exists a solution $\lambda \equiv \lambda_r > 0$ of $h_r(\lambda_r) = 0$, that is a root of (1.14) with

$$\left(\fint_{Q_r^{\lambda_r}} |Du|^{p-1}\, dx\, dt \right)^{1/(p-1)} = \beta = \lambda_r^{(p-2)/(p-1)} A_r(\lambda_r).$$

Observe that the numbers $A_r(\lambda_r)$ are uniformly bounded whenever $r > 1$ by (1.30), and therefore the relation

$$\begin{aligned} \lambda_r &\leq c\lambda_r^{(p-2)/(p-1)} A_r(\lambda_r) + c_1 \lambda_r^\gamma t_0^{-(n+1)/[2(p-1)]} \\ &\leq \frac{\lambda_r}{4} + c(p)[A_r(\lambda_r)]^{p-1} + c_1 \lambda_r^\gamma t_0^{-(n+1)/[2(p-1)]} \end{aligned}$$

which is a consequence of (1.32) and of $h_r(\lambda_r) = 0$, implies that the numbers λ_r are uniformly bounded for $r > 1$. On the other hand, by Theorem 1.1 (in the version for SOLA) and the previous inequality we have

$$|Du(x_0, t_0)|^{1-\gamma} \leq \lambda_r^{1-\gamma} \leq c[A_r(\lambda_r)]^{p-1}\lambda_r^{-\gamma} + ct_0^{-(n+1)/[2(p-1)]}.$$

Letting $r \to \infty$ in the previous inequality (recall that $\gamma < 0$) by (1.30) we obtain

$$|Du(x_0, t_0)|^{1-\gamma} \leq ct_0^{-(n+1)/[2(p-1)]}$$

and (1.29) follows as $(n+1)/[2(p-1)(1-\gamma)] = (n+1)/\theta$. □

Remark 1.4. Notice that in the previous proof it is sufficient to assume that $u \in L^{p-1}_{\rm loc}(\mathbb{R}; W^{1,p-1}_{\rm loc}(\mathbb{R}^n))$ (so that we have a local SOLA) and that (1.30) holds. Notice that (1.30) in particular holds for the Barenblatt solution and indeed this is a general fact typical of solutions u to Cauchy problems whenever the initial trace of u is compactly supported, i.e. that the source term is concentrated on $t = 0$ and has a compact support. See for example [10, Chapter 11, Theorem 2.1] and [29].

1.6. **Techniques employed, and plan of the paper.** The proof of Theorem 1.1 is rather delicate and involved, and employs and extends virtually all the known aspects of the gradient regularity theory for evolutionary p-Laplacean type equations. Some very hidden details are actually needed. Indeed, a preliminary part of the proof deals with a rather wide revisitation of DiBenedetto & Friedman's regularity theory of the gradient of solutions to the p-Laplacean system

$$w_t - \operatorname{div}(|Dw|^{p-2}Dw) = 0 \tag{1.33}$$

developed in [11] and explained in detail in [10]. Here comes a first difficulty: the Hölder continuity proofs given in [10, 11] are actually suited for the special structure in (1.33) and cannot be extended to general equations if not of the special form

$$w_t - \operatorname{div}(g(|Dw|)Dw) = 0\,, \qquad g(|Dw|) \approx |Dw|^{p-2}\,. \tag{1.34}$$

The point that makes such proofs very linked to the structure in (1.34) is that they are actually based on a linearization process, which do not extend to general structures, as

$$w_t - \operatorname{div} a(Dw) = 0\,. \tag{1.35}$$

On the other hand, the methods in [10] are devised to work directly for the case of the p-Laplacean system. While Hölder continuity of the gradient has been proved assuming a regular boundary datum [30], the literature does not contain a proof of right form of the gradient Hölder continuity a priori estimates that are needed to develop in turn potential estimates in the elliptic case for general equations as in (1.35) featuring the needed a priori local estimates to work in the framework of a suitable perturbation techniques.

A peculiarity of our approach is indeed in the following: since we are dealing in the most general case with problems involving measure data, we need to deal with estimates below the natural growth exponent. Actually, in some cases solutions are not even such that $Du \in L^2$ (or at least no uniform control is achievable for the quantities $\|Du\|_{L^2}$ in the corresponding approximation processes). On the other hand, in our setting we shall need a priori estimates where the "natural integrability space" for the (spatial) gradient here is L^{p-1}. For this reason, even when considering the model case (1.33), the a priori estimates available in [10, 11] do not suffice for our purposes, and another path must be taken. To overcome such points we revisit the Hölder regularity gradient theory available and extend it to the case of general homogeneous equations as (1.35). This is done in Section 3 and has two main outcomes. The first is Theorem 3.1 below, which is a fundamental block in the proof of the potential estimates and provides a "homogeneous" decay estimate for the excess functional

$$E_q(Dw, Q_r^\lambda) := \left(\fint_{Q_r^\lambda} |Dw - (Dw)_{Q_r^\lambda}|^q \,dx\,dt \right)^{1/q}, \qquad q \geq 1$$

in an intrinsic cylinder Q_r^λ. Note that the exponent q is arbitrary and not necessary linked in any particular way to p. The main assumption (3.4) serves to consider a nondegenerate condition that ensures the possibility of a homogeneous decay estimate when the equation is considered in an intrinsic cylinder Q_r^λ. The second outcome is Theorem 3.2 below, that features a quantitative estimate that will play an important role in the proof of the main potential estimate in Theorem 1.1.

After this preliminary section we pass to the proof of Theorem 1.1. The first step is the derivation of a few local comparison estimates between the solution considered u, and solutions of homogenous equations, again on intrinsic cylinders. This serves to start the iteration mechanism leading to the desired potential estimates. The proof of Theorem 1.1 is now rather delicate, and rests on an iteration procedure combined with an exit time argument devised to rule out possible degenerate behaviors of the equation and ultimately allowing to use Theorem 3.1. The essence is the following: either the gradient Du stays bounded from above by some fraction of λ on every scale of a suitable chain of shrinking nested intrinsic cylinders

$$\cdots \subset Q^\lambda_{r_{i+1}} \subset Q^\lambda_{r_i} \subset Q^\lambda_{r_{i-1}} \subset \cdots \tag{1.36}$$

and then the proof is finished, or otherwise this does not happen. In this case we start arguing from the exit time - i.e. the first moment the bound via the fraction of the potential fails when considering such a chain. We have then that the gradient stays above a certain fraction of the potential at every scale, and this helps to rule out possible degenerate behaviors. Ultimately, this allows to verify the applicability conditions of Theorem 3.1 by using L^∞ gradient a priori estimates for related homogenous equations on $Q^\lambda_{r_i}$, that in turn homogenize since we are on intrinsic cylinders. This allows us to proceed with the iteration. A main difficulty at this stage is that all this must be realized in a suitable intrinsic scale that is in the sequence considered in (1.36), where λ is the one appearing in (1.16); therefore *the choice of the intrinsic scale must be done a priori.* Here a very delicate and subtle balance must be realized between the speed of the shrinking of the cylinders

$$\frac{r_{i+1}}{r_i} = \delta_1 \in (0,1)$$

and the constant c appearing in (1.14), and therefore in the chain (1.36) via λ (see (4.20) below). One of the crucial points of the proof is that both δ_1 and c must in the end depend only on n, p, ν, L, and such a choice must be done a priori in a way that makes later possible the application of Theorem 3.1 in the context of the exit time argument employed, avoiding dangerous vicious circles.

We would like to finally remark that the techniques introduced in this paper are the starting point for further developments: the subquadratic case can be treated too (see [25]) while new perturbation methods for parabolic systems can be implemented [28].

The main results of this paper have been announced in the Nota Lincea [23]; see also [38] for further announcements and related results.

2. Main notation and definitions

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

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

the open ball with center x_0 and radius $r > 0$; when not important, or clear from the context, we shall omit denoting the center as follows: $B_r \equiv B(x_0, r)$. Unless otherwise stated, different balls in the same context will have the same center. We shall also denote $B \equiv B_1 = B(0,1)$ if not differently specified. In a similar fashion we shall denote by

$$Q_r(x_0, t_0) := B(x_0, r) \times (t_0 - r^2, t_0)$$

the standard parabolic cylinder with vertex (x_0, t_0) and width $r > 0$. When the vertex will not be important in the context or it will be clear that all the cylinders occurring in a proof will share the same vertex, we shall omit to indicate it, simply denoting Q_r. With $\lambda > 0$ being a free parameter, we shall often consider cylinders of the type

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

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

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

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

$$\partial_{\mathrm{par}} Q := \mathcal{A} \times \{t_1\} \cup \partial\mathcal{A} \times (t_1, t_2) \tag{2.2}$$

the usual parabolic boundary of Q, and this is nothing else but the standard topological boundary without the upper cap $\mathcal{A} \times \{t_2\}$. Accordingly, we shall denote the prabloic closure of a set as

$$\lfloor Q \rfloor_{\mathrm{par}} := Q \cup \partial_{\mathrm{par}} Q\,. \tag{2.3}$$

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; of course $|\mathcal{O}|$ denotes the Lebesgue measure of $\mathcal{O}$. A similar notation is adopted if the integral is only in space or time. In the rest of the paper we shall use several times the following elementary property of integral averages:

$$\left(\fint_{\mathcal{O}} |g - (g)_{\mathcal{O}}|^q\,dx\,dt\right)^{1/q} \le 2\left(\fint_{\mathcal{O}} |g - \gamma|^q\,dx\,dt\right)^{1/q}, \tag{2.4}$$

whenever $\gamma \in \mathbb{R}^n$ and $q \ge 1$. The oscillation of g on A is instead defined as

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

Given a real valued function h and a real number k, we shall denote

$$(h-k)_+ := \max\{h-k, 0\} \qquad \text{and} \qquad (h-k)_- := \max\{k-h, 0\}\,.$$

In this paper by a (local) weak solution to (1.1) we shall mean a function

$$u \in C^0(-T, 0; L^2(\Omega)) \cap L^p(-T, 0; W^{1,p}(\Omega)) \tag{2.5}$$

such that

$$-\int_{\Omega_T} u\varphi_t\,dx\,dt + \int_{\Omega_T} \langle a(Du), D\varphi\rangle\,dx\,dt = \int_{\Omega_T} \varphi\,d\mu \tag{2.6}$$

holds whenever $\varphi \in C_c^\infty(\Omega_T)$. As in this paper we are considering only a priori estimates (see the discussion in Section 1.4) we shall restrict ourselves to examine the case when μ is an integrable function. Notice that by density the identity (2.6) remains valid whenever $\varphi \in W_0^{1,p}(\Omega_T)$ has compact support. We recall that here Du stands for the spatial gradient of u: $Du = (u_{x_i})_{1\le i\le n}$.

Remark 2.1 (Warning for the reader)**.** When dealing with parabolic equations, a standard difficulty in using test functions arguments involving the solution is that we start with solutions that, enjoying the regularity in (2.5), do not have in general time derivatives in any reasonable sense. There are several, by now standard, ways to overcome this point, for instance using a regularization procedure via so-called Steklov averages. See for instance [10, Chapter 2] for their definition and their standard use. In this paper, in order to concentrate the attention only on significant issues and to skip irrelevant details, and following a by now standard custom (see for instance [11]), we shall argue on a formal level, that is assuming when using test functions argument, that the solution has square integrable time derivatives. Such arguments can easily be made rigorous using in fact Steklov averages as for instance in [10]. We shall remark anyway this thing in other places in the paper, when regularizations procedures will be needed and we will instead proceed formally.

With $s \geq 0$ being the one defined in (1.3), we define

$$V(z) = V_s(z) := (s^2 + |z|^2)^{\frac{p-2}{4}} z \,, \qquad z \in \mathbb{R}^n \,, \tag{2.7}$$

which is easily seen to be a locally bi-Lipschitz bijection of $\mathbb{R}^n$. For basic properties of the map $V(\cdot)$ we refer to [35, Section 2.2] and related references. The strict monotonicity properties of the vector field $a(\cdot)$ implied by the left hand side in $(1.3)_2$ can be recast using the map V. Indeed there exist constants $c, \tilde{c} \equiv c, \tilde{c}(n, p, \nu) \geq 1$ such that the following inequality holds whenever $z_1, z_2 \in \mathbb{R}^n$:

$$\tilde{c}^{-1}|z_2 - z_1|^p \leq c^{-1}|V(z_2) - V(z_1)|^2 \leq \langle a(z_2) - a(z_1), z_2 - z_1 \rangle \,. \tag{2.8}$$

3. Gradient Hölder theory and homogeneous decay estimates

In this section we concentrate on homogeneous equations of the type

$$w_t - \operatorname{div} a(Dw) = 0 \tag{3.1}$$

in a given cylinder $Q = B \times (t_1, t_2)$, where $B \subset \mathbb{R}^n$ is a given ball. The degree of initial regularity of the solution considered is given by the usual energy function spaces

$$w \in C^0(t_1, t_2; L^2(B)) \cap L^p(t_1, t_2; W^{1,p}(B)) \,. \tag{3.2}$$

Most of the times we shall consider such equations defined in suitably intrinsic cylinders Q_r^λ. More precisely, without specifying this all the times, on every occasion we are dealing with a function named w and an intrinsic cylinder as Q_r^λ, it goes without saying that w solves (3.1) on Q_r^λ. In the following, we shall denote

$$\|Dw(x,t)\| := \max_i |w_{x_i}(x,t)|$$

which is equivalent to the usual norm of Du defined by $|Dw|^2 := \sum |w_{x_i}|^2$ via the obvious relations

$$||Dw|| \leq |Dw| \leq \sqrt{n}||Dw|| \,. \tag{3.3}$$

Moreover, everywhere in the following, when considering the sup operator we shall actually mean esssup. The main result of this section is

Theorem 3.1. *Suppose that w is a weak solution to* (3.1) *in Q_r^λ and consider numbers*

$$A, B, q \geq 1 \qquad \text{and} \qquad \varepsilon \in (0,1) \,.$$

Then there exists a constant $\delta_\varepsilon \in (0,1/2)$ *depending only on* $n,p,\nu,L,A,B,\varepsilon$ *but otherwise independent of* s,q, *of the solution* w *considered and of the vector field* $a(\cdot)$, *such that if*

$$\frac{\lambda}{B} \leq \sup_{Q^\lambda_{\delta_\varepsilon r}} \|Dw\| \leq s + \sup_{Q^\lambda_r} \|Dw\| \leq A\lambda \tag{3.4}$$

holds, then

$$E_q(Dw, \delta_\varepsilon Q^\lambda_r) \leq \varepsilon E_q(Dw, Q^\lambda_r) \tag{3.5}$$

holds too, where E_q *denote the excess functional*

$$E_q(Dw, Q^\lambda_\varrho) := \left(\fint_{Q^\lambda_\varrho} |Dw - (Dw)_{Q^\lambda_\varrho}|^q \, dx\, dt \right)^{1/q}, \qquad \varrho \leq r\,. \tag{3.6}$$

Moreover, (3.5) *remains true replacing* δ_ε *by a smaller number* δ, *and* δ_ε *is a non-decreasing function of* ε, $1/A$ *and* $1/B$.

The proof of the previous result is in Section 3.3 below. The main novelty in Theorem 3.1 is the following. It is readily seen that equations as (3.1) are not homogeneous as long as $p \neq 2$; in other words, by multiplying a solution w by a constant $c > 0$, we do not get solutions to a similar equation. The main drawback of this basic phenomenon is the lack of homogeneous regularity estimates. In fact, we shall see that basically all the a priori estimates of solutions involve a scaling deficit - in general the exponent $p/2$ or $p-1$ as for instance in (1.21) - which reflects the anisotropicity of the problem in question and prevents the estimates to be homogeneous. On the other hand, the iteration method we are going to exploit for the proof of Theorem 1.1 necessitates homogeneous decay estimates for the excess functional. The key will be then to implement a suitable iteration based on intrinsic cylinders in a way that (3.5) will be satisfied and the iteration will only involve homogeneous estimates. Ultimately, Theorem 3.1 reproduces in the case $p \neq 2$ the homogeneous decay estimates known for the case $p = 2$, and indeed in this case Theorem 3.1 is known to hold without assuming conditions as (3.4). The novelty here, as in the whole paper, is for the case $p > 2$.

The proof of Theorem 3.1 will take several steps. A delicate revisitation of the gradient Hölder continuity estimates derived in [10] is presented in the next section, and it differs from the usual ones in two important respects. First, the proof holds for general parabolic equations, and not only for those having the quasidiagonal structure in (1.34). Indeed, we notice that large parts of the proof given in [10] heavily uses this fact to implement a linearization procedure which is impossible to implement for general structures as in (1.1). Second, estimates proposed here involve integrals below the natural growth exponents, and work directly using the L^q norms whenever $q > 1$ - compare with the definition of $E_q(\cdot)$ in Theorem 3.1. This point, in turn, requires delicate estimates and it is crucial since we are dealing with a priori estimates for equations involving measure data.

Remark 3.1. When proving Theorem 3.1 we shall argue under the additional assumption

$$s > 0\,. \tag{3.7}$$

This is by no mean restrictive. Indeed, by a simple approximation argument - see Section 3.6 below - it is possible to reduce to such a case as the previous inequality will not play any role in the quantitative estimates. It will only be used to derive qualitative properties of solutions, and, ultimately, to use that in this case Dw is differentiable in space (see (3.16) below). For this reason, and in order to emphasize these facts, we shall in several point of this Section give the proof directly in the general case $s \geq 0$, and this will in particular happen in Section 3.2, where we find the only point where a small difference occurs between the case $s = 0$ and the one $s > 0$ in the a priori estimates. This proof is intended to be *formal* when $s = 0$, this case being indeed later justified by approximation. In particular, we make this choice also in order to keep in Theorem 3.2, the treatment close to that of DiBenedetto [10], since we shall refer to this work to use a few arguments thereby developed.

3.1. **Basic Gradient Hölder continuity estimates.** Theorem 3.1 is basically a consequence of a series of intermediate lemmas allowing to reduce the oscillations of Dw when shrinking intrinsic cylinders. In this section w denotes a solution to (3.1) in a cylinder of the type $Q_r^\lambda \equiv Q$, enjoying the regularity indicated in (3.2). Moreover, as already observed in Remark 3.1, due to a standard approximation procedure in this Section we may assume that the equation in (3.1) is nondegenerate, that is, (3.7). In the following we shall use the standard notation

$$\|v\|_{V^2(Q)}^2 := \sup_{t_1<t<t_2} \int_B |v(x,t)|^2\,dx + \int_Q |Dv(x,t)|^2\,dx\,dt \tag{3.8}$$

whenever we are considering a cylinder of the type $Q = B \times (t_1, t_2)$. The space $V^2(Q)$ is the defined by all those $L^2(t_1,t_2;W^{1,2}(B))$ functions v such that the previous quantity is finite. Moreover we denote $V_0^2(Q) = V^2(Q) \cap L^2(t_1,t_2;W_0^{1,2}(B))$. The following Poincaré type inequality is then classical (see [10, Chapter 1, Corollary 3.1]):

$$\|v\|_{L^2(Q_1)}^2 \leq c(n)|\{|v|>0\}\cap Q_1|^{2/(n+2)}\|v\|_{V^2(Q_1)}^2 \tag{3.9}$$

and holds for all functions $v \in V_0^2(Q_1)$, where $Q_1 = B_1 \times (-1,0)$.

Proposition 3.1. *Assume that*

$$s + \sup_{Q_r^\lambda} \|Dw\| \leq A\lambda \tag{3.10}$$

holds for some constant $A \geq 1$. There exists a number $\sigma \equiv \sigma(n,p,\nu,L,A) \in (0,1/2)$ such that if

$$\frac{|\{(x,t)\in Q_r^\lambda \,:\, w_{x_i}(x,t) < \lambda/2\}|}{|Q_r^\lambda|} \leq \sigma \tag{3.11}$$

holds for some $i \in \{1,\dots,n\}$, then

$$w_{x_i} \geq \frac{\lambda}{4} \quad \textit{a.e. in } Q_{r/2}^\lambda\,.$$

Proof. Step 1: Rescaling. Without loss of generality we shall assume that the vertex of the cylinder coincides with the origin. We now make the standard intrinsic scaling by defining

$$v(x,t) := \frac{w(rx,\lambda^{2-p}r^2t)}{r}\,, \qquad (x,t)\in Q_1 \tag{3.12}$$

so that the newly defined function v solves

$$\lambda^{p-2}v_t - \operatorname{div} a(Dv) = 0\,. \tag{3.13}$$

From now on all the estimates will be recast in terms of the function v. Notice that with the new definition we still have

$$s + \|Dv\|_{L^\infty(Q_1)} \leq A\lambda \tag{3.14}$$

and assumption (3.11) translates into

$$|\{(x,t)\in Q_1 \,:\, v_{x_i}(x,t) < \lambda/2\}| \leq \sigma|Q_1|\,.$$

Our next aim is to show that

$$v_{x_i} \geq \frac{\lambda}{4} \qquad \text{a.e. in } Q_{1/2}\,. \tag{3.15}$$

The statement of the Proposition will then follow by scaling back to w.

Step 2: Iteration. In the following we shall proceed formally, all the details can be justified using Steklov averages [10]. We start by differentiating equation (3.13) in the x_i-direction; this is possible since (3.7) is in force and it turns out that

$$Dv \in L^2_{\mathrm{loc}}(-1,0;W^{1,2}_{\mathrm{loc}}(B,\mathbb{R}^n)) \cap C^0(-1,0;L^2_{\mathrm{loc}}(B_1,\mathbb{R}^n))\,. \tag{3.16}$$

The details can be found in [10, Chapter 8, Section 3]. Therefore, we obtain that v_{x_i} solves the following linear parabolic equation:

$$\lambda^{p-2}(v_{x_i})_t - \operatorname{div} \tilde{A}(x,t)Dv_{x_i} = 0\,, \qquad \text{where } \tilde{A}(x,t) := \partial a(Dv(x,t))\,. \tag{3.17}$$

The standard Caccioppoli's inequality for linear parabolic equations is now

$$\begin{aligned} &\sup_{-1<t<0} \lambda^{p-2} \int_{B_1} (v_{x_i}-k)_-^2 \eta^2(x,t)\,dx \\ &\quad + \int_{Q_1} (|Dv|^2+s^2)^{\frac{p-2}{2}} |D(v_{x_i}-k)_-|^2 \eta^2\,dx\,dt \\ &\qquad \leq c\lambda^{p-2} \int_{Q_1} (v_{x_i}-k)_-^2 \eta|\eta_t|\,dx\,dt \\ &\qquad\quad + c\int_{Q_1} (|Dv|^2+s^2)^{\frac{p-2}{2}} (v_{x_i}-k)_-^2 |D\eta|^2\,dx\,dt \\ &\qquad\quad + c\lambda^{p-2} \int_{B_1} (v_{x_i}-k)_-^2 \eta^2(x,-1)\,dx \end{aligned} \tag{3.18}$$

for a constant c depending only on n, p, ν, L; here $k \geq 0$ and $\eta \in C^\infty(Q_1)$ is a nonnegative cut-off function which vanishes on the lateral boundary of Q_1. Estimate (3.18) can be obtained by testing (3.17) with $(v_{x_i}-k)_-\eta^2$, and then arguing exactly as in [10, Chapter 2, Proposition 3.1]; it is necessary to observe here that the following inequalities are satisfied for all $\xi \in \mathbb{R}^n$ by $\tilde{A}(x,t)$ as a consequence of (1.3):

$$\begin{cases} \nu(s^2+|Dv(x,t)|^2)^{(p-2)/2}|\xi|^2 \leq \langle \tilde{A}(x,t)\xi,\xi\rangle \\ \quad |\tilde{A}(x,t)| \leq L(s^2+|Dv(x,t)|^2)^{(p-1)/2}\,. \end{cases} \tag{3.19}$$

We now let $k_0 = \lambda/2$ and for any integer $m \geq 0$ we define

$$k_m := k_0 - \frac{H}{8(1+A)}\left(1-\frac{1}{2^m}\right)\,, \qquad H := \sup_{Q_1}\,(v_{x_i}-k_0)_-\,.$$

Obviously $\{k_m\}$ is a decreasing sequence. For later convenience we also define the nonnegative cut-off function $\eta_m \in C^\infty(Q_m)$, where

$$Q_m := Q_{\varrho_m} \qquad \varrho_m := \frac{1}{2}+\frac{1}{2^{m+1}}\,, \qquad m \geq 0\,,$$

and in such a way that

$$0 \leq \eta_m \leq 1\,, \quad |D\eta_m|^2 + |(\eta_m)_t| \leq c(n)4^m\,, \quad \eta_m \equiv 1 \quad \text{on } Q_{m+1}\,. \tag{3.20}$$

Of course η_m is such that it vanishes outside Q_m and continuously on the parabolic boundary of Q_m. Notice that $Q_{\varrho_0} = Q_1$ and $Q_m \to Q_{1/2}$. Let us preliminary observe that we may assume that

$$4H \geq \lambda\,. \tag{3.21}$$

Indeed, we would otherwise have $H < \lambda/4$ that means

$$\sup_{Q_1} (v_{x_i} - k_0)_- = \lambda/2 - \inf_{Q_1} v_{x_i} < \frac{\lambda}{4}\,,$$

which immediately implies (3.15). Therefore we can assume that (3.21) holds. Moreover, by (3.21), we notice that for every $m \geq 1$ it holds that

$$k_m - k_{m+1} = \frac{H}{2^{m+4}(1+A)} \geq \frac{\lambda}{2^{m+6}(1+A)}\,, \tag{3.22}$$

and

$$k_m \geq \frac{\lambda}{4}\,, \qquad\qquad k_m \to k_\infty := k_0 - \frac{H}{8(1+A)} > \frac{\lambda}{4}\,. \tag{3.23}$$

Indeed, observe that (3.14) implies

$$\frac{H}{8(1+A)} \leq \frac{(\lambda/2 + A\lambda)}{8(1+A)} < \frac{\lambda}{4}\,.$$

Now, let us set, again for $m \geq 0$

$$A_m := \{(x,t) \in Q_m \,:\, v_{x_i} < k_m\}\,,$$

and define the truncated function

$$\tilde{v}_m := \begin{cases} 0 & \text{if} \quad v_{x_i} > k_m \\ k_m - v_{x_i} & \text{if} \quad k_m \geq v_{x_i} > k_{m+1} \\ k_m - k_{m+1} & \text{if} \quad v_{x_i} \leq k_{m+1}\,. \end{cases}$$

We have, as $\eta_m \equiv 1$ on Q_{m+1}, that

$$\begin{aligned} \lambda^{p-2}(k_m - k_{m+1})^2 |A_{m+1}| &= \lambda^{p-2} \|\tilde{v}_m\|^2_{L^2(A_{m+1})} \\ &\leq \lambda^{p-2} \|\tilde{v}_m\|^2_{L^2(Q_{m+1})} \\ &\leq \lambda^{p-2} \|\tilde{v}_m \eta_m\|^2_{L^2(Q_m)} \\ &\leq c\lambda^{p-2} \|\tilde{v}_m \eta_m\|^2_{V^2(Q_m)} |A_m|^{2/(n+2)}\,. \end{aligned} \tag{3.24}$$

In the last line we have applied inequality (3.9) to the function $\tilde{v}_m\eta_m$, which obviously non-negative; in this respect notice that

$$|\{\tilde{v}_m\eta_m > 0\} \cap Q_m| \leq |\{\tilde{v}_m > 0\} \cap Q_m| = |A_m|\,.$$

Then, observing that

$$\tilde{v}_m \leq (v_{x_i} - k_m)_-\,, \qquad\qquad |D\tilde{v}_m| \leq |D(v_{x_i} - k_m)_-| \chi_{Q_1 \setminus \{v_{x_i} < k_{m+1}\}}\,,$$

where $\chi_{Q_1 \setminus \{v_{x_i} < k_{m+1}\}}$ denotes the characteristic function of the set $Q_1 \setminus \{v_{x_i} < k_{m+1}\}$, we have, using the definition in (3.8), that

$$\begin{aligned} \lambda^{p-2} \|\tilde{v}_m\eta_m\|^2_{V^2(Q_m)} &\leq \lambda^{p-2} \|\tilde{v}_m\eta_m\|^2_{V^2(Q_1)} \\ &\leq \sup_{-1<t<0} \lambda^{p-2} \int_{B_1} (v_{x_i} - k_m)^2_- \eta_m^2(x,t)\, dx \\ &\qquad + \lambda^{p-2} \int_{Q_1} |D(v_{x_i} - k_m)_-|^2 \chi_{Q_1 \setminus \{v_{x_i} < k_{m+1}\}} \eta_m^2 \, dx\, dt \\ &\qquad + \lambda^{p-2} \int_{Q_1} (v_{x_i} - k_m)^2_- |D\eta_m|^2 \, dx\, dt\,. \end{aligned}$$

Now, notice that by (3.23) we have

$$\lambda \leq 4k_{m+1} \leq 4v_{x_i} \leq 4|Dv| \qquad \text{in} \quad Q_1 \setminus \{v_{x_i} < k_{m+1}\}$$

and, consequently,

$$
\begin{aligned}
\lambda^{p-2}\|\tilde{v}_m\eta_m\|^2_{V^2(Q_m)} \quad \leq \quad & \sup_{-1<t<0} \lambda^{p-2}\int_{B_1} (v_{x_i}-k_m)^2_-\eta_m^2(x,t)\,dx \\
& +c\int_{Q_1} (|Dv|^2+s^2)^{\frac{p-2}{2}}|D(v_{x_i}-k_m)_-|^2\eta_m^2\,dx\,dt \\
& +c\lambda^{p-2}\int_{Q_1}(v_{x_i}-k_m)^2_-|D\eta_m|^2\,dx\,dt\,.
\end{aligned}
$$

Combining the last inequality with (3.18) (obviously written with $k\equiv k_m$ and $\eta\equiv\eta_m$), since η_m is supported in Q_m, we also infer

$$
\begin{aligned}
\lambda^{p-2}\|\tilde{v}_m\eta_m\|^2_{V^2(Q_m)} \quad \leq \quad & c\lambda^{p-2}\int_{Q_1} (v_{x_i}-k_m)^2_-\eta_m|(\eta_m)_t|\,dx\,dt \\
& +c\int_{Q_1} (|Dv|^2+s^2)^{\frac{p-2}{2}}(v_{x_i}-k_m)^2_-|D\eta_m|^2\,dx\,dt \\
& +c\lambda^{p-2}\int_{Q_1}(v_{x_i}-k_m)^2_-|D\eta_m|^2\,dx\,dt
\end{aligned}
$$

and finally, using (3.14) and (3.20)

$$\lambda^{p-2}\|\tilde{v}_m\eta_m\|^2_{V^2(Q_m)} \leq cA^p4^m\lambda^p|A_m|\,.$$

This last inequality and (3.24) now give

$$\lambda^{p-2}(k_m-k_{m+1})^2|A_{m+1}| \leq c4^m\lambda^p|A_m|^{1+2/(n+2)}\,,$$

where $c\equiv c(n,p,\nu,L,A)$. Yet using (3.22) gives

$$|A_{m+1}| \leq c10^m|A_m|^{1+2/(n+2)}\,,$$

for every $m\geq 1$, for a constant c still depending only on n,p,ν,L,A. At this stage by using a standard iteration lemma [10, Chapter 1, Lemma 4.2] we have that there exists a number $\sigma\equiv\sigma(n,p,\nu,L,A)\in(0,1)$ such that if

$$|\{(x,t)\in Q_1\,:\,v_{x_i}(x,t)<\lambda/2\}| = |A_0| \leq \sigma|Q_1|\,,$$

then $|A_m|\to 0$ and this implies (3.15) by (3.23). The proof is complete. □

The dual version of the previous result is

Proposition 3.2. *Assume that* (3.10) *holds for some constant* $A\geq 1$. *There exists a number* $\sigma\equiv\sigma(n,p,\nu,L,A)\in(0,1/2)$ *such that if*

$$\frac{|\{(x,t)\in Q_r^\lambda\,:\,w_{x_i}(x,t)>-\lambda/2\}|}{|Q_r^\lambda|}\leq\sigma \tag{3.25}$$

holds for some $i\in\{1,\ldots,n\}$, *then*

$$w_{x_i}\leq-\frac{\lambda}{4}\quad a.e.\ in\ Q^\lambda_{r/2}\,.$$

Proof. Define $\tilde{w}:=-w$ and observe that this solves the equation $\tilde{w}_t-\operatorname{div}\tilde{a}(D\tilde{w})=0$, where $\tilde{a}(z):=-a(-z)$. Since the vector field $\tilde{a}(\cdot)$ still satisfies assumptions (1.3), we can then obtain Proposition 3.2 by simply applying Proposition 3.1 to $\tilde{w}$. Needless to say, a direct proof completely similar to the one of Proposition 3.1 is possible as well. □

Lemma 3.1. *Let* $\tilde{v}\in L^2(-1,0;W^{1,2}(B_1))$ *be a weak solution to the linear parabolic equation*

$$\tilde{v}_t-\operatorname{div}\,(B(x,t)D\tilde{v})=0\,,$$

where the matrix $B(x,t)$ has bounded and elliptic measurable entries, i.e.

$$\nu_0|\xi|^2 \leq \langle B(x,t)\xi,\xi\rangle\,, \qquad\qquad |B(x,t)| \leq L_0$$

hold whenever $\xi \in \mathbb{R}^n$, where $0<\nu_0 \leq L_0$ are fixed constant. Then there exists a constant $c_1 \equiv c_1(n,\nu_0,L_0) \geq 1$ such that

$$\sup_{Q_{1/2}} |\tilde v| \leq c_1 \left(\fint_{Q_1} |\tilde v|^q \, dx\, dt\right)^{1/q} \qquad\qquad q \in [1,2] \tag{3.26}$$

and further two constants $c_2 \equiv c_2(n,\nu_0,L_0) \geq 1$ and $\beta \equiv \beta(n,\nu_0,L_0) \in (0,1)$ such that

$$\left(\fint_{Q_\delta} |\tilde v - (\tilde v)_{Q_\delta}|^q \, dx\, dt\right)^{1/q} \leq c_2 \delta^\beta \left(\fint_{Q_1} |\tilde v - (\tilde v)_{Q_1}|^q \, dx\, dt\right)^{1/q}$$

holds whenever $q \in [1,2]$ and $\delta \in (0,1)$. The above inequalities still holds for $q \in (0,1)$, with additional dependence of the constants upon q.

Proof. The proof follows the one of [12, Proposition 4.1], where theorem is proved for the "worst possible case" $q = 1$ (the case $q = 2$ being the standard one). The proof given in [12] adapts to the case $q \in [1,2]$ in a straightforward way. Note that the constants involved are independent of q as we are assuming that this varies in a compact interval which stays bounded from zero. Again following [12, Proposition 4.1] it can be observed that the inequalities stated in the Lemma hold for $q \in (0,2]$, but the resulting constant c depends on q and blows up when $q \to 0$. □

Lemma 3.2. *Assume that in the cylinder Q_r^λ it holds that*

$$0 < \lambda/4 \leq ||Dw(x,t)|| \leq s + ||Dw(x,t)|| \leq A\lambda \qquad \forall\ (x,t) \in Q_r^\lambda\,, \tag{3.27}$$

where $A \geq 1$. Then there exist constants $\beta \in (0,1)$ and $c \geq 1$, both depending only on n,p,ν,L,A, such that

$$\left(\fint_{Q_{\delta r}^\lambda} |Dw - (Dw)_{Q_{\delta r}^\lambda}|^q \, dx\, dt\right)^{1/q} \leq c\delta^\beta \left(\fint_{Q_r^\lambda} |Dw - (Dw)_{Q_r^\lambda}|^q \, dx\, dt\right)^{1/q} \tag{3.28}$$

holds whenever $\delta \in (0,1)$ and $q \geq 1$.

Proof. Let us observe that by standard manipulations it is sufficient to prove the statement for the case $\delta \in (0,1/2)$; from this case the full one $\delta \in (0,1)$ follows after standard manipulations (see Proposition 3.3 below). We start as for the proof of Proposition 3.1 and rescale everything in the cylinder Q_1 as in (3.12), thereby getting a solution v in Q_1 to the equation (3.13). Moreover, by (3.27) it holds that

$$0 < \lambda/4 \leq ||Dv(x,t)|| \leq s + ||Dv(x,t)|| \leq A\lambda \qquad \forall\ (x,t) \in Q_1\,. \tag{3.29}$$

Then we differentiate (3.13), thereby obtaining (3.17). Therefore, dividing (3.17) by λ^{p-2} we see that each component v_{x_i} solves

$$(v_{x_i})_t - \operatorname{div}\,(B(x,t)Dv_{x_i}) = 0\,, \qquad\qquad B(x,t) := \lambda^{2-p}\tilde A(x,t)\,. \tag{3.30}$$

By virtue of (3.19) and (3.29) the matrix $B(x,t)$ is uniformly elliptic in the sense that

$$c^{-1}|\xi|^2 \leq \langle B(x,t)\xi,\xi\rangle \leq c\left(\frac{s+\lambda}{\lambda}\right)^{p-2} |\xi|^2 \leq c|\xi|^2 \tag{3.31}$$

holds whenever $\xi \in \mathbb{R}^n$, where $c \equiv c(n,p,\nu,L,A) \geq 1$. We end the proof by showing that there exist constants $\beta \in (0,1)$ and $c \geq 1$, both depending only on n,p,ν,L,A, such that for every $q \geq 1$

$$\left(\fint_{Q_\delta} |v_{x_i} - (v_{x_i})_{Q_\delta}|^q \, dx\, dt\right)^{1/q} \leq c\delta^\beta \left(\fint_{Q_1} |v_{x_i} - (v_{x_i})_{Q_1}|^q \, dx\, dt\right)^{1/q} \tag{3.32}$$

holds whenever $\delta \in (0,1/2)$ and $i \in \{1,\dots,n\}$. The case $q \in [1,2]$ is a direct consequence of Lemma 3.1. For the case $q > 2$ we instead argue as follows. Observing that $v_{x_i} - (v_{x_i})_{Q_\delta}$ is still a solution to (3.30), by (3.26) of Lemma 3.1 for $\delta \in (0,1)$ we have

$$\begin{aligned}\left(\fint_{Q_{\delta/2}} |v_{x_i} - (v_{x_i})_{Q_{\delta/2}}|^q \,dx\,dt\right)^{1/q} &\leq 2\left(\fint_{Q_{\delta/2}} |v_{x_i} - (v_{x_i})_{Q_\delta}|^q \,dx\,dt\right)^{1/q}\\ &\leq 2\sup_{Q_{\delta/2}} |v_{x_i} - (v_{x_i})_{Q_\delta}|\\ &\leq c\left(\fint_{Q_\delta} |v_{x_i} - (v_{x_i})_{Q_\delta}|^2 \,dx\,dt\right)^{1/2} \qquad (3.33)\end{aligned}$$

with $c \equiv c(n,p,\nu,L,A) \geq 1$. Applying (3.32) with $q = 2$ and Hölder's inequality, we have

$$\begin{aligned}\left(\fint_{Q_{\delta/2}} |v_{x_i} - (v_{x_i})_{Q_{\delta/2}}|^q \,dx\,dt\right)^{1/q} &\leq c\left(\fint_{Q_\delta} |v_{x_i} - (v_{x_i})_{Q_\delta}|^2 \,dx\,dt\right)^{1/2}\\ &\leq c\delta^\beta\left(\fint_{Q_1} |v_{x_i} - (v_{x_i})_{Q_1}|^2 \,dx\,dt\right)^{1/2}\\ &\leq c\delta^\beta\left(\fint_{Q_1} |v_{x_i} - (v_{x_i})_{Q_1}|^q \,dx\,dt\right)^{1/q} \qquad (3.34)\end{aligned}$$

from which (3.32) actually follows for $\delta \in (0,1/2)$; by scaling back to w this implies (3.28) for $\delta \in (0,1/2)$, as the index i is arbitrary. Finally, as observed at the beginning of the proof, if (3.28) holds whenever $\delta \in (0,1/2)$, added then it also holds for $\delta \in (0,1)$ (modulo enlarging the constant c of a factor depending on n) and the proof of (3.28) is finished. □

Summarizing the previous results yields

Proposition 3.3. *Assume that* (3.10) *is in force. There exists a positive number* $\sigma \equiv \sigma(n,p,\nu,L,A) \in (0,1/2)$ *such that if there exists* $i \in \{1,\dots,n\}$ *for which either* (3.11) *or* (3.25) *holds, then*

$$\left(\fint_{Q^\lambda_{\delta r}} |Dw - (Dw)_{Q^\lambda_{\delta r}}|^q \,dx\,dt\right)^{1/q} \leq c_d\delta^\beta\left(\fint_{Q^\lambda_r} |Dw - (Dw)_{Q^\lambda_r}|^q \,dx\,dt\right)^{1/q} \qquad (3.35)$$

holds whenever $\delta \in (0,1)$ *for constants* $\beta \equiv \beta(n,p,\nu,L,A) \in (0,1)$ *and* $c_d \equiv c_d(n,p,\nu,L,A) \geq 1$. *Moreover, it holds that*

$$\|Dw\| \geq \frac{\lambda}{4} \qquad \text{a.e. in } Q^\lambda_{r/2}\,. \qquad (3.36)$$

Proof. If there exists $i \in \{1,\dots,n\}$ for which either (3.11) or (3.25) holds then Proposition 3.1 or 3.2 applies and hence (3.36) follows immediately. We can therefore apply Lemma 3.2 (in the cylinder $Q^\lambda_{r/2}$). As an outcome we get

$$\left(\fint_{Q^\lambda_{\delta r/2}} |Dw - (Dw)_{Q^\lambda_{\delta r/2}}|^q \,dx\,dt\right)^{1/q} \leq c\delta^\beta\left(\fint_{Q^\lambda_{r/2}} |Dw - (Dw)_{Q^\lambda_{r/2}}|^q \,dx\,dt\right)^{1/q}$$

whenever $\delta \in (0,1)$ and with the dependences of the constants specified in Lemma 3.2. In turn we have that

$$\left(\fint_{Q^\lambda_{r/2}} |Dw - (Dw)_{Q^\lambda_{r/2}}|^q \,dx\,dt\right)^{1/q} \leq 2\left(\fint_{Q^\lambda_{r/2}} |Dw - (Dw)_{Q^\lambda_r}|^q \,dx\,dt\right)^{1/q}$$

$$\leq 2^{(n+2)/q+1}\left(\fint_{Q_r^\lambda}|Dw-(Dw)_{Q_r^\lambda}|^q\,dx\,dt\right)^{1/q}.$$

This means we have that (3.35) holds for $\delta\in(0,1/2)$. Finally to show that (3.35) holds for $\delta\in[1/2,1)$ one may proceed as in the last group of inequalities, enlarging again the constant of a factor $2^{(n+2)/q+1}$ and the proof is complete. □

The next step deals with a degenerate behavior and analyzes the case ruled out by the previous Proposition 3.3.

Proposition 3.4. *Assume that* (3.10) *holds, while neither* (3.11) *nor* (3.25) *hold for any* $i\in\{1,\dots,n\}$. *Then it is possible to find* $\sigma_1\in(0,1)$ *and* $\eta\in(1/2,1)$, *depending only on* n,p,ν,L,A, *such that*

$$\|Dw\|\leq \eta A\lambda \qquad \text{a.e. in } Q_{\sigma_1 r}^\lambda\,. \tag{3.37}$$

Proof. As usual, we assume without loss of generality that the vertex of cylinder Q_r^λ is the origin. The proof closely follows the one for [10, Chapter 9, Proposition 1.2]. We therefore ask the reader to keep track of the various parts of the proof of [10, Chapter 9, Proposition 1.2] since we shall only report the main modifications. We divide the rest of the proof in three steps.

Step 1: Rescaling from a good instant. Assume now that none of the conditions in (3.11) and (3.25) holds; therefore for every $i\in\{1,\dots,n\}$ it happens that

$$\frac{|\{(x,t)\in Q_r^\lambda\,:\,w_{x_i}(x,t)\geq\lambda/2\}|}{|Q_r^\lambda|}<1-\sigma \tag{3.38}$$

and

$$\frac{|\{(x,t)\in Q_r^\lambda\,:\,w_{x_i}(x,t)\leq-\lambda/2\}|}{|Q_r^\lambda|}<1-\sigma\,. \tag{3.39}$$

We fix one index $i\in\{1,\dots,n\}$ and argue for w_{x_i} referring to (3.38); the same reasoning will obviously work for the other gradient components. Later on we shall give the modification necessary for dealing with condition (3.39). Now, proceeding exactly as in [10, Chapter 9, Lemma 12.1] we have that there exists a time level t^* (the "good instant")

$$-\lambda^{2-p}r^2\leq t^*\leq-\frac{\sigma}{2}\lambda^{2-p}r^2 \tag{3.40}$$

such that

$$\frac{|\{x\in B_r\,:\,w_{x_i}(x,t^*)\geq\lambda/2\}|}{|B_r|}\leq\frac{1-\sigma}{1-\sigma/2}\,.$$

By scaling we define the new function

$$v(x,t):=\frac{w(rx,-t^*t)}{rA\lambda}\,,\qquad\qquad (x,t)\in Q_1\,,$$

so that

$$\frac{s}{A\lambda}+\sup_{Q_1}\|Dv\|\leq 1 \tag{3.41}$$

holds and moreover

$$\frac{|\{x\in B_1\,:\,v_{x_i}(x,-1)\geq 1/(2A)\}|}{|B_1|}\leq\frac{1-\sigma}{1-\sigma/2}\,.$$

A fortiori, as $\sigma<1/2$ and $A\geq1$, we also have

$$\frac{|\{x\in B_1\,:\,v_{x_i}(x,-1)\geq(1-\sigma)\}|}{|B_1|}\leq\frac{1-\sigma}{1-\sigma/2}\,. \tag{3.42}$$

Let us now define

$$\nu^* := \frac{-t^*}{(A\lambda)^{2-p} r^2}\,, \tag{3.43}$$

so that (3.40) yields

$$\sigma A^{p-2}/2 \le \nu^* \le A^{p-2}\,. \tag{3.44}$$

We moreover define the new vector field

$$\tilde{a}(z) := \frac{\nu^* a(A\lambda z)}{(A\lambda)^{p-1}}\,.$$

It is now straightforward to check that $\tilde{v}$ weakly solves the equation

$$v_t - \operatorname{div} \tilde{a}(Dv) = 0 \qquad \text{in} \quad Q_1\,. \tag{3.45}$$

Moreover the vector field $\tilde{a}(\cdot)$ satisfies the following ellipticity and growth assumptions whenever $z, \xi \in \mathbb{R}^n$:

$$\begin{cases} |\tilde{a}(z)| + |\partial \tilde{a}(z)| \left(|z|^2 + \left(\frac{s}{A\lambda}\right)^2\right)^{1/2} \le \nu^* L \left(|z|^2 + \left(\frac{s}{A\lambda}\right)^2\right)^{(p-1)/2} \\ \nu^* \nu \left(|z|^2 + \left(\frac{s}{A\lambda}\right)^2\right)^{(p-2)/2} |\xi|^2 \le \langle \partial \tilde{a}(z)\xi, \xi\rangle\,. \end{cases}$$

Step 2: Switch to a nondegenerate regime. Again following [10, Chapter 9, Section 1.2] we now want to prove that there exists a number $\eta \in (1/2, 1)$, which can be determined as a functions of the parameters n, p, ν, L, A, such that

$$|\{(x,t) \in Q_{1/2} \,:\, v_{x_i}(x,t) > \eta\}| = 0\,. \tag{3.46}$$

Scaling back to w by (3.40) this in turn implies

$$|\{(x,t) \in Q^{\lambda}_{\sigma r/4} \,:\, w_{x_i}(x,t) > \eta A\lambda\}| = 0\,. \tag{3.47}$$

In order to prove (3.46) we will re-exploit the strategy for the proof of the similar statement in [10, Chapter 9, Theorem 12.1], which is in turn based on the use of logarithmic inequalities and De Giorgi type iterations. More precisely, we again differentiate equation (3.45) in the x_i-direction, thereby obtaining

$$(v_{x_i})_t - \operatorname{div} \tilde{A}(x,t) D v_{x_i} = 0\,, \qquad \text{where } \tilde{A}(x,t) := \partial \tilde{a}(Dv(x,t))\,. \tag{3.48}$$

The simple remark is to observe that in order to prove (3.46) one may proceed exactly as in [10, Chapter 9, Theorem 12.1], with $z(x,t)$ replaced by $v_{x_i}(x,t)$, and with equation (3.48) replacing the differential inequality $z_t - \operatorname{div} A_*(x,t) z_{x_i} \le 0$ which is [10, Chapter 9, (12.4)]. The only difference is that while in [10] the matrix considered is already uniformly elliptic, the matrix $\tilde{A}(x,t)$ we are considering in (3.48) is not (in the sense that the lower bound on the first eigenvalue depends on s). This point can be anyway easily seen to be inessential when proving (3.46) and it is in fact possible to adopt here exactly the same treatment proposed in [10]. Indeed, it is sufficient to note that the test functions needed to follow the same arguments of [10] are essentially

$$\Psi((v_{x_i} - k)_+) \qquad \text{and} \qquad (v_{x_i} - k)_+\,, \qquad k \ge 1/4\,, \tag{3.49}$$

where Ψ is the standard logarithmic function defined in [10, Chapter 2, (3.12)]; see also (3.51) below. See in particular the choice in [10, Chapter 9, Lemma 12.2]. In other words all the resulting integral inequalities are supported in sets contained in $\{|Dv| \ge 1/4\}$. Observe now that on such sets the matrix $\tilde{A}(x,t)$ becomes uniformly elliptic as

$$\frac{\nu^*}{16^{(p-2)/2}} \quad \le \quad \nu^* \left(|Dv|^2 + \left(\frac{s}{A\lambda}\right)^2\right)^{(p-2)/2}$$

$$\leq \quad 4n^{(p-2)/2}\nu^* \qquad \text{on } \{|Dv| \geq 1/4\}\,. \tag{3.50}$$

Here we also used (3.41) and (3.3). Therefore we may proceed as in the case of standard quadratic equations - exactly as in [10, Chapter 9, Sections 12-13] - getting the same inequalities of [10] and thereby proving (3.46). It is here important to note that all the constants involved depend only on n, p, ν, L, A since the number ν^* appearing in (3.43) and (3.50) depends in turn only on the number σ determined in Proposition 3.3, which in turn depends only on n, p, ν, L, A. For the reader's convenience we give a short road map to the proof, recalling the main steps in [10, Chapter 9, Sections 12-13]. Inequality (3.42) is exactly the same relation presented in [10, Chapter 9, (12.6)], therefore, considering as in [10, Chapter 9, Lemma 12.2] the logarithmic function

$$\Psi(v_{x_i}) := \log^+\left[\frac{\sigma}{\sigma - (v_{x_j} - (1-\sigma))_+ + \eta_0}\right] \tag{3.51}$$

with $0 < \eta_0 < \sigma$, we gain the following inequality:

$$\int_{B_r\times\{t\}} \Psi^2(v_{x_i})\,dx \leq \int_{B_1\times\{-1\}} \Psi^2(v_{x_i})\,dx + \frac{c}{(1-r)^2}\int_{Q_1} \Psi(v_{x_i})\,dx\,dt$$

for every $r \in (0,1)$ and every $t \in (-1,0)$. The constant c depends only on n, p, ν, L, ν^* via (3.50) and ultimately via n, p, ν, L, A by (3.44). The last inequality is indeed exactly [10, Chapter 9, (12.7)]. We remark that this is a crucial point where we are using the fact that, thanks to (3.50), the equation (3.48) becomes uniformly parabolic when using the one in (3.51) as test function (compare [10, Chapter 2, Proposition 3.2]). At this point we proceed as in [10, Chapter 9, Lemma 13.1] thereby obtaining that for a proper choice of $\eta_0 \equiv \eta_0(n, p, \nu, L, A) \in (0, \sigma)$ we have

$$\frac{|\{x \in B_1 \,:\, v_{x_i}(x,t) > (1-\eta_0)\}|}{|B_1|} \leq 1 - \frac{\sigma^2}{4}\,, \qquad \forall\, t \in (-1,0)\,.$$

Finally, from this inequality we can proceed in the proof of [10, Chapter 9, Lemma 13.2, Theorem 12.1], that eventually leads to (3.46). Notice that also this proof is based on the use of test functions as the second one appearing in (3.49), and therefore the equation (3.48) becomes uniformly parabolic on the supports selected by such a truncation test function; recall (3.50).

Step 3: Final size reduction. Observe now that we have proved (3.47) but we also need to prove that

$$|\{(x,t) \in Q^\lambda_{\sigma r/4} \,:\, w_{x_i}(x,t) < -\eta A\lambda\}| = 0 \tag{3.52}$$

to conclude with (3.37). This can be easily observed since we also know that (3.39) holds for all $i \in \{1,\ldots,n\}$. At this point we may reduce to the case already treated in Steps 1 and 2 exactly as in Proposition 3.2, that is by passing to $-w_{x_i}$ and reducing to the case (3.38). Finally, all in all we have shown that both (3.47) and (3.52) hold for every choice of $i \in \{1,\ldots,n\}$ and this implies (3.37) with $\sigma_1 = \sigma/4$. Observe that $\sigma \equiv \sigma(n, p, \nu, L, A)$ is the number determined in Propositions 3.1-3.2 and later occurring in Proposition 3.3; the proof is complete. □

We conclude this section with a result that will come into the play when proving that estimates are independent of s.

Proposition 3.5. *Assume that*

$$\sup_{Q^\lambda_r} \|Dw\| \leq A\lambda \qquad \textit{and} \qquad \gamma\lambda \leq s \leq \gamma_1 A\lambda \qquad \textit{where } 0 < \gamma \leq \gamma_1\,. \tag{3.53}$$

Then

$$\left(\fint_{Q^\lambda_{\delta r}} |Dw-(Dw)_{Q^\lambda_{\delta r}}|^q\,dx\,dt\right)^{1/q} \leq \tilde{c}_d\delta^{\beta_1}\left(\fint_{Q^\lambda_{r}} |Dw-(Dw)_{Q^\lambda_{r}}|^q\,dx\,dt\right)^{1/q} \tag{3.54}$$

holds whenever $\delta \in (0,1)$ *for constants* $\beta_1 \equiv \beta_1(n,p,\nu,L,A,\gamma,\gamma_1) \in (0,1)$ *and* $\tilde{c}_d \equiv \tilde{c}_d(n,p,\nu,L,A,\gamma,\gamma_1) \geq 1$.

Proof. We again start as for the proof of Proposition 3.1 and rescale everything in the cylinder Q_1 as in (3.12), thereby getting a solution v in Q_1 to the equation (3.13). Then we differentiate the equation thereby obtaining (3.17). By dividing (3.17) by λ^{p-2} we see that each component v_{x_i} solves (3.30). This time we use (3.53) to deduce (3.31) for $c \equiv c(n,p,\nu,L,A,\gamma)$:

$$c^{-1}|\xi|^2 \equiv c^{-1}\gamma^{p-2}|\xi|^2 \leq \langle B(x,t)\xi,\xi\rangle \leq c(1+\gamma_1)^{p-2}|\xi|^2 \equiv c|\xi|^2\,.$$

The rest of the proof follows exactly as in Lemma 3.2. □

3.2. **Alternatives and Iteration.** In this section we streamline the results of the previous one and organize them in a way that will be useful for the next developments. When considering an intrinsic cylinder of the type Q^λ_r, i.e. a cylinder such that

$$s + \sup_{Q^\lambda_r} ||Dw|| \leq A\lambda\,, \tag{3.55}$$

by Propositions 3.3 and 3.4 we have then two possibilities:

- **The Nondegenerate Alternative**. This means that we can apply Proposition 3.3 and therefore we have that

$$\left(\fint_{Q^\lambda_{\delta r}} |Dw-(Dw)_{Q^\lambda_{\delta r}}|^q\,dx\,dt\right)^{1/q} \leq c_d\delta^{\beta}\left(\fint_{Q^\lambda_{r}} |Dw-(Dw)_{Q^\lambda_{r}}|^q\,dx\,dt\right)^{1/q} \tag{3.56}$$

  for every $\delta \in (0,1)$, where the constants $\beta \equiv \beta(n,p,\nu,L,A) \in (0,1)$ and $c_d \equiv c_d(n,p,\nu,L,A) \geq 1$ are those defined in Proposition 3.2
- **The Degenerate Alternative**. In this case we can instead apply Proposition 3.4 and we reduce the size of the gradient in a suitable inner cylinder

$$\sup_{Q^\lambda_{\sigma_1 r}} ||Dw|| \leq \eta A\lambda\,, \qquad \eta \equiv \eta(n,p,\nu,L,A) \in (0,1)$$

  where $\sigma_1 \equiv \sigma_1(n,p,\nu,L,A) \in (0,1)$

The Degenerate and the Nondegenerate alternative can be combined to obtain what we shall call the Degenerate Iteration. This describes a situation when the Degenerate Alternative holds a certain number of times when considering a suitable chain of shrinking intrinsic cylinders, and therefore the size of the gradient exhibits a geometric decay. For technical reasons we shall choose a slightly worse decay parameter $\eta_1 \in (\eta,1)$ rather that η.

The Degenerate Iteration. By starting with a condition as (3.55) in an intrinsic cylinder Q^λ_r, following [10] we consider the number $\eta \equiv \eta(n,p,\nu,L,A) \in (0,1)$ defined in Proposition 3.4 and then define

$$\eta_1 := \frac{1+\eta}{2}\,, \qquad \text{so that } \eta<\eta_1<1 \quad \text{and} \quad \eta_1-\eta = \frac{1-\eta}{2}\,. \tag{3.57}$$

Obviously $\eta_1 \equiv \eta_1(n,p,\nu,L,A) \in (0,1)$. We define the sequences

$$\begin{cases} \lambda_{j+1} := \eta_1\lambda_j \\ \lambda_0 := \lambda\,, \end{cases} \qquad \begin{cases} R_{j+1} := c_0 R_j \\ R_0 := r\,. \end{cases}$$

The number $c_0 \in (0,1)$ is defined via the numbers σ_1 appearing in the Degenerate Alternative, and via η_1 defined in (3.57), as follows:

$$c_0 := \frac{\sigma_1 \eta_1^{(p-2)/2}}{2} \in (0, 1/2)$$

so that c_0 is a quantity depending only on n, p, ν, L, A. With such a choice the following inclusions hold:

$$Q^{\lambda}_{R_{j+1}} \subset Q^{\lambda_{j+1}}_{R_{j+1}} \subset Q^{\lambda_j}_{\sigma_1 R_j} \subset Q^{\lambda_j}_{R_j} \subset Q^{\lambda}_{r}\,, \qquad \forall\, j \in \mathbb{N}\,. \tag{3.58}$$

Here, as in the following, all the cylinders share the same vertex. Assume that the Degenerate Alternative holds for the initial cylinder $Q^{\lambda}_{R_0} \equiv Q^{\lambda}_{r}$, then we have

$$\sup_{Q^{\lambda_1}_{R_1}} ||Dw|| \leq \sup_{Q^{\lambda_0}_{\sigma_1 R_0}} ||Dw|| \leq \eta A\lambda \leq \eta_1 A\lambda = A\lambda_1\,.$$

We are therefore led to consider the same situation in the cylinder $Q^{\lambda_1}_{R_1}$; assume now that

$$s + \sup_{Q^{\lambda_1}_{R_1}} ||Dw|| \leq A\lambda_1$$

holds and that again the Degenerate Alternative occurs. We then have

$$\sup_{Q^{\lambda_2}_{R_2}} ||Dw|| \leq \sup_{Q^{\lambda_1}_{\sigma_1 R_1}} ||Dw|| \leq A\lambda_2 = \eta_1 A\lambda_1\,.$$

Proceeding in a similar fashion, and assuming that the Degenerate Alternative can be applied m times and that

$$s + \sup_{Q^{\lambda_j}_{R_j}} ||Dw|| \leq A\lambda_j \tag{3.59}$$

holds, if then Proposition 3.4 can be applied on $Q^{\lambda_j}_{R_j}$, we have

$$\sup_{Q^{\lambda_{j+1}}_{R_{j+1}}} ||Dw|| \leq \sup_{Q^{\lambda_j}_{\sigma_1 R_j}} ||Dw|| \leq \eta_1^{j+1} A\lambda = A\lambda_{j+1}\,. \tag{3.60}$$

In particular, by (3.58) we get

$$\sup_{Q^{\lambda}_{R_{j+1}}} ||Dw|| \leq \eta_1^{j+1} A\lambda\,.$$

From now on we shall denote

$$Q^{\lambda}_{j} \equiv Q^{\lambda}_{R_j}$$

and in the next section, when proving Theorem 3.1, we shall see that a lower bound of the type in (3.4) prevents the Degenerate Alternative to occur more than a finite number of times, say $\tilde{m}$, *which can be quantitatively determined* in terms of the given parameters n, p, ν, L, A and B. Therefore, after $\tilde{m}$ steps of Degenerate Iteration, the Nondegenerate Alternative will hold and consequently (3.5) will follow, with a constant c depending on $\tilde{m}$, and therefore ultimately on n, p, ν, L, A, B.

3.3. **Proof of Theorem 3.1.** For the proof of Theorem 3.1 we need the notation and the content of the *Degenerate Iteration* described in the previous section. As specified in Remark 3.1 the proof below formally covers the case $s = 0$ but is actually valid only in the case $s > 0$, since it relies on the results in the previous sections, indeed proved under this additional assumption. Since all the estimates are independent of s the case $s = 0$ can be eventually reached by a simple approximation argument - see Section 3.6 below.

Step 1: The Degenerate Iteration always stops after a controlled number of steps. With $\eta_1 \equiv \eta_1(n,p,\nu,L,A) \in (0,1)$ being defined in (3.57) to build the degenerate iteration, we define $m \in \mathbb{N}$ as the smallest integer such that

$$\eta_1^m A\lambda < \frac{\lambda}{2B}\,. \tag{3.61}$$

Observe that this determines $m \geq 1$ as a function of the parameters n,p,ν,L,A,B. Now, consider a number $\underline{\delta} \leq c_0^{m+1}$ so that

$$Q^\lambda_{\underline{\delta}r} \subset Q^\lambda_{m+1} \tag{3.62}$$

and assume that

$$\frac{\lambda}{B} \leq \sup_{Q^\lambda_{\underline{\delta}r}} ||Dw||\,. \tag{3.63}$$

This in turn implies

$$A\lambda_m \equiv \eta_1^m A\lambda < \lambda/(2B) \leq \sup_{Q^\lambda_{\underline{\delta}r}} ||Dw|| \leq \sup_{Q^\lambda_{m+1}} ||Dw||\,. \tag{3.64}$$

Let us define

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

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

$$\tilde{m} \leq m\,. \tag{3.65}$$

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

$$\sup_{Q^\lambda_{m+1}} ||Dw|| \leq \sup_{Q^{\lambda m+1}_{R_{m+1}}} ||Dw|| \leq \eta_1^{m+1} A\lambda\,,$$

contradicting (3.64). Thus (3.65) holds. From now on we will look for a number $\delta_\varepsilon = \underline{\delta}$, which is smaller or equal than c_0^{m+1}.

Step 2: The first nondegenerate case. In this Step we assume that the first stopping time of the Degenerate Iteration, that is $\tilde{m}$, satisfies $\tilde{m} \leq m$, where m has been defined in (3.61). Now we analyze the situation at level $\tilde{m}$, and in particular the reasons why the degenerate alternative cannot occur. There are basically two reasons for the Degenerate Iteration to stop. The first is when (3.59) is not satisfied, i.e.

$$s + \sup_{Q^{\lambda_{\tilde{m}}}_{R_{\tilde{m}}}} ||Dw|| > A\lambda_{\tilde{m}} = \eta_1^{\tilde{m}} A\lambda \tag{3.66}$$

and therefore we cannot even try to verify the Degenerate Alternative, that requires (3.66) as a preliminary starting condition. In this Step we analyze this case. Note that by (3.4), $\tilde{m} \geq 1$. Note also that since the Degenerate Alternative holds at level $\tilde{m}-1$, we have

$$\sup_{Q^{\lambda_{\tilde{m}}}_{R_{\tilde{m}}}} ||Dw|| \leq \eta A\lambda_{\tilde{m}-1} = \eta\eta_1^{\tilde{m}-1} A\lambda\,. \tag{3.67}$$

Comparing (3.66) and (3.67) yields

$$s > \eta_1^{\tilde{m}} A\lambda - \sup_{Q^{\lambda_{\tilde{m}}}_{R_{\tilde{m}}}} ||Dw|| \geq \eta_1^{\tilde{m}-1}(\eta_1 - \eta) A\lambda \geq \gamma A\lambda \geq \gamma\lambda\,, \tag{3.68}$$

where, since $\tilde{m} \leq m$, we can set

$$\gamma := \eta_1^{m-1}(\eta_1 - \eta) > 0\,. \tag{3.69}$$

Notice that $\gamma \equiv \gamma(n,p,\nu,L,A,B)$ as we have already seen that m depends only on n,p,ν,L,A,B; see also (3.57). By (3.4) and (3.68) we have

$$\sup_{Q_r^\lambda} ||Dw|| \le A\lambda \qquad \text{and} \qquad \gamma\lambda \le s \le A\lambda$$

and we can apply Proposition 3.5 with the choice of γ made in (3.69) and $\gamma_1 = 1$, directly in the starting cylinder Q_r^λ. Therefore if

$$\delta \le \left(\frac{\varepsilon}{\tilde{c}_d}\right)^{1/\beta_1}, \tag{3.70}$$

we have

$$\left(\fint_{Q_{\delta r}^\lambda} |Dw - (Dw)_{Q_{\delta r}^\lambda}|^q \, dx\, dt\right)^{1/q} \le \varepsilon \left(\fint_{Q_r^\lambda} |Dw - (Dw)_{Q_r^\lambda}|^q \, dx\, dt\right)^{1/q}. \tag{3.71}$$

Step 3: The second nondegenerate case. In this Step we again assume that the first stopping time of the Degenerate Iteration, that is $\tilde{m}$, satisfies $\tilde{m} \le m$, with m as in (3.61), but we consider the case which is complementary to the one of Step 2. This is when

$$s + \sup_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} ||Dw|| \le A\lambda_{\tilde{m}}$$

and the Nondegenerate Alternative holds in the cylinder $Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}$; observe that here it may happen that $\tilde{m} = 0$. We can therefore use (3.56) in such a cylinder. Let us define

$$\tilde{\delta}_\varepsilon := \tilde{\delta} c_0^m \quad \text{with} \quad \tilde{\delta} \le c_0\,. \tag{3.72}$$

The number $\tilde{\delta}$ will be fixed in a few lines, in a way that makes it depending on ε, and this justifies the notation in the line above. We observe that the following inclusions:

$$Q_{\tilde{\delta}_\varepsilon r}^\lambda = \tilde{\delta} c_0^{m-\tilde{m}} Q_{\tilde{m}}^\lambda \subset \tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}} \subset Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}} \subset Q_r^\lambda \tag{3.73}$$

hold as a consequence of (3.58) and (3.72). Therefore

$$\begin{aligned}
&\left(\fint_{Q_{\tilde{\delta}_\varepsilon r}^\lambda} |Dw - (Dw)_{Q_{\tilde{\delta}_\varepsilon r}^\lambda}|^q \, dx\, dt\right)^{1/q} \\
&\qquad \le 2\left(\fint_{Q_{\tilde{\delta}_\varepsilon r}^\lambda} |Dw - (Dw)_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx\, dt\right)^{1/q} \\
&\qquad \le c\left(\frac{|\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}|}{|Q_{\tilde{\delta}_\varepsilon r}^\lambda|} \fint_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx\, dt\right)^{1/q} \\
&\qquad = c\left(\frac{|Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}|}{|Q_{\tilde{m}}^\lambda|} \fint_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx\, dt\right)^{1/q}.
\end{aligned}$$

On the other hand, using (3.56) with $\delta = \tilde{\delta} c_0^{m-\tilde{m}}$ and in the cylinder $Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}$, and keeping again (3.73) in mind, we have

$$\begin{aligned}
&\left(\fint_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{\tilde{\delta} c_0^{m-\tilde{m}} Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx\, dt\right)^{1/q} \\
&\qquad \le c(\tilde{\delta} c_0^{m-\tilde{m}})^\beta \left(\fint_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx\, dt\right)^{1/q}
\end{aligned}$$

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

where in the last estimate we used that $c_0 \leq 1$ and that we are assuming $\tilde m \leq m$. Connecting the last two groups of inequalities and continuing with the estimate, and again keeping (3.73) in mind, we have

$$\begin{aligned}
&\left(\fint_{Q^{\lambda}_{\tilde\delta_\varepsilon r}} |Dw-(Dw)_{Q^{\lambda}_{\tilde\delta_\varepsilon r}}|^q\,dx\,dt\right)^{1/q}\\
&\leq c\tilde{\delta}^{\beta}\left(\frac{|Q^{\lambda_{\tilde m}}_{R_{\tilde m}}|}{|Q^{\lambda}_{\tilde m}|}\fint_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} |Dw-(Dw)_{Q^{\lambda}_{r}}|^q\,dx\,dt\right)^{1/q}\\
&\leq c\tilde{\delta}^{\beta}\left(\frac{|Q^{\lambda}_{r}|}{|Q^{\lambda}_{\tilde m}|}\fint_{Q^{\lambda}_{r}} |Dw-(Dw)_{Q^{\lambda}_{r}}|^q\,dx\,dt\right)^{1/q}\\
&\leq \frac{\tilde c\tilde{\delta}^{\beta}}{c_0^{\tilde m(n+2)/q}}\left(\fint_{Q^{\lambda}_{r}} |Dw-(Dw)_{Q^{\lambda}_{r}}|^q\,dx\,dt\right)^{1/q}\\
&\leq \frac{\tilde c\tilde{\delta}^{\beta}}{c_0^{m(n+2)}}\left(\fint_{Q^{\lambda}_{r}} |Dw-(Dw)_{Q^{\lambda}_{r}}|^q\,dx\,dt\right)^{1/q},
\end{aligned}$$

where $\tilde c \equiv \tilde c(n,p,\nu,L,A)$. Notice that if we impose that

$$\tilde\delta \leq \left(\frac{c_0^{m(n+2)}\varepsilon}{\tilde c}\right)^{1/\beta}, \tag{3.74}$$

then we have

$$\fint_{Q^{\lambda}_{\tilde\delta_\varepsilon r}} |Dw-(Dw)_{Q^{\lambda}_{\tilde\delta_\varepsilon r}}|^q\,dx\,dt \;\leq \varepsilon^q \fint_{Q^{\lambda}_{r}} |Dw-(Dw)_{Q^{\lambda}_{r}}|^q\,dx\,dt\,. \tag{3.75}$$

Step 4: Determining the number δ_ε. By looking at conditions (3.70) and (3.74), we are led to define

$$\delta_\varepsilon := \tilde\delta c_0^m \quad \text{with} \quad \tilde\delta := \min\left\{\left(\frac{c_0^{m(n+2)}\varepsilon}{\tilde c}\right)^{1/\beta}, \left(\frac{\varepsilon}{\tilde c_d}\right)^{1/\beta_1}, c_0\right\} \tag{3.76}$$

and notice that both $\tilde\delta$ and δ_ε depend only on $n,p,\nu,L,A,B,\varepsilon$; moreover $\tilde\delta \leq c_0$. The number δ_ε defined in (3.76) is the one we are looking for and it does not depend on the solution w (neither on the vector field $a(\cdot)$) since it works both in Step 2 and Step 3 (therefore the choice does not depend on the reason why the Degenerate Iteration stopped, a fact that could imply a subtle dependence on the solution considered). Indeed, notice that since $\delta_\varepsilon \leq c_0^{m+1}$ then δ_ε perfectly works in Step 1 and therefore, assuming the first bound in (3.4) with this choice of δ_ε, the estimate on the stopping time (3.65) holds and we can consider Step 2 and Step 3. Then, if the case of Step 2 occurs then (3.5) follows from (3.71) with $\delta = \delta_\varepsilon$. On the other hand, if the case of Step 3 occurs, we notice that (3.5) follows from (3.75) with $\tilde\delta_\varepsilon = \delta_\varepsilon$. The proof of Theorem 3.1 is complete.

3.4. **Spatial gradient Hölder continuity.** This section is confined to observe that the arguments of the previous sections imply the Hölder continuity of the spatial gradient of general homogeneous equations.

Theorem 3.2. *Let w be a weak solution to* (3.1) *in a given cylinder Q. Then Dw is locally Hölder continuous in Q. Moreover, let $Q_r^\lambda \subset Q$ be an intrinsic cylinder such that*

$$s + \sup_{Q_r^\lambda} ||Dw|| \leq A\lambda \tag{3.77}$$

holds for a certain constant $A \geq 1$. Then

$$|Dw(x,t) - Dw(x_1,t_1)| \leq c_h \lambda \left(\frac{\varrho}{r}\right)^\alpha \tag{3.78}$$

holds whenever $(x,t),(x_1,t_1) \in Q_\varrho^\lambda$ for a constants $c_h \equiv c_h(n,p,\nu,L,A) \geq 1$ and $\alpha \equiv \alpha(n,p,\nu,L,A) \in (0,1)$ which is independent of s, of the solution w considered and of the vector field $a(\cdot)$. Here $Q_\varrho^\lambda \subset Q_r^\lambda$ are intrinsic cylinders sharing the same vertex.

Proof. As in the case of Theorem 3.1, we shall give a proof which formally includes the case $s = 0$ but that it is only valid for the one in which $s > 0$; the case $s = 0$ can be again reached by approximation. This approach is particularly useful here as the proof is a close revisitation of the one given by DiBenedetto in [10, Chapter 9, Section 3]. The case $s = 0$ is completely the same while the case $s > 0$ needs a few additional arguments; actually, we shall only deal with the case $s > 0$, as observed in Remark 3.1, but to stay closer to the presentation in [10], we shall also deal with the case $s = 0$, therefore assuming that the Degenerate and Nondegenerate alternative are valid in this case (following the approach in [10]).

Case $s = 0$. In the case $s = 0$ the proof is exactly the same of [10, Chapter 9, Section 3]. One first proves [10, Chapter 9, Lemma 2.1] using the alternatives described in Section 3.2 together with the Degenerate Iteration, and then the rest of the proof follows (here the lemma uses E_q instead of E_2, but it is the same; the reader may also uses directly E_2 if he/she likes to follow [10] more closely). For more on the proof of [10, Chapter 9, Lemma 2.1] see also the Case $s > 0$ below. Once [10, Chapter 9, Section 3] has been proved one may proceed exactly as in [10, Chapter 9, Lemma 3.1 & Lemma 3.2]. The only observation to make is that [10, Chapter 9, Lemma 3.1 & Lemma 3.2] perfectly work leading to inequality (3.78) without requiring that $(x,t),(x_0,t_0) \in K \subset Q_r^\lambda$ where K is a compact subset such that $\operatorname{dist}(K, Q_r^\lambda) > 0$. It is indeed sufficient to consider the situation where we replace K by the cylinder Q_ϱ^λ, which shares the same vertex with Q_r^λ, as in fact in Theorem 3.2 (observe that, when applying the arguments in [10], we here take $\Omega_T = Q_r^\lambda$ thanks to the starting assumption (3.77)). We do no think it is the case to report the full proof of Theorem 3.2 since it is rather long and at this point completely similar to the one explained in [10, Chapter 9, Section 3]. We observe that the stability of the constants when $p \to 2$ follows essentially from two points: first, the proof of the alternatives we give in Sections 3.1 and 3.2 does not depend on linearization methods, but it is rather direct and does only the quadratic structure of the differentiated equation (3.13) (after scaling). Second: all the choices of the constants in the subsequents parts of the proofs in can be made uniform when $p > 2$ moves in a small neighborhood of 2. In particular, the choice of β in [10, page 253], must be replaced by $\alpha_0/[N(p-2)+1]$.

Case $s > 0$. This case, which is nondegenerate, is not considered in [10], where DiBenedetto confined himself to the (important) model case (1.33). The difference with the previous case is that here the Degenerate Iteration described in Section 3.2 always stops after a finite number of steps since $s > 0$ and condition (3.59) cannot obviously verified for every λ_m as $\lambda_m \to 0$; since the number of steps the Degenerate Iteration works clearly depends on s, we have therefore to be careful since we want estimates that are independent of $s > 0$. Basically, this is the only missing part to reproduce DiBenedetto's arguments is [10, Chapter 9, Lemma 2.1], that we shall readapt now. We therefore ask the reader to keep track of the proof of [10, Chapter 9, Lemma 2.1]. Let us define

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

This number is called n_0 in [10] and R_{n_0} is called *the switching radius.* Observe that in the case $s = 0$ it may happen that $\tilde{m} = \infty$ (as also treated in [10]). Now, the main point in proving the analog of [10, Chapter 9, Lemma 2.1] is in showing that

$$\sup_{Q_{R_i}^{\lambda_i}} ||Dw|| \leq A\lambda_i \qquad\qquad \forall \; i \leq \tilde{m}$$

and then, when $\tilde{m}$ is finite, the decay estimate

$$\fint_{Q_{\delta R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{Q_{\delta R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx \, dt \leq c^q \delta^{\beta q} \fint_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} |Dw - (Dw)_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}}|^q \, dx \, dt \tag{3.79}$$

holds whenever $\delta \in (0, 1)$, where $c \geq 1$ and $\beta \in (0, 1)$ depend only on n, p, ν, L, A. Notice that a small difference here is given by the fact that in [10] it is $A = 1$ and $q = 2$, but this makes no essential problem, as it only affects the constants appearing in the statement of Theorem 3.2, that indeed depend on A.

In the case $\tilde{m} = 0$ there is nothing to say, since the Nondegenerate Alternative occurs at the very first moment and then one proceed as in [10, Chapter 9, Section 2]; more precisely (3.79) follows by (3.56) applied in the starting cylinder Q_r^λ. So we argue on the case $\tilde{m} \geq 1$. Now, there are basically two reasons why the Degenerate Iteration stops at $\tilde{m}$. The first is when the starting condition (3.59), i.e.

$$s + \sup_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} ||Dw|| \leq A\lambda_{\tilde{m}} \, ,$$

is in force but then the Nondegenerate Alternative holds. In this case notice that (3.79) follows by (3.56) (that is Proposition 3.3), applied with the choice $Q_r^\lambda \equiv Q_{R_m}^{\lambda_{\tilde{m}}}$. Estimate (3.79) is exactly [10, Chapter 9, (2.5)], where $n_0 = \tilde{m}$. At this point [10, Chapter 9, Lemma 2.1] follows and it is possible to proceed exactly as in [10], with the remarks made above for the case $s = 0$.

The other reason why the Degenerate Alternative stops at step $\tilde{m}$ is that the starting condition (3.59) is simply not satisfied at step $\tilde{m}$ since s is large, that is

$$s + \sup_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} ||Dw|| > A\lambda_{\tilde{m}} = \eta_1^{\tilde{m}} A\lambda \, . \tag{3.80}$$

In this case we cannot proceed with the Nondegenerate Alternative, and therefore we rather proceed in a different way as the equation is automatically nondegenerate on this scale. Indeed, notice that since the Degenerate Alternative held at level $\tilde{m}-1$ then we have

$$\sup_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} ||Dw|| \leq \eta A\lambda_{\tilde{m}-1} = \eta\eta_1^{\tilde{m}-1} A\lambda \leq A\lambda_{\tilde{m}} \, .$$

Comparing the last inequality with (3.80) yields

$$s > \eta_1^{\tilde{m}} A\lambda - \sup_{Q_{R_{\tilde{m}}}^{\lambda_{\tilde{m}}}} ||Dw|| \geq (\eta_1 - \eta)\eta_1^{\tilde{m}-1} A\lambda =: \gamma A\lambda_{\tilde{m}-1} \geq \gamma\lambda_{\tilde{m}}$$

where $\gamma = \eta_1 - \eta$ (see also (3.57)) and therefore $\gamma \equiv \gamma(n, p, \nu, L, A) > 0$. On the other hand, since the Degenerate Alternative held at level $\tilde{m} - 1$, the starting condition (3.59) must hold at level $\tilde m - 1$ and therefore $s \le A\lambda_{\tilde m-1} = (A/\eta_1)\lambda_{\tilde m}$. Summarizing, letting $\gamma_1 := A/\eta_1$, we have

$$\sup_{Q^{\lambda_{\tilde m}}_{R_{\tilde m}}} ||Dw|| \le A\lambda_{\tilde m} \qquad \text{and} \qquad \gamma\lambda_{\tilde m} \le s \le \gamma_1 A\lambda_{\tilde m}$$

and we can therefore apply Proposition 3.5 on the cylinder $Q^{\lambda_{\tilde m}}_{R_{\tilde m}}$, with the choice of γ, γ_1 made here; notice that both γ and γ_1 depend only on n, p, ν, L, A. In turn this yields (3.79) and therefore the analog of Lemma [10, Chapter 9, Lemma 2.1]. Again, the rest of the proof follows as in [10], in the case $s = 0$. □

Corollary 3.1. *Under the assumptions and notations of Theorem 3.2 it holds that*

$$|V(Dw(x,t)) - V(Dw(x_1,t_1))| \le c_v\lambda^{p/2}\left(\frac{\varrho}{r}\right)^{\alpha}.$$

Proof. Let us preliminary observe that (3.77) implies

$$s + ||Dw||_{L^\infty(Q^\lambda_r)} \le c(n,p,\nu,L,A)\lambda\,.$$

Therefore, by using [35, (2.2)] and (3.78), it follows that

$$\begin{aligned}
&|V(Dw(x,t)) - V(Dw(x_1,t_1))| \\
&\qquad \le c(s + |Dw(x,t)| + |Dw(x_1,t_1)|)^{(p-2)/2}|Dw(x,t) - Dw(x_1,t_1)| \\
&\qquad \le c(s+\lambda)^{(p-2)/2}\lambda(\varrho/R)^{\alpha}\,.
\end{aligned}$$

The statement follows using (3.78) again. □

Remark 3.2. The dependence on A of the constant c_h appearing in (3.78) is linear, i.e. $c_h = \tilde c_h A$, where $\tilde c_h$ depends only on n, p, ν, L.

Remark 3.3. The statement about the local Hölder continuity of Du of Theorem 3.2 remains valid for solutions to non-homogeneous equations as

$$w_t - \operatorname{div} a(Dw) = g \in L^\infty\,. \tag{3.81}$$

See for instance [11, 25, 28]. This remark is important when considering the additional regularity hypotheses made for instance in Theorems 1.1-1.3, and in particular, the one of the a priori continuity of the gradient Du, as discussed in Sections 1.4 and 4.3. We also remark that Theorem 3.1 continues to hold in the case of solutions to the p-Laplacean system, as it will be shown, starting from the techniques introduced here, in the forthcoming paper [28]; moreover, the singular case $1 < p < 2$ is demonstrated in [25].

3.5. **Further a priori estimates for homogeneous equations.** The following result is taken from [10, Chapter 8, Theorem 5.1], and in the form suitable for general equations can be retrieved from [24].

Theorem 3.3. *Suppose that w is a weak solution to* (3.1) *in Q^λ_r, $\lambda, r > 0$. Then there exists a constant $c_3 \ge 1$, depending only on n, p, ν, L, but otherwise independent of s, of the solution w considered and of the vector field $a(\cdot)$, such that*

$$\sup_{\frac12 Q^\lambda_r} ||Dw|| \le c_3\lambda + c_3\lambda^{2-p}\fint_{Q^\lambda_r}(|Dw|+s)^{p-1}\,dx\,dt\,.$$

Consequently, if

$$\fint_{Q^\lambda_r}(|Dw|+s)^{p-1}\,dx\,dt \le \lambda^{p-1}$$

then

$$\sup_{\frac12 Q^\lambda_r} ||Dw|| \le 2c_3\lambda\,.$$

3.6. **The approximation scheme.** The approximation method needed to assume $s > 0$ in the proof of Theorems 3.1 and 3.2 is at the center of discussion here; the general scheme is completely standard, but the occurrence of a few subtle differences might nevertheless deserve some explanation in the case of Theorem 3.1. Therefore in the following we take for granted Theorem 3.1 in the case $s > 0$ and we prove it for the case $s = 0$; the only point will be that, when passing from the case $s > 0$ to the case $s = 0$ the constants will increase, but in a universal way, that is depending only on the parameters already appearing in the statement of Theorem 3.1. Now we proceed with the approximation. We start mollifying the vector field $a(\cdot)$ as follows. Let $\sigma > 0$ (actually denoting a sequence converging to zero) and let $\theta_\sigma \in C_0^\infty(B_\sigma(0))$ be a standard mollifier with $B_\sigma(0) \subset \mathbb{R}^n$, such that $\int_{\mathbb{R}^n} \theta_\sigma(z)\, dz = 1$. Define

$$a_\sigma(z) := \int_{\mathbb{R}^n} \theta_\sigma(z-\xi)a(\xi)\, d\xi.$$

The we define w_σ as the unique solution to the following Cauchy-Dirichlet problem:

$$\begin{cases} (w_\sigma)_t - \operatorname{div} a_\sigma(Dw_\sigma) = 0 & \text{in } Q_r^\lambda \\ w_\sigma = w & \text{on } \partial_{\rm par} Q_r^\lambda\,. \end{cases} \tag{3.82}$$

Exactly as for instance in [24] - but this is actually standard - it follows that $a_\sigma(\cdot)$ satisfies (1.3) with new constants ν, L and with s replaced now by $s_\sigma = \sigma$; without loss of generality we shall consider σ small enough to have $s_\sigma \leq \lambda$. Again as in [24] it follows that up to not relabeled subsequences (i.e. we still keep the notation $_\sigma$)

$$\begin{cases} Dw_\sigma \to Dw \quad \text{strongly in } L^p \text{ and a.e.} \\ \displaystyle -\!\!\!\!\!\!\int_{Q_r^\lambda} (|Dw_\sigma| + s_\sigma)^p\, dx\, dt \leq c_a^p -\!\!\!\!\!\!\int_{Q_r^\lambda} (|Dw| + s_\sigma)^p\, dx\, dt\,, \end{cases} \tag{3.83}$$

where c_a depends only on n, p, ν, L. Before going on let us recall a basic result (see see [10, Chapter 8, Theorem 5.1]) asserting

$$-\!\!\!\!\!\!\int_{Q_r^\lambda} (|Dw_\sigma| + s_\sigma)^p\, dx\, dt \leq c^p\lambda^p \Longrightarrow \sup_{Q_{r/2}^\lambda} \|Dw_\sigma\| \leq \tilde{c}_a c\lambda \tag{3.84}$$

for a new constant $\tilde{c}_a \equiv \tilde{c}_a(n, p, \nu, L)$. Now we assume that the Theorem 3.1 holds for the case $s > 0$, and fix A, B, ε in the "$s = 0$"-version of Theorem 3.1 we want to prove. Take the choice

$$\varepsilon \to 2^{-(n+2)}\varepsilon =: \tilde{\varepsilon}\,, \qquad A \to 2\tilde{c}_a c_a A =: \tilde{A}\,, \qquad B \to 2B$$

and determine the number $\delta_{\tilde{\varepsilon}}(\tilde{A})$ in Theorem 3.1 for the case $s > 0$ (remark that $\delta_{\tilde{\varepsilon}}(\tilde{A})$ also depends on n, p, ν, L via the new constants in (1.3) for $a_\sigma(\cdot)$; it of course also depends on B). We claim that now Theorem 3.1 for the case $s = 0$ holds with the choice

$$\delta_\varepsilon(A) := \delta_{\tilde{\varepsilon}}(\tilde{A})/2\,, \tag{3.85}$$

and indeed the assumptions in question are now

$$\frac{\lambda}{B} \leq \sup_{Q_{\delta_{\tilde{\varepsilon}} r/2}^\lambda} \|Dw\| \leq \sup_{Q_r^\lambda} \|Dw\| \leq \tilde{A}\lambda\,. \tag{3.86}$$

Let us now observe that for yet another not relabelled subsequence we may assume that

$$\frac{\lambda}{2B} \leq \sup_{Q_{\delta_{\tilde{\varepsilon}} r/2}^\lambda} \|Dw_\sigma\|\,.$$

Indeed, were this not the case, by using the convergence in (3.83) we would immediately contradict the first inequality in (3.86). On the other hand, thanks to (3.83)-(3.84) it follows

$$s_\sigma + \sup_{Q^\lambda_{r/2}} \|Dw_\sigma\| \leq \tilde{A}\lambda\,. \tag{3.87}$$

We can therefore apply Theorem 3.1 in the case $s \equiv s_\sigma > 0$ thereby obtaining

$$E_q(Dw_\sigma, (\delta_{\tilde{\varepsilon}}/2)Q^\lambda_r) \leq 2^{-(n+2)}\varepsilon E_q(Dw_\sigma, Q^\lambda_{r/2})$$

letting $\sigma \to 0$, (3.83) and (3.87) yield

$$E_q(Dw, (\delta_{\tilde{\varepsilon}}/2)Q^\lambda_r) \leq 2^{-(n+2)}\varepsilon E_q(Dw, Q^\lambda_{r/2}) \leq \varepsilon E_q(Dw, Q^\lambda_r)$$

and this proves Theorem 3.1 in the case $s = 0$, with the choice in (3.85). Finally, the approximation argument to deduce Theorem 3.2 in the case $s = 0$ from the case $s > 0$ is completely standard, and follows along the lines of the one for Theorem 3.1.

4. Proof of the intrinsic potential estimate

In this section we give the proof Theorem 1.1, which in turn implies Theorems 1.2 and 1.3 as seen in the Introduction. First, in Section 4.1 we propose a few comparison estimates necessary to implement the iteration procedure that will lead, in Section 4.2 below, to the proof of the intrinsic potential estimate (1.16).

4.1. **Comparison results.** In the rest of the section we consider in a fixed parabolic cylinder $Q \equiv Q^\lambda_\rho(x_0, t_0) \subseteq \Omega_T$ the unique solution

$$w \in C^0(t_0 - \lambda^{2-p}\varrho^2, t_0; L^2(B(x_0, \varrho))) \cap L^p(t_0 - \lambda^{2-p}\varrho^2, t_0; W^{1,p}(B(x_0, \varrho))) \tag{4.1}$$

to the following Cauchy-Dirichlet problem:

$$\begin{cases} w_t - \operatorname{div} a(Dw) = 0 & \text{in } Q^\lambda_\varrho \\ w = u & \text{on } \partial_{\mathrm{par}} Q^\lambda_\varrho\,. \end{cases} \tag{4.2}$$

Then we establish a comparison estimate between u and w in the next

Lemma 4.1. *Let u be as in Theorem 1.1 and w as in* (4.2). *Let q be such that*

$$0 < q < p - 1 + \frac{1}{n+1}\,. \tag{4.3}$$

Then there exists a constant c depending only on n, p, ν, q such that it holds that

$$\begin{aligned} &\left(\fint_Q \left(|Du - Dw|^q + |V(Du) - V(Dw)|^{2q/p} \right) dx\, dt \right)^{1/q} \\ &\qquad \leq c \left(\frac{|\mu|(Q)}{|Q|^{(n+1)/(n+2)}} \right)^{(n+2)/[(p-1)n+p]}\,. \end{aligned} \tag{4.4}$$

Proof. As also earlier in the paper, the calculations below will be done on a formal level; they can be anyway made rigorous by a standard use of Steklov averages. This is precisely the point where (1.25) is required. We shall in the following denote $Q_t := B \times \{t\} \equiv B(x_0, \varrho) \times \{t\}$, whenever $t \in (-\lambda^{2-p}r^2, 0)$. In the rest of the proof, without loss of generality we shall assume $q \geq 1$ and that the vertex of the cylinder (x_0, t_0) coincides with the origin.

Step 1: Preliminary estimates. We will first prove that

$$\sup_\tau \int_{Q_\tau} |u - w|\, dx \leq |\mu|(Q)\,, \qquad -\lambda^{2-p}r^2 < \tau < 0 \tag{4.5}$$

and

$$\int_Q \frac{|V(Du)-V(Dw)|^2}{(\alpha+|u-w|)^\xi}\,dx\,dt \le c\frac{\alpha^{1-\xi}}{\xi-1}|\mu|(Q) \tag{4.6}$$

hold for $\alpha > 0$ and $\xi > 1$, where $c \equiv c(n,p,\nu) \ge 1$. For this, choose the test function(s)

$$\eta_{1,\varepsilon} = \pm\min\{1,(u-w)_\pm/\varepsilon\}\phi\,, \quad \varepsilon>0\,,$$

where $\phi \in C^\infty(\mathbb{R})$ is a nonincreasing function depending only on t, such that $0\le\phi\le 1$ and $\phi(t)=0$ for all $t\ge\tau$ for $\tau\in(-\lambda^{2-p}r^2,0)$; obviously $\phi_t\le 0$. In the following, when choosing the function ϕ according to our needs, we shall do it in a way that $\int_{\mathbb{R}}|\phi_t|\,dt=1$. A direct calculation gives

$$D\eta_{1,\varepsilon} = \frac{1}{\varepsilon}D(u-w)\chi_{\{0<(u-w)_\pm<\varepsilon\}}\phi\,.$$

Now we test subtracted equations of u and w with $\eta_{1,\varepsilon}$, thereby obtaining

$$\int_Q (u-w)_t\eta_{1,\varepsilon}\,dx\,dt + \int_Q \langle a(Du)-a(Dw), D\eta_{1,\varepsilon}\rangle\,dx\,dt = \int_Q \eta_{1,\varepsilon}\,d\mu\,. \tag{4.7}$$

Observe that since

$$\begin{aligned}(u-w)_t\min\{1,(u-w)_\pm/\varepsilon\} &= \partial_t\int_0^{u-w}\min\{1,s_\pm/\varepsilon\}\,ds\\ &= \pm\partial_t\int_0^{(u-w)_\pm}\min\{1,s/\varepsilon\}\,ds\,,\end{aligned}$$

integration by parts then yields

$$\int_Q (u-w)_t\eta_{1,\varepsilon}\,dx\,dt = \int_Q\int_0^{(u-w)_\pm}\min\{1,s/\varepsilon\}\,ds(-\phi_t)\,dx\,dt\,. \tag{4.8}$$

Thus it follows that

$$\begin{aligned}&\int_Q\int_0^{(u-w)_\pm}\min\{1,s/\varepsilon\}\,ds(-\phi_t)\,dx\,dt\\ &+\frac{1}{\varepsilon}\int_Q\langle a(Du)-a(Dw),D(u-w)\rangle\,\chi_{\{0<(u-w)_\pm<\varepsilon\}}\,dx\,dt \le |\mu|(Q)\,,\end{aligned} \tag{4.9}$$

where we used

$$\int_Q \eta_{1,\varepsilon}\,d\mu \le |\mu|(Q)\,.$$

Observe that both the terms in the left hand side of (4.9) are nonnegative by (2.8). Next, by the dominated convergence theorem we have

$$\int_Q\int_0^{(u-w)_\pm}\min\{1,s/\varepsilon\}\,ds(-\phi_t)\,dx\,dt \to \int_Q (u-w)_\pm(-\phi_t)\,dx\,dt$$

as $\varepsilon\to 0$. Using this information together with (4.9) and (2.8) yields

$$\int_Q |u-w|(-\phi_t)\,dx\,dt \le |\mu|(Q)\,.$$

Letting ϕ approximate the characteristic function of $(-\infty,\tau)$, taking any $\tau\in(-\lambda^{2-p}r^2,0)$, gives

$$\int_{Q_\tau}|u-w|\,dx \le |\mu|(Q)\,,$$

from which (4.5) follows. We also get that

$$\sup_{\varepsilon>0}\int_Q\langle a(Du)-a(Dw),D\eta_{1,\varepsilon}\rangle\,dx\,dt \le |\mu|(\Omega)\,. \tag{4.10}$$

We now test (4.7) with

$$\eta_{2,\varepsilon} = \frac{\eta_{1,\varepsilon}}{(\alpha + (u-w)_\pm)^{\xi-1}}\,, \qquad \xi > 1\,,\ \varepsilon, \alpha > 0\,,$$

to get

$$\int_Q (u-w)_t \eta_{2,\varepsilon}\,dx\,dt + \int_Q \langle a(Du) - a(Dw), D\eta_{2,\varepsilon}\rangle\,dx\,dt = \int_Q \eta_{2,\varepsilon}\,d\mu\,.$$

For the first term on the left we get by integration by parts - as for (4.8) - that

$$\int_Q (u-w)_t \eta_{2,\varepsilon}\,dx\,dt = \int_Q \int_0^{(u-w)_\pm} \frac{\min\{1, s/\varepsilon\}}{(\alpha+s)^{\xi-1}}\,ds(-\phi_t)\,dx\,dt\,.$$

Thus we arrive by (4.5) at

$$\sup_{\varepsilon>0} \int_Q (u-w)_t \eta_{2,\varepsilon}\,dx\,dt \le \alpha^{1-\xi} \sup_t \int_{Q_t} |u-w|\,dx \int_{\mathbb{R}} |\phi_t|\,dt \le \alpha^{1-\xi}|\mu|(\Omega)\,.$$

For the elliptic term we notice

$$\begin{aligned}
&\int_Q \langle a(Du) - a(Dw), D\eta_{2,\varepsilon}\rangle\,dx\,dt\\
&\quad = \int_Q \langle a(Du) - a(Dw), D\eta_{1,\varepsilon}\rangle \frac{1}{(\alpha + (u-w)_\pm)^{\xi-1}}\,dx\,dt\\
&\quad\quad + (1-\xi)\int_Q \langle a(Du) - a(Dw), D(u-w)_\pm\rangle \frac{\eta_{1,\varepsilon}}{(\alpha + (u-w)_\pm)^{\xi}}\,dx\,dt\,.
\end{aligned}$$

The first integral on the right can be majorized using (4.10) as

$$\begin{aligned}
&\int_Q \langle a(Du) - a(Dw), D\eta_{1,\varepsilon}\rangle \frac{1}{(\alpha + (u-w)_\pm)^{\xi-1}}\,dx\,dt\\
&\quad \le \alpha^{1-\xi} \sup_{\varepsilon>0} \int_\Omega \langle a(Du) - a(Dw), D\eta_{1,\varepsilon}\rangle\,dx\,dt \le \alpha^{1-\xi}|\mu|(Q)\,.
\end{aligned}$$

But since

$$\left|\int_Q \eta_{2,\varepsilon}\,d\mu\right| \le \alpha^{1-\xi}|\mu|(Q)\,,$$

we obtain

$$(\xi - 1)\int_Q \frac{\langle a(Du) - a(Dw), D(u-w)_\pm\rangle}{(\alpha + (u-w)_\pm)^{\xi}} \eta_{1,\varepsilon}\,dx\,dt \le 3\alpha^{1-\xi}|\mu|(Q)$$

and therefore, using (2.8) and the definition of $\eta_{1,\varepsilon}$ we obtain

$$\int_Q \frac{|V(Du) - V(Dw)|^2}{(\alpha + |u-w|)^{\xi}} \min\{1, |u-w|/\varepsilon\}\,dx\,dt \le \frac{c\alpha^{1-\xi}}{\xi-1}|\mu|(Q)\,.$$

Letting $\varepsilon \to 0$ yields (4.6).

Step 2: Comparison estimates. Fix now

$$\frac{\xi q}{p-q} = \frac{n+1}{n} q =: \tilde q \qquad \Longleftrightarrow \qquad \xi = \frac{n+1}{n}(p-q)\,,$$

so that $\xi > 1$ iff (4.3) holds. Define

$$\alpha = \left(\fint_Q |u-w|^{\tilde q}\,dx\,dt\right)^{1/\tilde q}$$

and assume that $\alpha > 0$, for if it is not, then $u = w$ and (4.4) follows. The parabolic Sobolev's inequality (see for instance [10, Chapter 1, Proposition 3.1]) yields

$$\alpha \le c(n,q)\left[\fint_Q |D(u-w)|^q\,dx\,dt \left(\sup_\tau \int_{Q_\tau} |u-w|\,dx\right)^{q/n}\right]^{n/[q(n+1)]}\,,$$

and thus by (4.5) that

$$
\begin{aligned}
\alpha &\leq c|\mu|(Q)^{1/(n+1)}\left(\fint_Q |D(u-w)|^q \,dx\,dt\right)^{n/(q(n+1))} \\
&\leq c|\mu|(Q)^{1/(n+1)}\left(\fint_Q |V(Du)-V(Dw)|^{2q/p} \,dx\,dt\right)^{n/(q(n+1))}. \qquad (4.11)
\end{aligned}
$$

Applying then Hölder's inequality, together with (4.6) and (4.11), we obtain

$$
\begin{aligned}
&\fint_Q |V(Du)-V(Dw)|^{2q/p} \,dx\,dt \\
&= \fint_Q \left(\frac{|V(Du)-V(Dw)|^2}{(\alpha+|u-w|)^\xi}\right)^{q/p} (\alpha+|u-w|)^{\xi q/p} \,dx\,dt \\
&\leq \left(\fint_Q \frac{|V(Du)-V(Dw)|^2}{(\alpha+|u-w|)^\xi} \,dx\,dt\right)^{q/p} \left(\fint_Q (\alpha+|u-w|)^{\xi q/(p-q)} \,dx\,dt\right)^{(p-q)/p} \\
&\leq c\left(\frac{|\mu|(Q)}{|Q|}\alpha^{1-\xi}\right)^{q/p} \alpha^{\xi q/p} \\
&\leq c\left(\frac{|\mu|(Q)^{(n+2)/(n+1)}}{|Q|}\left(\fint_Q |V(Du)-V(Dw)|^{2q/p} \,dx\,dt\right)^{n/(q(n+1))}\right)^{q/p}.
\end{aligned}
$$

But since

$$
1-\frac{n}{p(n+1)} = \frac{(p-1)n+p}{p(n+1)},
$$

we end up with (4.4) and the proof is complete after recalling the first inequality in (2.8). □

Corollary 4.1. *Let u and w be as in Lemma 4.1. Suppose that the intrinsic relation*

$$
\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)} \leq \lambda \qquad (4.12)
$$

is satisfied. Then there exists a constant $c_4 = c_4(n,p,\nu) \geq 1$ such that

$$
\left(\fint_{Q_\varrho^\lambda} |Du-Dw|^{p-1} \,dx\,dt\right)^{1/(p-1)} \leq c_4\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)} \qquad (4.13)
$$

and

$$
\left(\fint_{Q_\varrho^\lambda} |V(Du)-V(Dw)|^\gamma \,dx\,dt\right)^{1/\gamma} \leq c_4\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/\gamma} \qquad (4.14)
$$

hold, where

$$
\gamma := 2(p-1)/p \geq 1. \qquad (4.15)
$$

Proof. Write

$$
\begin{aligned}
\left(\frac{|\mu|(Q_\varrho^\lambda)}{|Q_\varrho^\lambda|^{(n+1)/(n+2)}}\right)^{\frac{n+2}{(p-1)n+p}} &= c(n,p)\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{(2-p)(n+1)/(n+2)}\varrho^{n+1}}\right)^{\frac{n+2}{(p-1)n+p}} \\
&= c(n,p)\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{\frac{n+2}{(p-1)n+p}} \lambda^{\frac{2-p}{(p-1)n+p}}.
\end{aligned}
$$

We now use (4.12) in the form

$$\lambda^{\frac{2-p}{(p-1)n+p}} \leq \left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{\frac{2-p}{[(p-1)n+p](p-1)}}$$

so that

$$\left(\frac{|\mu|(Q_\varrho^\lambda)}{|Q_\varrho^\lambda|^{(n+1)/(n+2)}}\right)^{\frac{n+2}{(p-1)n+p}} \leq c(n,p)\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{\frac{n+2}{(p-1)n+p}+\frac{2-p}{[(p-1)n+p](p-1)}} .$$

The proof of (4.13) and (4.14) now follows (eventually taking the largest constant) using the previous inequality together with Lemma 4.1 - used with the obvious choice $q = p - 1$ - and the identity

$$\frac{n+2}{(p-1)n+p} + \frac{2-p}{[(p-1)n+p](p-1)} = \frac{1}{p-1} .$$

□

Remark 4.1. The intrinsic bound in (4.12) reflects the fact that when switching to the intrinsic geometry, the gradient, which, as suggested by (1.16), dimensionally speaking is comparable to λ, scales according to the density of the intrinsic potential

$$\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}$$

exactly as it happens in the elliptic case [12].

Remark 4.2. The kind of "intrinsic comparison estimate" introduced in Lemma 4.1 and Corollary 4.1 eventually revealed to useful also in other contexts where integrability results for solutions are inferred from those of the assigned data μ; see for instance [1, 7].

4.2. **Proof of Theorem 1.1.** The proof goes in five steps. We shall use large demagnifying constants as $400, 800, 2400$ and the like, to emphasize the role of certain choices in the proof. In the rest of the proof, given a measurable vector valued map, typically a gradient, $g\colon Q \to \mathbb{R}^n$, where Q is a cylinder, we shall denote its L^{p-1}-excess functional in Q as

$$E(g,Q) := E_{p-1}(g,Q) = \left(\fint_Q |g-(g)_Q|^{p-1}\,dx\,dt\right)^{1/(p-1)} . \tag{4.16}$$

Compare with the definition in (3.6).

Step 1: Setting of the constants and basic inequalities. In the following all the cylinders will have (x_0,t_0) as vertex, therefore we shall as usual omit denoting the vertex simply writing $Q_\varrho^\lambda(x_0,t_0) \equiv Q_\varrho^\lambda$. We start taking λ of the form

$$\lambda := H_1\beta + H_2\int_0^{2r}\left(\frac{|\mu|(Q_\varrho^\lambda)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho} , \tag{4.17}$$

and fix the constants $H_1, H_2 \geq 1$ in due course of the proof, in a way that makes them depending only on n,p,ν,L; β is assumed to satisfy (1.15). At the end, when proving (1.16), we shall take $c := \max\{H_1,H_2\}$. We look at Theorem 3.3 and let

$$A := 10c_3 . \tag{4.18}$$

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

$$\varepsilon = \frac{1}{2^5 2^{(n+2)/(p-1)}} , \qquad B := 400n , \qquad q = p-1 . \tag{4.19}$$

Since A depends itself on n,p,ν,L, this ultimately fixes a positive constant $\delta_\varepsilon \in (0,1/2)$ depending only on n,p,ν,L; we may assume $\delta_\varepsilon \le (\log 2)^{p-1}$. Now define

$$Q_i := Q^\lambda_{r_i}\,, \qquad r_i = \delta_1^i r\,, \qquad \delta_1 := \delta_\varepsilon/2 \tag{4.20}$$

whenever $i \ge 0$ is an integer; again $\delta_1 \equiv \delta_1(n,p,\nu,L) \in (0,1/4)$. We also set

$$H_1 := 400\delta_1^{-(n+2)/(p-1)} \tag{4.21}$$

so that

$$\left(\fint_{Q_0} (|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} + \delta_1^{-(n+2)/(p-1)}E(Du,Q_0) \le \frac{\lambda}{100}\,. \tag{4.22}$$

Next, we again look at Theorem 3.2 and with the choice of A made in (4.18) we consider the exponent α determined by A; again we observe that $\alpha \equiv \alpha(n,p,\nu,L) \in (0,1)$. In the same way, referring to Theorem 3.2, we determine the corresponding constant $c_h \equiv c_h(n,p,\nu,L,A)$ with the choice of A made in (4.18). As A is itself depending on n,p,ν,L we have that c_h depends again only on n,p,ν,L. We take now $k \ge 2$ as the smallest integer (larger or equal than 2) so that

$$c_h\delta_1^{(k-1)\alpha} \le \frac{\delta_1^{(n+2)/(p-1)}}{800}\,. \tag{4.23}$$

Then k depends only upon n,p,ν,L as also δ_1 and c_h do. With k fixed, choose in turn $H_2 \equiv H_2(n,p,\nu,L)$ as follows:

$$H_2 := 2400c_4\delta_1^{-(k+3)(n+2)/(p-1)}\,, \tag{4.24}$$

where $c_4 \equiv c_4(n,p,\nu)$ has been fixed in Corollary 4.1. Now observe that

$$\begin{aligned}
&\int_0^{2r}\left(\frac{|\mu|(Q^\lambda_\varrho)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho}\\
&= \sum_{i=0}^\infty \int_{r_{i+1}}^{r_i}\left(\frac{|\mu|(Q^\lambda_\varrho)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho} + \int_r^{2r}\left(\frac{|\mu|(Q^\lambda_\varrho)}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho}\\
&\ge \sum_{i=0}^\infty \left(\frac{|\mu|(Q_{i+1})}{\lambda^{2-p}{r_i}^{n+1}}\right)^{1/(p-1)}\int_{r_{i+1}}^{r_i}\frac{d\varrho}{\varrho} + \left(\frac{|\mu|(Q_0)}{\lambda^{2-p}(2r)^{n+1}}\right)^{1/(p-1)}\int_r^{2r}\frac{d\varrho}{\varrho}\\
&= \delta_1^{(n+1)/(p-1)}\log\left(\frac{1}{\delta_1}\right)\sum_{i=0}^\infty\left(\frac{|\mu|(Q_{i+1})}{\lambda^{2-p}{r_{i+1}}^{n+1}}\right)^{1/(p-1)}\\
&\qquad +2^{-(n+1)/(p-1)}\log 2\left(\frac{|\mu|(Q_0)}{\lambda^{2-p}r^{n+1}}\right)^{1/(p-1)}\\
&\ge \delta_1^{(n+2)/(p-1)}\sum_{i=0}^\infty\left(\frac{|\mu|(Q_i)}{\lambda^{2-p}{r_i}^{n+1}}\right)^{1/(p-1)}\,.
\end{aligned} \tag{4.25}$$

Therefore, by (4.17) and the choice in (4.24) it follows that

$$8c_4\delta_1^{-(k+2)(n+2)/(p-1)}\sum_{i=0}^\infty\left(\frac{|\mu|(Q_i)}{\lambda^{2-p}r_i^{n+1}}\right)^{1/(p-1)} \le \frac{\lambda}{300}\,. \tag{4.26}$$

In particular, we have

$$\left(\frac{|\mu|(Q_i)}{\lambda^{2-p}r_i^{n+1}}\right)^{1/(p-1)} \le \frac{\delta_1^{(k+2)(n+2)/(p-1)}\lambda}{2400c_4} \le \frac{\lambda}{2400c_4} \le \lambda\,, \qquad \forall\, i \ge 0\,. \tag{4.27}$$

We are now ready to state some conditional estimates of later use, that we assemble in a lemma.

Lemma 4.2. *Let $w_i \equiv w$ be the comparison function of Lemma 4.1 defined in* (4.2) *with $Q_\varrho^\lambda \equiv Q_i$, i.e. w_i solves*

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

If

$$\left(\fint_{Q_i} (|Dw_i| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)} \leq \lambda\,, \tag{4.28}$$

then

$$s + \sup_{Q_{i+1}} ||Dw_i|| \leq s + \sup_{\frac{1}{2}Q_i} ||Dw_i|| \leq A\lambda\,. \tag{4.29}$$

Moreover, with $k \equiv k(n,p,\nu,L) \geq 2$ being the integer defined via (4.23)*, it holds that*

$$2\delta_1^{-(n+2)/(p-1)} E(Dw_i, Q_{i+k}) \leq \frac{\lambda}{400} \tag{4.30}$$

and

$$\left(1 + 2\delta_1^{-(n+2)/(p-1)}\right) \left(\fint_{Q_{i+k}} |Du - Dw_i|^{p-1}\, dx\, dt \right)^{1/(p-1)} \leq \frac{\lambda}{400}\,. \tag{4.31}$$

Proof. First, by Theorem 3.3 in view of (4.28), it follows that

$$s + \sup_{\frac{1}{2}Q_i} ||Dw_i|| \leq s + 2c_3\lambda\,.$$

The choice of H_1 in (4.21) implies that

$$s \leq \frac{\lambda}{400} \tag{4.32}$$

and thus we have that the choice of A in (4.18) gives (4.29) as $Q_{i+1} \subset (1/2)Q_i$. At this point, as a consequence of Theorem 3.2 (applied with $Q_\varrho^\lambda \equiv Q_{i+k}$ and $Q_r^\lambda \equiv Q_{i+1}$) and (4.23) we have

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

in turn implying (4.30) as, trivially,

$$E(Dw_i, Q_{i+k}) = \left(\fint_{Q_{i+k}} |Dw_i - (Dw_i)_{Q_{i+k}}|^{p-1}\, dx\, dt \right)^{1/(p-1)} \leq \operatorname*{osc}_{Q_{i+k}} Dw_i\,.$$

Finally, by (4.27) we may apply Corollary 4.1 so that

$$\begin{aligned} &\left(\fint_{Q_{i+k}} |Du - Dw_i|^{p-1}\, dx\, dt \right)^{1/(p-1)} \\ &\quad \leq \left(\frac{|Q_i|}{|Q_{i+k}|} \right)^{1/(p-1)} \left(\fint_{Q_i} |Du - Dw_i|^{p-1}\, dx\, dt \right)^{1/(p-1)} \\ &\quad \leq c_4 \delta_1^{-k(n+2)/(p-1)} \left(\frac{|\mu|(Q_i)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)} \\ &\quad \leq \frac{\delta_1^{(n+2)/(p-1)}}{2400} \lambda\,, \end{aligned}$$

where in the last estimate we used (4.27). Now (4.31) follows as a consequence of the trivial estimate $(1+2\delta_1^{-(n+2)/(p-1)}) \leq 3\delta_1^{-(n+2)/(p-1)}$. The proof of the lemma is complete. □

We finally remark that, all in all, the constants A, δ_1, k, H_1, H_2 have been determined in a way that makes them depending only on n, p, ν, L.

Step 2: The exit time argument. Define now, whenever $i \geq 0$

$$C_i := \left(\fint_{Q_i} (|Du| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)} + \delta_1^{-(n+2)/(p-1)} E(Du, Q_i)\,. \tag{4.33}$$

Now, observe that (4.22) reads also as

$$C_0 \leq \frac{\lambda}{100}\,.$$

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

$$C_{i_e} \leq \frac{\lambda}{100}, \qquad C_{i_e+m} > \frac{\lambda}{100}, \qquad \forall\, m \geq 1\,. \tag{4.34}$$

Indeed, on the contrary, we could find an increasing subsequence $\{j_i\}$ such that

$$C_{j_i} \leq \frac{\lambda}{100} \qquad \text{for every } i\,,$$

and then, as the gradient is supposed to be continuous, obviously

$$|Du(x_0, t_0)| = \lim_{i\to\infty} \left(\fint_{Q_{j_i}} |Du|^{p-1}\, dx\, dt \right)^{1/(p-1)} \leq \frac{\lambda}{100}\,, \tag{4.35}$$

and the proof would be finished. Therefore, from now on, for the rest of the proof, we shall argue under the additional assumption (4.34).

Step 3: After the exit. The next result exploits the effect of arguing "after the exit time".

Lemma 4.3. *Assume that for $i \geq i_e$ it holds that*

$$\left(\fint_{Q_i} (|Dw_i| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)} \leq \lambda\,. \tag{4.36}$$

Then the inequality

$$\frac{\lambda}{400n} \leq \sup_{\frac{\delta_\varepsilon}{2} Q_i} ||Dw_i|| \tag{4.37}$$

also holds. Here $\delta_\varepsilon = 2\delta_1$ has been determined in Step 1 when applying Theorem 3.1 with the choice (4.19).

Proof. Observe that by (4.36) we may use Lemma 4.2. By using (2.4), triangle inequality, (4.30) and (4.31), we have

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

$$\leq \left(\fint_{Q_{i+k}} (|Dw_i| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)} + 2\delta_1^{-(n+2)/(p-1)} E(w_i, Q_{i+k})$$
$$+ \left(1 + 2\delta_1^{-(n+2)/(p-1)}\right) \left(\fint_{Q_{i+k}} |Du - Dw_i|^{p-1}\, dx\, dt \right)^{1/(p-1)}$$
$$\leq \left(\fint_{Q_{i+k}} (|Dw_i| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)} + \frac{\lambda}{200}\,.$$

The previous inequality and (4.34) then give

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

for all integers $i \geq i_e$ (recall here that $k \geq 2$ so that $i + k > i_e$; it is also useful to recall (3.3)). Finally, using (4.32) it also follows that

$$\frac{\lambda}{400n} \leq \sup_{Q_{i+1}} ||Dw_i||\,.$$

In turn, observe that by the definition of δ_1 in (4.20) we have

$$Q_{i+1} = Q^{\lambda}_{\delta_1^{i+1} r} = Q^{\lambda}_{\delta_\varepsilon \delta_1^i r/2} = (\delta_\varepsilon/2) Q_i \tag{4.38}$$

so that (4.37) follows and the lemma is proved. □

Step 4: Excess decay. The following lemma exploits a decay property enjoyed by the excess functional after the exit time.

Lemma 4.4. *Let $i \geq i_e$ and assume that the condition*

$$s + |(Du)_{Q_i}| + E(Du, Q_i) \leq \frac{\lambda}{2} \tag{4.39}$$

holds. Then the inequality

$$E(Du, Q_{i+1}) \leq \tfrac{1}{4} E(Du, Q_i) + 4c_4 \delta_1^{-(n+2)/(p-1)} \left(\frac{|\mu|(Q_i)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)} \tag{4.40}$$

holds too.

Proof. Let us first show that we are able to use both Lemma 4.2 and 4.3. In fact, by Corollary 4.1 and (4.27), we have

$$\left(\fint_{Q_i} (|Dw_i| + s)^{p-1}\, dx\, dt \right)^{1/(p-1)}$$
$$\leq s + \left(\fint_{Q_i} |Du|^{p-1}\, dx\, dt \right)^{1/(p-1)} + \left(\fint_{Q_i} |Du - Dw_i|^{p-1}\, dx\, dt \right)^{1/(p-1)}$$
$$\leq s + |(Du)_{Q_i}| + E(Du, Q_i) + c_4 \left(\frac{|\mu|(Q_i)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)}$$
$$\leq s + |(Du)_{Q_i}| + E(Du, Q_i) + \frac{\lambda}{1200}$$
$$\leq \lambda\,.$$

Since (4.28) is satisfied, at this point we can apply both Lemma 4.2 and Lemma 4.3 to get (4.29) and (4.37), respectively; summarizing we have

$$\frac{\lambda}{400n} \leq \sup_{\frac{\delta_\varepsilon}{2} Q_i} ||Dw_i|| \leq s + \sup_{\frac{1}{2} Q_i} ||Dw_i|| \leq A\lambda\,.$$

The last inequality allows to apply Theorem 3.1 to $w_i (= w)$, with the choice made in (4.19), in the cylinder $(1/2)Q_i (= Q_r^\lambda$ in the notation of Theorem 3.1), thereby obtaining

$$E(Dw_i, Q_{i+1}) = E(Dw_i, (\delta_\varepsilon/2)Q_i) \leq \frac{1}{2^5 2^{(n+2)/(p-1)}} E(Dw_i, (1/2)Q_i)\,, \tag{4.41}$$

where we have kept (4.38) in mind. In turn, let us estimate as follows:

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

Connecting the last inequality with (4.41) gives

$$E(Dw_i, Q_{i+1}) \leq \tfrac{1}{16} E(Dw_i, Q_i)\,. \tag{4.42}$$

On the other hand we have

$$\begin{aligned} &E(Du, Q_{i+1}) \\ &= \left(\fint_{Q_{i+1}} |Du - (Du)_{Q_{i+1}}|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\leq 2\left(\fint_{Q_{i+1}} |Du - (Dw)_{Q_{i+1}}|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\leq 2E(Dw_i, Q_{i+1}) + 2\left(\fint_{Q_{i+1}} |Du - Dw_i|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\leq 2E(Dw_i, Q_{i+1}) + 2\delta_1^{-(n+2)/(p-1)} \left(\fint_{Q_i} |Du - Dw_i|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\leq 2E(Dw_i, Q_{i+1}) + 2c_4\delta_1^{-(n+2)/(p-1)} \left(\frac{|\mu|(Q_i)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)} \end{aligned} \tag{4.43}$$

and, similarly

$$E(Dw_i, Q_i) \leq 2E(Du, Q_i) + 2c_4 \left(\frac{|\mu|(Q_i^\lambda)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)}\,.$$

Connecting this last inequality with (4.42) and (4.43) yields (4.40). The proof of Lemma 4.4 is complete. □

Step 5: Iteration and conclusion. We conclude the proof via an iteration procedure that starts from the exit time i_e; in other words we shall consider only indexes $i \geq i_e$, where i_e has been defined in (4.34). Denote in short

$$A_i := E(Du, Q_i), \qquad k_i = |(Du)_{Q_i}|\,.$$

Recall that by the definitions in (4.33) and (4.34), and since $p \geq 2$, we have

$$s + k_{i_e} + \delta_1^{-(n+2)/(p-1)} A_{i_e} \leq C_{i_e} \leq \frac{\lambda}{100}\,. \tag{4.44}$$

We now prove, by induction, that

$$s + k_j + A_j \leq \frac{\lambda}{4} \tag{4.45}$$

holds whenever $j \geq i_e$. Indeed, by (4.44) the case $j = i_e$ of the previous inequality holds. Then, assume by induction that (4.45) holds whenever $j \in \{i_e, \ldots, i\}$, and tis means in particular that (4.39) is verified for all $j \in \{i_e, \ldots, i\}$. Applying Lemma 4.4 estimate (4.40) implies

$$A_{j+1} \leq \tfrac{1}{4} A_j + 4c_4 \delta_1^{-(n+2)/(p-1)} \left(\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^{n+1}} \right)^{1/(p-1)} \tag{4.46}$$

for all $j \in \{i_e, \ldots, i\}$. It immediately follows by (4.45) (assumed for all $j \in \{i_e, \ldots, i\}$) and (4.27) that

$$A_{i+1} \leq \frac{\lambda}{16} + 4c_4 \delta_1^{-(n+2)/(p-1)} \left(\frac{|\mu|(Q_i)}{\lambda^{2-p} r_i^{n+1}} \right)^{1/(p-1)} \leq \frac{\lambda}{16} + \frac{\lambda}{600} \leq \frac{\lambda}{14} . \tag{4.47}$$

Furthermore, summing up (4.46) for $j \in \{i_e, \ldots, i\}$ gives

$$\sum_{j=i_e}^{i+1} A_j \leq A_{i_e} + \frac{1}{4} \sum_{j=i_e}^{i} A_j + 4c_4 \delta_1^{-(n+2)/(p-1)} \sum_{j=i_e}^{i+1} \left(\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^{n+1}} \right)^{1/(p-1)} ,$$

yielding

$$\sum_{j=i_e}^{i+1} A_j \leq 2A_{i_e} + 8c_4 \delta_1^{-(n+2)/(p-1)} \sum_{j=i_e}^{i+1} \left(\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^{n+1}} \right)^{1/(p-1)} .$$

Next, using the previous inequality and Hölder's inequality ($p - 1 \geq 1$), we have

$$\begin{aligned}
k_{i+1} - k_{i_e} &= \sum_{j=i_e}^{i} (k_{j+1} - k_j) \\
&\leq \sum_{j=i_e}^{i} \fint_{Q_{j+1}} |Du - (Du)_{Q_j}| \, dx \, dt \\
&\leq \sum_{j=i_e}^{i} \left(\fint_{Q_{j+1}} |Du - (Du)_{Q_j}|^{p-1} \, dx \, dt \right)^{1/(p-1)} \\
&\leq \delta_1^{-(n+2)/(p-1)} \sum_{j=i_e}^{i} \left(\fint_{Q_j} |Du - (Du)_{Q_j}|^{p-1} \, dx \, dt \right)^{1/(p-1)} \\
&= \delta_1^{-(n+2)/(p-1)} \sum_{j=i_e}^{i} A_j \\
&\leq 2\delta_1^{-(n+2)/(p-1)} A_{i_e} + 8c_4 \delta_1^{-2(n+2)/(p-1)} \sum_{j=i_e}^{i+1} \left(\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^{n+1}} \right)^{1/(p-1)}
\end{aligned}$$

and thus it follows that

$$k_{i+1} \leq k_{i_e} + 2\delta_1^{-(n+2)/(p-1)} A_{i_e} + 8c_4 \delta_1^{-2(n+2)/(p-1)} \sum_{j=0}^{\infty} \left(\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^{n+1}} \right)^{1/(p-1)} .$$

In turn, by (4.44) and (4.26) the previous estimate yields

$$k_{i+1} \leq \frac{\lambda}{25} .$$

The last inequality together with (4.32) and (4.47) allows to verify the induction step, i.e.

$$s + k_{i+1} + A_{i+1} \leq \frac{\lambda}{4}\,. \tag{4.48}$$

Therefore (4.45) holds for every $i \geq i_e$. Estimate (1.16) finally follows with the choice (announced at the beginning) $c := \max\{H_1, H_2\}$, since, as Du is here assumed to be continuous, it holds that

$$|Du(x_0, t_0)| = \lim_{i\to\infty} k_i \leq \frac{\lambda}{4}\,.$$

4.3. **General measure data and Theorem 1.4.** We describe how to pass to the limit in Theorem 1.1, justifying the content of Section 1.4 and Theorem 1.4. We therefore start with the approximation settled up in [4] and outlined in Section 1.4. By possibly passing to a subsequence we may assume that

$$Du_h \in C^0, \qquad Du_h \to Du \text{ in } L^{p-1} \qquad \text{and} \qquad Du_h \to Du \ \ a.e. \tag{4.49}$$

Notice also that the first claim in the previous line follows by the regularity theory available for solutions to equations with a good right hand side; see remark 3.3.

In the following we shall keep the notation introduced in the proof of Theorem 1.1; the idea is not really to pass to the limit in the statement of Theorem 1.1, but rather passing to the limit in its proof. With (x_0, t_0) being a Lebesgue point of Du, we proceed as for Theorem 1.1, but we take H_2 as defined (4.24) with a larger constant, let's say $H_2 := 10^4 c_4 \delta_1^{-(k+3)(n+2)/(p-1)}$. Therefore, instead of (4.26) we have the better bound

$$8c_4\delta_1^{-(k+2)(n+2)/(p-1)} \sum_{i=0}^{\infty} \left(\frac{|\mu|(\lfloor Q_i \rfloor_{\rm par})}{\lambda^{2-p} r_i^{n+1}}\right)^{1/(p-1)} \leq \frac{\lambda}{600}\,. \tag{4.50}$$

Notice that in the previous line we are using (4.50) with the parabolic closure $\lfloor Q_i \rfloor_{\rm par}$ (which has been defined in (2.3)) instead of Q_i appearing in (4.50). This comes from (4.25), which in turn involves a one dimensional integral in the definition of the potential, and therefore we can w.l.o.g. assume the that $|\mu|(\partial_{\rm par} Q_i) = 0$, for every $i \geq 0$. Next, we define, in analogy to (4.33), the quantities

$$C_i^h := \left(\fint_{Q_i} (|Du_h| + s)^{p-1}\, dx\, dt\right)^{1/(p-1)} + \delta_1^{-(n+2)/(p-1)} E(Du_h, Q_i)\,. \tag{4.51}$$

We now jump to Step 2 of the proof of Theorem 1.1 and we notice that we can again argue under the additional assumption (4.34), otherwise (4.35) holds since (x_0, t_0) is a Lebesgue point and we are done. Let us now fix an integer $M \geq i_e$; in view of (4.49), there exits $K_M \in \mathbb{N}$ such that the following holds whenever $h \geq K_M$:

$$C_{i_e}^h \leq \frac{\lambda}{99}, \qquad C_j^h > \frac{\lambda}{100}, \qquad \forall\, j \in \{i_e + 1, \ldots, M\}\,. \tag{4.52}$$

Moreover, as a consequence Definition (1) (especially, keep (1.27) in mind), and of the new choice of the constant H_2, we can also assume that following truncated version of (4.26) holds, whenever $h \geq K_M$:

$$8c_4\delta_1^{-(k+2)(n+2)/(p-1)} \sum_{i=0}^{M+1} \left(\frac{|\mu_h|(Q_i)}{\lambda^{2-p} r_i^{n+1}}\right)^{1/(p-1)} \leq \frac{\lambda}{500}\,. \tag{4.53}$$

The idea is now to replicate the proof of Theorem 1.1 for each u_h with $h \geq K_M$, starting from Lemma 4.2, except of course Step 2, and replacing the final induction of Step 5 by a finite induction that starts from the exit index i_e and stops at $M+1$. Indeed, following the proof of Theorem 1.1, and applying the arguments developed there to u_h, it is easy to see that (4.52)-(4.53) are sufficient to perform all the steps and making the iteration procedure of Step 5 until $M+1$, thereby in particular obtaining, in (4.48)

$$s + |(Du_h)_{Q_{M+1}}| + E(Du_h, Q_{M+1}) \leq \frac{\lambda}{4}\,.$$

By first letting $h \to \infty$ and then $M \to \infty$ in the previous inequality, and finally recalling that (x_0, t_0) is a Lebesgue point of Du, we conclude with (1.14) and Theorem 1.1 remains valid for general SOLA as prescribed in Theorem 1.4. Theorems 1.2 and 1.3 both follow for SOLA as a corollary of Theorem 1.1 as already shown in Section 1.4.

5. Alternative forms of the potential estimates

Our purpose here is to prove the following alternative form of Theorem 1.1:

Theorem 5.1. *Let u be a solution to* (1.1) *such that Du is continuous in Ω_T and that $\mu \in L^1$. There exists a constant $c \geq 1$, depending only on n, p, ν, L, such that if $\lambda > 0$ is a generalized root of*

$$\lambda^{p/2} = c\beta^{p/2} + c\int_0^{2r} \left(\frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\lambda^{2-p}\varrho^{n+1}}\right)^{p/[2(p-1)]} \frac{d\varrho}{\varrho}\,, \tag{5.1}$$

and if β satisfies (1.15) *where $Q_{2r}^\lambda \subset \Omega_T$ is an intrinsic cylinder with vertex at (x_0, t_0), then $|Du(x_0, t_0)| \leq \lambda$.*

The main difference with Theorem 1.1 is in that we require a slightly less restrictive convergence assumption on the potential when considering the right hand side of (5.1). Indeed, let us recall the following elementary inequality for sequences $\{a_k\}$:

$$\sum_{k=0}^\infty a_k^q \leq \left(\sum_{k=0}^\infty a_k\right)^q, \qquad q \geq 1, \qquad a_k \geq 0\,,\ \forall\, k \in \mathbb{N}\,. \tag{5.2}$$

We now apply the previous fact with $q = p/2$ to perform the following computation, where $\lambda > 0$ and $r > 0$ are fixed numbers:

$$\begin{aligned}
\int_0^{r} \left(\frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\lambda^{2-p}\varrho^{n+1}}\right)^{p/[2(p-1)]} \frac{d\varrho}{\varrho} &= \sum_{k=0}^\infty \int_{r/2^{k+1}}^{r/2^k} \left(\frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\lambda^{2-p}\varrho^{n+1}}\right)^{p/[2(p-1)]} \\
&\lesssim \sum_{k=0}^\infty \left(\frac{|\mu|(Q_{r/2^k}^\lambda(x_0,t_0))}{\lambda^{2-p}(r/2^{k+1})^{n+1}}\right)^{p/[2(p-1)]} \\
&\leq \left[\sum_{k=0}^\infty \left(\frac{|\mu|(Q_{r/2^k}^\lambda(x_0,t_0))}{\lambda^{2-p}(r/2^{k+1})^{n+1}}\right)^{1/(p-1)}\right]^{p/2} \\
&\lesssim \left[\int_0^{2r} \left(\frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\lambda^{2-p}\varrho^{n+1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\right]^{p/2}.
\end{aligned}$$

Such an improvement has consequences for instance when applying Theorem 1.2 in order to get L^∞ criteria for Du in the setting of Lorentz spaces for the right hand side μ: it leads to slightly better exponents than those provided by Theorem 1.1.

The proof of Theorem 5.1 follows the lines of the one for Theorem 1.1 - that we actually preferred to give first in order to avoid to bother the reader immediately with so many technical complications - and therefore we shall confine ourselves to give the main modifications.

The main point is that we shall consider excess functionals based on the function $V(\cdot)$ considered in (2.7). In this respect we start recalling an additional property of the function $V(\cdot)$ defined in (2.7). Indeed, whenever $g\colon Q \to \mathbb{R}^n$ is an $L^{\gamma p/2}(Q)$ integrable map for $\gamma \geq 1$ and Q is a cylinder, a basic property of the function $V(\cdot)$ when $p \geq 2$ is given by the equivalences

$$\begin{aligned}&\fint_Q |V(g) - (V(g))_Q|^\gamma \, dx\, dt \\ &\qquad \approx \fint_Q |g-(g)_Q|^{p\gamma/2} + (s+|(g)_Q|)^{(p-2)\gamma/2}|g-(g)_Q|^\gamma \, dx\, dt\,. \end{aligned} \tag{5.3}$$

From now on, the number γ will be the one defined (4.15).

5.1. **A form of Theorem 3.1.** Given a measurable vector valued map $g\colon Q \to \mathbb{R}^n$, we define the new excess functional $\tilde{E}(\cdot)$ as

$$\tilde{E}(g,Q) = \left(\fint_Q |V(g) - (V(g))_Q|^\gamma \, dx\, dt\right)^{1/\gamma}. \tag{5.4}$$

The main result of this section is to provide an alternative form of Theorem 3.1.

Theorem 5.2. *Under the assumptions and the notation of Theorem 3.1, there exists $\delta_\varepsilon \in (0,1/2)$ depending only on $n, p, \nu, L, A, B, \varepsilon$, but otherwise independent of s, of the solution w considered and of the vector field $a(\cdot)$, such that the decay excess estimate $\tilde{E}(Dw, \delta_\varepsilon Q_r^\lambda) \leq \varepsilon \tilde{E}(Dw, Q_r^\lambda)$ holds.*

The proof of the above statement rests in turn on an different formulation of the Nondegenerate Alternative from Section 3.2. For this we need a few preliminary lemmas.

Lemma 5.1. *With the assumptions and the notations of Lemma 3.2 there exists a constant $c \geq 1$, depending only on n, p, ν, L, A, such that the decay excess estimate $\tilde{E}(Dw, Q_{\delta r}^\lambda) \leq c\delta^\beta \tilde{E}(Dw, Q_r^\lambda)$ holds for every $\delta \in (0,1)$.*

Proof. Lemma 3.2, and in particular (3.28) used in the case $q = p\gamma/2$, together with (5.3), tells us that we may confine ourselves to prove that

$$\begin{aligned}&\fint_{Q_{\delta r}^\lambda} (s+|(Dw)_{Q_{\delta r}^\lambda}|)^{(p-2)\gamma/2}|Dw-(Dw)_{Q_{\delta r}^\lambda}|^\gamma \, dx\, dt \\ &\qquad\qquad \leq c\delta^{\beta\gamma} \fint_{Q_r^\lambda} (s+|(Dw)_{Q_r^\lambda}|)^{(p-2)\gamma/2}|Dw-(Dw)_{Q_r^\lambda}|^\gamma \, dx\, dt \\ &\qquad\qquad\quad + c\delta^{\beta\gamma} \fint_{Q_r^\lambda} |Dw-(Dw)_{Q_r^\lambda}|^{p\gamma/2} \, dx\, dt \end{aligned} \tag{5.5}$$

holds whenever $\delta \in (0,1)$. Using (3.27) yields

$$\begin{aligned}(s+|(Dw)_{Q_{\delta r}^\lambda}|)^{(p-2)\gamma/2} &\leq c(s+\lambda)^{(p-2)\gamma/2} \leq c\fint_{Q_r^\lambda} (s+|Dw|)^{(p-2)\gamma/2} \, dx\, dt \\ &\leq c(s+|(Dw)_{Q_r^\lambda}|)^{(p-2)\gamma/2} + c\fint_{Q_r^\lambda} |Dw-(Dw)_{Q_r^\lambda}|^{(p-2)\gamma/2} \, dx\, dt\,. \end{aligned}$$

Using the previous inequality and applying (3.28) with $q = \gamma$ gives then

$$\begin{aligned}&\fint_{Q_{\delta r}^\lambda} (s+|(Dw)_{Q_{\delta r}^\lambda}|)^{(p-2)\gamma/2}|Dw-(Dw)_{Q_{\delta r}^\lambda}|^\gamma \, dx\, dt \\ &\leq c\delta^{\beta\gamma} \fint_{Q_r^\lambda} (s+|(Dw)_{Q_r^\lambda}|)^{(p-2)\gamma/2}|Dw-(Dw)_{Q_r^\lambda}|^\gamma \, dx\, dt \end{aligned}$$

$$+c\delta^{\beta\gamma} \fint_{Q_r^\lambda} |Dw-(Dw)_{Q_r^\lambda}|^{(p-2)\gamma/2}\,dx\,dt \fint_{Q_r^\lambda} |Dw-(Dw)_{Q_r^\lambda}|^{\gamma}\,dx\,dt\,.$$

In turn, applying Hölder's inequality twice with exponents $p/(p-2)$ and $p/2$ (this is necessary only when $p>2$) we see that the product of the last two integrals can be estimated by

$$\fint_{Q_r^\lambda} |Dw-(Dw)_{Q_r^\lambda}|^{p\gamma/2}\,dx\,dt$$

and this in turn yields (5.5); the proof is complete. □

The proof of the following variant of Proposition 3.5 can be now achieved arguing as in Lemma 5.1 and using Proposition 3.5 itself.

Lemma 5.2. *With the assumptions and the notations of Proposition 3.5 there exists a constant $c\geq 1$, depending only on $n,p,\nu,L,A,\gamma,\gamma_1$, such that the decay excess estimate $\tilde{E}(Dw,Q_{\delta r}^\lambda)\leq c\delta^{\beta_1}\tilde{E}(Dw,Q_r^\lambda)$ holds for every $\delta\in(0,1)$.*

Proof of Theorem 5.2. The proof essentially rests on the alternative formulation of the Nondegenerate Alternative, where instead of (3.56) we can now apply the inequality $\tilde{E}(Dw,Q_{\delta r}^\lambda)\leq c\delta^{\beta}\tilde{E}(Dw,Q_r^\lambda)$; this is basically a consequence of Lemma 5.1 applied in Proposition 3.3 instead of Lemma 3.2. This observation being made, the rest of the proof proceeds exactly as for Theorem 3.1, by using Lemma 5.2 instead of Proposition 3.5, and modulo an obvious change of the constants. □

5.2. **Proof of Theorem 5.1.** The proof is now completely similar to the one for Theorem 1.1; we just describe the main changes. First, instead of using the excess in (4.16) we shall use the one in (5.4). Moreover, we shall do the following replacement of the various quantities considered:

$$\left(\fint_{Q_i}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} \longleftrightarrow \left(\fint_{Q_i}(|V(Du)|+s^{p/2})^{\gamma}\,dx\,dt\right)^{1/\gamma} \tag{5.6}$$

and

$$\left(\fint_{Q_i}|Du-Dw_i|^{p-1}\,dx\,dt\right)^{1/(p-1)} \longleftrightarrow \left(\fint_{Q_i}|V(Du)-V(Dw_i)|^{\gamma}\,dx\,dt\right)^{1/\gamma}.$$

Passing from one quantity to another just needs to observe that

$$(|z|+s)^{p-1}\leq 2^{p-2}(|V(z)|+s^{p/2})^{\gamma}\leq 2^{p-2+\gamma}(|z|+s)^{p-1}\,. \tag{5.7}$$

In particular, while the first quantity in (5.6) was controlled by multiples of λ in Theorem 1.1, the second one will be now controlled by multiples of $\lambda^{p/2}$. The choice of the constants will be slightly different and will be adjusted taking into account the constants appearing in the inequalities involved (5.7); in particular $\delta_1=\delta_\varepsilon/2$ will be defined according to Theorem 5.2. Moreover, the constant $\delta_1^{-(n+2)/(p-1)}$ will be replaced by $\delta_1^{-(n+2)/\gamma}$ everywhere. We describe the main modifications according to the various steps.

Step 1. Here we observe that (1.15) and (5.7) give

$$\frac{\lambda^{p/2}}{H_1}\geq\beta^{p/2}\geq\frac{1}{2}\left(\fint_{Q_r^\lambda}(|V(Du)|+s^{p/2})^{\gamma}\,dx\,dt\right)^{1/\gamma}.$$

Therefore, choosing $H_1\equiv H_1(n,p,\nu,L)$ large enough allows to verify the initial smallness condition

$$\left(\fint_{Q_0}(|V(Du)|+s^{p/2})^{\gamma}\,dx\,dt\right)^{1/\gamma}+\delta_1^{-(n+2)/\gamma}E(Du,Q_0)\leq\frac{\lambda^{p/2}}{100}\,. \tag{5.8}$$

Estimate(4.27) must be replaced by

$$\left(\frac{|\mu|(Q_i)}{\lambda^{2-p}r_i^{n+1}}\right)^{1/\gamma} \leq \frac{\delta_1^{(k+2)(n+2)/\gamma}\lambda^{p/2}}{2400c_4} \leq \frac{\lambda^{p/2}}{2400c_4} \leq \lambda^{p/2}\,, \qquad \forall\ i \geq 0$$

with possibly new valus of c_4 and k. Accordingly, in Lemma 4.2 we use the new excess $\tilde{E}(\cdot)$ and (4.14) instead of (4.13). Moreover instead of Theorem 3.2 we shall apply Corollary 3.1.

Step 2. The definition in (4.33) has to be replaced by

$$C_i = \left(\fint_{Q_i} (|V(Du)| + s^{p/2})^\gamma \,dx\,dt\right)^{1/\gamma} + \delta_1^{-(n+2)/\gamma}\tilde{E}(Du, Q_i)$$

while the exit time condition looks like $C_{i_e+m} \geq \lambda^{p/2}/100$, for every $m \geq 1$.

Steps 3, 4 & 5. Here the main difference is that Theorem 5.2 must be applied to get an excess decay estimate for the excess $\tilde{E}(\cdot)$. Specifically (4.40) is replaced by

$$\tilde{E}(Du, Q_{i+1}) \leq \tfrac{1}{4}\tilde{E}(Du, Q_i) + 4c_4\delta_1^{-(n+2)/\gamma}\left(\frac{|\mu|(Q_i)}{\lambda^{2-p}r_i^{n+1}}\right)^{1/\gamma}$$

under the assumption $s^{p/2} + k_i + A_i \leq \lambda^{p/2}/4$ where $k_i := |(V(Du))_{Q_i}|$ and $A_i := \tilde{E}(Du, Q_i)$. Proceeding by induction we then prove that the inequality in the last line always holds whenever $i > i_e$; we conclude with

$$|Du(x_0, t_0)|^{p/2} \leq \lim_{i\to\infty} k_i \leq \frac{\lambda^{p/2}}{4}$$

and this finishes the proof.

Acknowledgements. The authors are supported by the ERC grant 207573 "Vectorial Problems" and by the Academy of Finland project "Potential estimates and applications for nonlinear parabolic partial differential equations". The authors also thank Paolo Baroni for remarks on a preliminary version of the paper. Last but not least, acknowledgments go to the referees for their careful reading of the paper and for the suggestions aimed at improving the presentation.

References

[1] Baroni P.: Marcinkiewcz estimates for degenerate parabolic equations with measure data. *Preprint 2012.*

[2] Boccardo L.: Elliptic and parabolic differential problems with measure data. *Boll. Un. Mat. Ital. A (7)* 11 (1997), 439461.

[3] Boccardo L. & Gallouët T.: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* 87 (1989), 149–169.

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

[5] Brézis H. & Strauss W. A.: Semi-linear second-order elliptic equations in L^1. *J. Math. Soc. Japan* 25 (1973), 565–590.

[6] Dal Maso G. & Murat F. & Orsina L. & Prignet A.: Renormalized solutions of elliptic equations with general measure data. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 28 (1999), 741–808.

[7] DiCastro A. & Palatucci G.: Nonlinear parabolic problems with lower order terms and related integral estimates. *Nonlinear Analysis TMA* 75 (2012), 4177–4197.

[8] DiBenedetto E.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Analysis TMA* 7 (1983), 827–850.

[9] DiBenedetto E.: On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 13 (1986), 487–535.

[10] DiBenedetto E.: *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993.

[11] DiBenedetto E. & Friedman A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. reine ang. Math. (Crelles J.)* 357 (1985), 1–22.

[12] Duzaar F. & Mingione G.: Gradient estimates via non-linear potentials. *Amer. J. Math.* 133 (2011), 1093–1149.

[13] Duzaar F. & Mingione G.: Gradient estimates via linear and nonlinear potentials. *J. Funct. Anal.* 259 (2010), 2961–2998.

[14] Havin M. & Maz'ja V. G.: A nonlinear potential theory. *Russ. Math. Surveys* 27 (1972), 71–148.

[15] Hedberg L.I. & Wolff T.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* 33 (1983), 161–187.

[16] Kilpeläinen T.: Singular solutions to p-Laplacian type equations. *Ark. Mat.* 37 (1999), 275–289.

[17] Kilpeläinen T. & Malý J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* 172 (1994), 137–161.

[18] Kilpeläinen T. & Kuusi T. & Tuhola-Kujanpää A.: Superharmonic functions are locally renormalized solutions. *Ann. Inst. H. Poincarè Anal. Non Linèaire* 28 (2011), 775–795.

[19] Kinnunen J. & Lukkari T. & Parviainen M.: An existence result for superparabolic functions. *J. Funct. Anal.* 258 (2010), 713–728.

[20] Kinnunen J. & Lukkari T. & Parviainen M.: Local approximation of superharmonic and superparabolic functions in nonlinear potential theory. *Preprint 2012.*

[21] Kinnunen J. & Parviainen M.: Stability for degenerate parabolic equations. *Adv. Calc. Var.* 3 (2010), 2948.

[22] Kuusi T. & Korte R.: A note on the Wolff potential estimate for solutions to elliptic equations involving measures. *Adv. Calc. Var.* 3 (2010), 99–113.

[23] Kuusi T. & Mingione G.: Nonlinear potential estimates in parabolic problems. *Rendiconti Lincei - Matematica e Applicazioni* 22 (2011), 161–174.

[24] Kuusi T. & Mingione G.: Potential estimates and gradient boundedness for nonlinear parabolic systems. *Rev. Mat. Iberoamericana* 28 (2012), 535–576.

[25] Kuusi T. & Mingione G.: Gradient regularity for nonlinear parabolic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (V)*, to appear.

[26] Kuusi T. & Mingione G.: A surprising linear type estimate for nonlinear elliptic equations. *C. R. Acad. Sci. Paris (Ser. I), Math.* 349 (2011), 889–892.

[27] Kuusi T. & Mingione G.: Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.*, DOI: 10.1007/s00205-012-0562-z.

[28] Kuusi T. & Mingione G.: New perturbation methods for nonlinear parabolic problems. *J. Math. Pures Appl. (IX)* 98 (2012), 390–427.

[29] Kuusi T. & Parviainen M.: Existence for a degenerate Cauchy problem. *Manuscripta Math.* 128 (2009), 213–249.

[30] Lieberman G.: Boundary and initial regularity for solutions of degenerate parabolic equations. *Nonlinear Analysis TMA* 20 (1993), 551–569.

[31] Lindqvist P.: On the definition and properties of p-superharmonic functions. *J. reine angew. Math. (Crelles J.)* 365 (1986), 67–79.

[32] Lindqvist P.: Notes on the p-Laplace equation. *Univ. Jyväskylä, Report* 102, (2006).

[33] Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Differential Equations* 76 (1988), 203–212.

[34] Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis.* University of Washington, St. Louis.

[35] Mingione G.: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (5)* 6 (2007), 195–261.

[36] Mingione G.: Gradient estimates below the duality exponent. *Math. Ann.* 346 (2010), 571–627.

[37] Mingione G.: Gradient potential estimates. *J. Europ. Math. Soc.* 13 (2011), 459–486.

[38] Mingione G.: Nonlinear measure data problems. *Milan J. Math.* 79 (2011), 429–496.

[39] Ragnedda F. & Vernier Piro S. & Vespri V.: Pointwise estimates for the fundamental solutions of a class of singular parabolic problems. *Preprint 2012.*

[40] Trudinger N.S. & Wang X.J.: On the weak continuity of elliptic operators and applications to potential theory. *Amer. J. Math.* 124 (2002), 369–410.

[41] Trudinger N.S. & Wang X.J.: Quasilinear elliptic equations with signed measure data. *Disc. Cont. Dyn. Systems A* 23 (2009), 477–494.

[42] Vázquez J.L.: *Smoothing and decay estimates for nonlinear diffusion equations. Equations of porous medium type.* Oxford Lecture Series in Math. Appl., 33. Oxford University Press, Oxford, 2006. xiv+234 pp.

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

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