Dark Side manuscript No.
(will be inserted by the editor)

Linear potentials in nonlinear potential theory

TUOMO KUUSI & GIUSEPPE MINGIONE

Abstract

Pointwise gradient bounds via Riesz potentials like those available for the linear Poisson equation actually hold for general quasilinear degenerate equations of p-Laplacean type. The regularity theory of such equations completely reduces to that of the classical Poisson equation up to the C^1-level.

1. Results

In this paper we prove the following:

Theorem 1.1. *Let $u \in C^1(\Omega)$ be a weak solution to the equation*

$$-\mathrm{div}\,(|Du|^{p-2}Du) = \mu\,, \tag{1.1}$$

where μ is a Borel measure with finite mass, $p \geq 2$, and $\Omega \subset \mathbb{R}^n$ is an open set with $n \geq 2$. Then there exists a constant c, depending only on n and p, such that the pointwise Riesz potential estimate

$$|Du(x)|^{p-1} \leq c\mathbf{I}_1^{|\mu|}(x,R) + c\left(\fint_{B(x,R)} |Du|\,dy\right)^{p-1} \tag{1.2}$$

holds whenever $B(x,R) \subseteq \Omega$ is a ball centered at x and with radius R.

In (1.2), $\mathbf{I}_1$ denotes the classical, linear (truncated) Riesz potential of $|\mu|$, which is suited to problems defined in bounded domains, and it is defined by

$$\mathbf{I}_1^{|\mu|}(x,R) := \int_0^R \frac{|\mu|(B(x,\varrho))}{\varrho^{n-1}}\,\frac{d\varrho}{\varrho}\,,$$

while we refer to Section 2 below for further notation. Theorem 1.1, proposed above in the form of an a priori estimate valid for energy solutions, actually extends to the case when u is rather a so-called very weak solution not necessarily belonging to $W^{1,p}(\Omega)$. This type of low integrability is typical when dealing with measure data problems [1,2,7,15,24]. The extension goes via a standard approximation argument briefly recalled in Section 5 below.

The surprising character of Theorem 1.1 mainly relies on the fact that, although considering degenerate quasilinear equations, the gradient can be pointwise estimated via Riesz potentials exactly as it happens for solutions to the hyper-classical Poisson equation

$$-\triangle u = \mu\,, \tag{1.3}$$

for which estimate (1.2) is an immediate consequence of the classical representation formula via Green's functions. Indeed, we have

Corollary 1.2. *Let $u \in W^{1,p}(\mathbb{R}^n)$ be a local weak solution to the equation* (1.1) *with $p \geq 2$ and μ being a Borel measure with locally finite mass. Then there exists a constant c, depending only on n, p, such that the following estimate holds for every Lebesgue point $x \in \mathbb{R}^n$ of Du:*

$$|Du(x)|^{p-1} \leq c \int_{\mathbb{R}^n} \frac{d|\mu|(y)}{|x-y|^{n-1}}\,.$$

The other thing that makes Theorem 1.1 somehow unexpected is that, starting from the seminal papers of Kilpeläinen & Malý [18,19], with different approaches offered by Trudinger & Wang [38,39] (that also worked in the subelliptic setting, starting by the ideas in [37,36]) and Korte & Kuusi [20] and Duzaar & Mingione [9], it is a standard fact that solutions to non-homogeneous Laplacean equations with measure data as (1.1) can be pointwise estimated in a natural way by means of classical nonlinear Wolff potentials [14], i.e.

$$\mathbf{W}^{\mu}_{\beta,p}(x,R) := \int_0^R \left(\frac{|\mu|(B(x,\varrho))}{\varrho^{n-\beta p}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho} \qquad \beta \in (0, n/p]\,.$$

In particular, the main result of [18,19] (see also [38,20,9] different approaches) claims that the following pointwise estimate holds:

$$|u(x)|^{p-1} \leq c\left[\mathbf{W}^{\mu}_{1,p}(x,R)\right]^{p-1} + c\left(\fint_{B(x,R)} |u|\,dy\right)^{p-1}\,. \tag{1.4}$$

The previous inequality is *sharp*, in the sense that $\mathbf{W}^{\mu}_{1,p}$ cannot be replaced by any other weaker potential. It plays a crucial role in the theory of quasilinear elliptic equations, as for instance shown in the work of Phuc & Verbitsky [34,35]. Estimate (1.4) has been eventually upgraded to the gradient level in [9,32], where a gradient estimate has been provided using again Wolff potentials:

$$|Du(x)|^{p-1} \leq c\left[\mathbf{W}^{\mu}_{1/p,p}(x,R)\right]^{p-1} + c\left(\fint_{B(x,R)} |Du|\,dy\right)^{p-1}\,. \tag{1.5}$$

Finally, a unifying approach allowing to view both estimate (1.4) and (1.5) as particular cases of a general family of "universal potential estimates" has been finally given by the authors in [21]. Estimate (1.2) obviously improves the one in (1.5) as, by using elementary manipulations together with the fact that $p \geq 2$, we have

$$\begin{aligned}
\mathbf{I}_1^{|\mu|}(x,R) &= \int_0^R \frac{|\mu|(B(x_0,\varrho))}{\varrho^{n-1}}\,\frac{d\varrho}{\varrho} && (1.6)\\
&\lesssim \sum_{k=0}^{\infty} \frac{|\mu|(B(x,R/2^k))}{(R/2^k)^{n-1}} && (1.7)\\
&\lesssim \left[\sum_{k=0}^{\infty}\left(\frac{|\mu|(B(x,R/2^k))}{(R/2^k)^{n-1}}\right)^{1/(p-1)}\right]^{p-1} && (1.8)\\
&\lesssim \left[\int_0^{2R}\left(\frac{|\mu|(B(x_0,\varrho))}{\varrho^{n-1}}\right)^{1/(p-1)}\,\frac{d\varrho}{\varrho}\right]^{p-1}\\
&= \left[\mathbf{W}^{\mu}_{1/p,p}(x,2R)\right]^{p-1}. && (1.9)
\end{aligned}$$

The main novelty of this paper is actually the fact that, when switching to the gradient regularity theory, Wolff potentials are not anymore necessary and *and the whole theory "linearizes" as in the case of the standard Poisson equation* $-\triangle u = \mu$. As a matter of fact - taking also into account Theorem 1.5 below - this paper shows the surprising fact that, when considered up to the C^1-level, *there is no difference between the regularity theory of general quasilinear degenerate equations and the one of the classical Poisson equation*. We interpret this fact by observing that, while an equation as (1.1) looks genuinely nonlinear in terms of the solution u, it rather looks *as a linear one* when considering the "stress tensor" $|Du|^{p-2}Du$; see Section 1.2 below and, in particular, (1.14).

We conjecture that estimate (1.2) is sharp in the sense that it cannot be improved by using different types of nonlinear potentials whatsoever. We also observe that the case $p < 2$ of estimate (1.2) has been proved in [10], but in this case it does not improve (1.5) in that (1.9) is in general false for $p < 2$ and the real problematic issue is in passing from nonlinear potentials to linear ones when the latter provide better bounds and Wolff potentials must be bypassed.

From the viewpoint of the standard regularity theory, the ultimate effect of the validity of estimate (1.2), as already mentioned above, is that *the integrability theory of solutions to p-Laplacean type equations is now completely reduced to the linear one*, i.e. there is basically no difference between degenerate quasilinear equations as (1.1) and the classical, linear Poisson equation. As explained for instance in [31] (where the case $p = 2$ is considered for general quasilinear equations) estimate (1.12) allows to recover, in a sharp way, all the gradient integrability results available for Du obtainable in terms of the measure μ and additional borderline cases. For instance, when the model equation (1.1) is considered, the results in [1,2,6,9,16,31] can be now recovered. For an overview of the relevant *Nonlinear Calderón-Zygmund theory* we refer to [33].

We also notice that a connection between Riesz potentials and the p-Laplacean operator has been found by Lindqvist & Manfredi in [26], where it is proved that certain Riesz potentials are p-superharmonic functions.

Theorem 1.1 is in turn a particular case of a result valid for more general equations, and in turn opens the way to provide a new, sharp continuity criterium for the gradient; see Theorems 1.3 and 1.5 below. We also observe that the results of this paper also allows provide another viewpoint to the ones in [21].

1.1. General elliptic equations

We describe the more general context to which our results apply. In the rest of the paper Ω will denote a bounded and open domain of $\mathbb{R}^n$, with $n \geq 2$, while $p \geq 2$. We shall consider general nonlinear, possibly degenerate elliptic equations with p-growth of the type

$$-\text{div}\, a(Du) = \mu \qquad \text{in } \Omega \tag{1.10}$$

whenever μ is a Borel measure with finite mass. The vector field $a\colon \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be C^1-regular and satisfying the following *growth and ellipticity assumptions*:

$$\begin{cases} |a(z)| + |a_z(z)|(|z|^2+s^2)^{1/2} \leq L(|z|^2+s^2)^{(p-1)/2} \\ \nu(|z|^2+s^2)^{(p-2)/2}|\xi|^2 \leq \langle a_z(z)\xi, \xi\rangle \end{cases} \tag{1.11}$$

whenever $z, \xi \in \mathbb{R}^n$, where $0 < \nu \leq L$ and $s \geq 0$ are fixed parameters. A model case for the previous situation is clearly given by considering the p-Laplacean equation in (1.1) or by its nondegenerate version (when $s > 0$)

$$-\text{div}\,[(|Du|^2+s^2)^{(p-2)/2}Du] = \mu\,.$$

The result is now

Theorem 1.3. *Let $u \in W^{1,p}(\Omega)$ be a weak solution to the equation* (1.10) *under the assumptions* (1.11) *with $p \geq 2$, where μ is a Borel measure with finite total mass defined on Ω. Then there exists a constant c, depending only on n, p, ν, L, such that the pointwise estimate*

$$|Du(x)|^{p-1} \leq c\mathbf{I}_1^{|\mu|}(x,R) + c\left(\fint_{B(x,R)} (|Du|+s)\,dy\right)^{p-1} \tag{1.12}$$

holds whenever $B(x,R) \subseteq \Omega$ and $x \in \Omega$ is a Lebesgue point of Du.

More general cases when $u \notin W^{1,p}(\Omega)$ also follow via approximation; see Section 5 below. Theorem 1.3 yields in turn the following, immediate

Corollary 1.4. *Let $u \in W^{1,p}(\Omega)$ be as in Theorem 1.3. Then*

$$\mathbf{I}_1^{|\mu|}(\cdot, R) \in L^\infty_{\mathrm{loc}}(\Omega) \textit{ for some } R > 0 \quad \Longrightarrow \quad Du \in L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^n)\,.$$

In particular, there exists a constant c, depending only on n, p, ν, L, such that the following estimate holds whenever $B_R \subseteq \Omega$:

$$\|Du\|_{L^\infty(B_{R/2})} \leq c \left\| \mathbf{I}_1^{|\mu|}(\cdot, R) \right\|_{L^\infty(B_R)}^{1/(p-1)} + c \fint_{B_R} (|Du| + s)\, dy\,.$$

The previous result is striking as it states that the classical, sharp Riesz potential criterium implying the Lipschitz continuity of solutions to the Poisson equations remains valid when considering the p-Laplacean operator. More analogies actually take place. Indeed, the techniques used for Theorem 1.1 also yield gradient continuity criteria for solutions to nonlinear equations, and we have

Theorem 1.5. *Let $u \in W^{1,p}(\Omega)$ be as in Theorem 1.3. If*

$$\lim_{R \to 0} \mathbf{I}_1^{|\mu|}(x, R) = 0 \textit{ locally uniformly in } \Omega \textit{ w.r.t. } x\,, \tag{1.13}$$

then Du is continuous in Ω.

An intermediate VMO-regularity result for Du is also given in Theorem 4.3 below. Theorem 1.5 admits the following relevant corollary, providing gradient continuity when μ is a function belonging to a borderline Lorentz space:

Corollary 1.6. *Let $u \in W^{1,p}(\Omega)$ be as in Theorem 1.3. If $\mu \in L(n, 1)$ locally holds in Ω, that is if*

$$\int_0^\infty |\{x \in \Omega' \,:\, |\mu(x)| > t\}|^{1/n}\, dt < \infty \quad \textit{for every open subset } \Omega' \Subset \Omega\,,$$

then Du is continuous in Ω.

We recall that Lorentz spaces interpolate Lebesgue spaces in the sense that $L^\gamma \subset L(n, 1) \subset L^n$ holds for every $\gamma > n$, all the inclusions being strict. Corollary 1.6 strikingly extends the known results available in the literature where the condition $\mu \in L(n, 1)$ is found to be a sufficient one for gradient boundedness of solutions [4, 12]. This condition is sharp already in the case of the Poisson equation [3] and we remark that the the two-dimensional case $n = 2$ oddly remained an open problem in [4, 12], essentially for technical reasons; this gap is settled here. Moreover, in these papers the gradient boundedness was proved while here we prove the continuity. The results extends also to fully nonlinear elliptic equations as $F(D^2u) = 0$ as eventually shown in [8], and upgrades to the optimal level a previous continuity criterion obtained in [11], claiming the continuity of the gradient in the case $\mu \in L(n, 1/(p-1))$. Indeed, $L(n, 1) \subset L(n, 1/(p-1))$, this inclusion being strict for $p > 2$. For more on Lorentz spaces we for instance refer to [13].

Finally, another immediate corollary of Theorem 1.5 is concerned with those measures satisfying special density properties.

Corollary 1.7. *Let $u \in W^{1,p}(\Omega)$ be as in Theorem 1.3. Assume that the measure μ satisfies the density condition*

$$|\mu|(B_R) \le cR^{n-1}h(R) \qquad \textit{with} \qquad \int_0^R h(\varrho)\,\frac{d\varrho}{\varrho} < \infty$$

for every ball $B_R \subset \mathbb{R}^n$, where $c \ge 0$. Then Du is continuous in Ω.

The previous results upgrades, in an optimal fashion and to the gradient level, analogous criteria for the continuity of u obtained by Lieberman [23].

1.2. Techniques

The techniques introduced in this paper considerably depart from those used in the literature to obtain nonlinear potential estimates in that here, in order to obtain linear type estimates for solutions to a non-linear equation, *we introduce a local linearization technique* that in fact incorporates, in a pointwise fashion and in a maybe not easy recognizable way, several ingredients from several different theories. In particular, since estimate (1.2) provides integrability estimates for the gradient, then exit time arguments of the type used in the classical linear Calderón-Zygmund theory are used here; on the other hand, since in the best possible cases the arguments provide $C^{0,1} - C^1$-regularity results (as in Theorem 1.5), basic elements of the classical De Giorgi-Nash-Moser theory must be used in the proof. Finally, since on the other hand the a priori estimates found cover the case when μ is assumed to be a Borel measure, then the technology for measure data problems must be employed as well [1,2,7]. Mixing-up all these different ingredients in a single proof is really delicate, and the final result is a proof that probably deserves a preliminary road-map to be approached.

Let us disclose the heuristic strategy for the model case (1.1). The idea is to conceive equation (1.1) as a decoupled system

$$-\mathrm{div}\, v = \mu \qquad \text{and} \qquad v := |Du|^{p-2}Du \tag{1.14}$$

and therefore to proceed for a linear estimate for v, via a Riesz potential. This brave argument immediately finds an obstruction in the fact that the first equation in (1.14) does not find necessarily a solution of the type $v := |Du|^{p-2}Du$. To bypass it we introduce a delicate linearization argument allowing, in a sense, to treat $|Du|^{p-2}$ as a Muckenhoupt weight, thanks to the fact that, in a certain sense, we can reduce to the case that u is "almost" a solution to the p-harmonic equation

$$\mathrm{div}\,(|Du|^{p-2}Du) = 0\,. \tag{1.15}$$

First of all, since the function $\tilde u := \varepsilon u$ solves

$$-\mathrm{div}\,(|D\tilde u|^{p-2}D\tilde u) = \varepsilon^{p-1}\mu\,,$$

we can always assume that the total variation of μ is very small by taking ε small. This is actually a version of the good-λ inequality principle, that allows, via a delicate comparison argument, to consider that u is almost a solution to (1.15) (see Lemmas 2.4 and 3.1 below) in the sense that its distance to a real p-harmonic function is quantitatively small. At this stage we make crucial use of the methods and the regularity results from the theory of measure data equations [1,2] that we have to carefully exploit here. Using this information we can then start the linearization method via an exit time argument; more precisely, let us set

$$\lambda^{p-1} \approx \mathbf{I}_1^{|\mu|}(x,R) + \left(\fint_{B(x,R)} |Du|\, dy \right)^{p-1} \tag{1.16}$$

and let us consider a sequence of shrinking balls $B_i(x) \equiv B(x, \delta^i R)$, where $\delta \equiv \delta(n,p)$ is a suitably small, but yet universal, parameter. There are now two possibilities; the first is when the gradient averages are dominated by λ, that is

$$|(Du)_{B_i}| \leq \lambda$$

happens for infinitely many indexes i, and, at this stage, (1.2) follows at any Lebesgue point of the gradient. The other case, on which we obviously concentrate, is that there exists an "exit time index" $i_e \in \mathbb{N}$ such that

$$|(Du)_{B_i}| > \lambda \tag{1.17}$$

whenever $i > i_e$. Now, since Du is almost harmonic in B_i we can commute the information in average in (1.17) in an everywhere information

$$|Du(x)| \geq \lambda / M \tag{1.18}$$

for another universal, possibly large, constant $M \geq 1$. In other words, we see that the equation becomes non-degenerate. We now look at the equation (1.1) and use the information in (1.18) to get, formally

$$-\triangle v = -\lambda^{p-2} \triangle u \text{ `` } \lesssim \text{ '' } \operatorname{div}(|Du|^{p-2} Du) = \mu\,, \tag{1.19}$$

where $v := \lambda^{p-2} u$, and therefore, the standard linear Riesz representation formula for solutions to the Poisson equation - this time applied to v - tells that

$$\lambda^{p-2} |Du(x)| \lesssim \mathbf{I}_1^{|\mu|}(x,R) + \lambda^{p-2} \fint_{B(x,R)} |Du|\, dy \tag{1.20}$$

from which estimate (1.2) follows by the definition of λ. Notice that, while the scheme of our proof of Theorem 1.1 follows the heuristic outlined above until (1.19), we cannot of course use the representation formula as in (1.20) since (1.19) holds only *formally*. At this stage the final linear type representation formula will be obtained along the proof by a careful induction argument. Needless to say, the whole heuristic argument outlined above must be made rigorous through an extremely careful control of all the constants involved.

The proof of Theorem 1.5 relies instead on a more sophisticated argument, and on the analysis of a potentially countable number of exit time moments. This marks a considerable difference between the case of potentials estimates for u and the ones for Du. In fact, in the case of potential bounds for u, this can be obtained for lower semicontinuous and positive solutions by using the Kilpeläinen & Malý [19] Wolff potential bound (1.4), the fact that the equation is invariant under translation (subtracting a constant from the solution still yields a solution) and the fact that positive solutions satisfy the weak Harnack inequality. No one of these ingredients is available for Du and a different path must be therefore taken.

The results in this paper have been announced in the CRAS note [22].

2. Preliminary results and notations

In the following we denote

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

When the center will not be relevant we shall simply denote $B_R \equiv B(x,R)$; moreover, we shall also denote σB the ball, concentric to B whose radius equals the one of B multiplied by $\sigma > 0$. A similar meaning is adopted when considering B/σ. We shall denote by c, δ, ε etc general positive constants; relevant functional dependence on the parameters will be emphasized by displaying them in parentheses; for example, to indicate a dependence of c on the real parameters n, p, ν, L we shall write $c \equiv c(n,p,\nu,L)$. Special constants will be denoted as c_0, c_h, etc. *All such constants, i.e. starting by c, are assumed to be larger or equal than one*. In the following, given a set $A \subset \mathbb{R}^n$ with positive measure and a map $g \in L^1(A,\mathbb{R}^n)$, we shall denote by

$$(g)_A := -\!\!\!\!\!\!\int_A g(y)\,dy$$

its integral average over the set A. Finally, we recall that, under the assumptions of Theorems 1.1-1.3, a weak solution $u \in W^{1,p}(\Omega)$ to (1.10) is a function such that

$$\int_\Omega \langle a(Du), D\varphi\rangle\,dy = \int_\Omega \varphi\,d\mu$$

holds whenever $\varphi \in W^{1,p}_0(\Omega) \cap L^\infty(\Omega)$.

With $s \geq 0$ introduced in (1.11), we define

$$V(z) \equiv V_s(z) := (s^2 + |z|^2)^{(p-2)/4} z\,, \qquad z \in \mathbb{R}^n\,, \tag{2.1}$$

which is easily seen to be a locally bi-Lipschitz bijection of $\mathbb{R}^n$. A basic property of the map $V(\cdot)$ is the following: For any $z_1, z_2 \in \mathbb{R}^n$, and any $s \geq 0$, the inequalities

$$\begin{aligned} c^{-1}\Big(s^2 + |z_1|^2 + |z_2|^2\Big)^{(p-2)/2} &\leq \frac{|V(z_2) - V(z_1)|^2}{|z_2 - z_1|^2} \\ &\leq c\Big(s^2 + |z_1|^2 + |z_2|^2\Big)^{(p-2)/2} \end{aligned} \tag{2.2}$$

hold for a constant $c \equiv c(n,p)$. We refer for instance to [30] for the last inequality and more properties of the map $V(\cdot)$. The strict monotonicity properties of the vector field $a(\cdot)$ implied by the left hand side in $(1.11)_2$ 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 $z_1, z_2 \in \mathbb{R}^n$:

$$\tilde{c}^{-1}|z_2 - z_1|^p \leq c^{-1}|V(z_2) - V(z_1)|^2 \leq \langle a(z_2) - a(z_1), z_2 - z_1 \rangle \, . \tag{2.3}$$

Let us now consider a weak solution $v \in W^{1,p}(A)$ to

$$\operatorname{div} a(Dv) = 0 \qquad \text{in } A\,, \tag{2.4}$$

where $A \subseteq \Omega$ is an open subset of Ω. The next lemma encodes in a suitable integral way the Hölder continuity properties of the Dv (see [9, Theorem 3.1] for a proof).

Theorem 2.1. *Let $v \in W^{1,p}(A)$ be a weak solution to* (2.4) *under the assumptions* (1.11)*. Then there exist constants $\beta \in (0,1]$ and $c_0 \geq 1$, both depending only on n, p, ν, L, such that the estimate*

$$\fint_{B_\varrho} |Dv - (Dv)_{B_\varrho}| \, dy \leq c_0 \left(\frac{\varrho}{R}\right)^{\beta} \fint_{B_R} |Dv - (Dv)_{B_R}| \, dy \tag{2.5}$$

holds whenever $B_\varrho \subseteq B_R \subseteq A$ are concentric balls.

The following result, which can be also obtained from the previous one, is essentially proved in [5, 27, 28, 21]. See for instance [27, (2.4)-(2.5)]

Theorem 2.2. *Let $v \in W^{1,p}(A)$ be a weak solution to* (2.4) *under the assumptions* (1.11)*. Then there exist constants $\beta \in (0,1]$ and $c_l, c_h \geq 1$, all depending only on n, p, ν, L, such that the estimates*

$$\|Dv\|_{L^\infty(B_{R/2})} \leq c_l \fint_{B_R} (|Dv| + s) \, dy \tag{2.6}$$

and

$$|Dv(x_1) - Dv(x_2)| \leq c_h \left(\fint_{B_R} (|Dv| + s) \, dy\right) \left(\frac{\varrho}{R}\right)^{\beta} \tag{2.7}$$

hold whenever $B_R \subseteq A$ is a ball and $x_1, x_2 \in B_\varrho \subset B_{R/2}$.

Next, a "density improvement lemma", based on Theorem 2.2.

Lemma 2.3. *Under the assumptions of Theorem 2.2, let $B \subset A$ be a ball and*

$$\frac{\lambda}{\Gamma} \leq \fint_{\sigma^m B} (|Dv| + s) \, dy \qquad \sup_{B/2} (|Dv| + s) \leq \Gamma\lambda \tag{2.8}$$

hold for some integer $m \geq 1$ and for numbers $\Gamma \geq 1$, $\lambda \geq 0$ and $\sigma \in (0, 1/4)$ such that

$$0 < \sigma \leq \left(\frac{1}{8c_h\Gamma^2}\right)^{1/\beta} , \tag{2.9}$$

where β and c_h are the constants appearing in Theorem 2.2. Then

$$\frac{\lambda}{4\Gamma} \leq |Dv| + s \qquad \textit{holds in } \sigma B\,.$$

Proof. The first inequality in (2.8) implies that there exists a point $x_0 \in \sigma^m B$ such that $|Dv(x_0)| + s > \lambda/(2\Gamma)$. On the other hand, as (2.7) and the second inequality in (2.8) give $|Dv(x) - Dv(x_0)| \leq c_h \Gamma \lambda (2\sigma)^\beta$ whenever $x \in \sigma B$, the last two inequalities and (2.9) then give

$$|Dv(x)| + s \geq |Dv(x_0)| + s - |Dv(x) - Dv(x_0)| \geq \frac{\lambda}{2\Gamma} - \frac{\lambda}{4\Gamma} = \frac{\lambda}{4\Gamma}$$

for all $x \in \sigma B$.

In the rest of the section the function $u \in W^{1,p}(\Omega)$ shall always be the one considered in Theorem 1.3. Let us now consider a ball $B_R \subset \Omega$ and define $v \in u + W_0^{1,p}(B_R)$ as the unique solution to the following Dirichlet problem:

$$\begin{cases} \operatorname{div} a(Dv) = 0 & \text{in } B_R \\ v = u & \text{on } \partial B_R. \end{cases} \tag{2.10}$$

Then, we give a comparison lemma.

Lemma 2.4. *Under the assumption* (1.11)*, let* $u \in W^{1,p}(\Omega)$ *be a solution to* (1.10)*, and* $v \in u + W_0^{1,p}(B_R)$ *be as in* (2.10)*. Then*

$$\int_{B_R} \frac{|V(Du) - V(Dv)|^2}{(\alpha + |u-v|)^\xi}\, dy \leq c \frac{\alpha^{1-\xi}}{\xi - 1} |\mu|(B_R) \tag{2.11}$$

holds whenever $\alpha > 0$ *and* $\xi > 1$*, where* $c \equiv c(n,p,\nu) \geq 1$*. Here the function* $V(\cdot)$ *has been defined in* (2.1)*.*

Proof. We begin with testing the weak formulation

$$\int_{B_R} \langle a(Du) - a(Dv), D\varphi \rangle \, dy = \int_{B_R} \varphi \, d\mu \,, \tag{2.12}$$

valid whenever $\varphi \in W_0^{1,p}(B) \cap L^\infty(B)$, by the functions

$$\eta_{1,\varepsilon} := \pm \min\{1, (u-v)_\pm/\varepsilon\} \,,$$

where $\varepsilon > 0$, and

$$\eta_{2,\varepsilon} := (\alpha + (u-v)_\pm)^{1-\xi} \eta_{1,\varepsilon} \,,$$

which both obviously belong to $W_0^{1,p}(B) \cap L^\infty(B)$. We recall the standard notation

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

Testing (2.12) by $\eta_{1,\varepsilon}$, yields

$$0 \leq \int_{B_R} \langle a(Du) - a(Dv), D\eta_{1,\varepsilon} \rangle \, dy = \int_{B_R} \eta_{1,\varepsilon} \, d\mu \leq |\mu|(B_R) \,. \tag{2.13}$$

We then test (2.12) by $\eta_{2,\varepsilon}$; as for the resulting term on the left hand side we notice

$$\begin{aligned}
&\int_{B_R} \langle a(Du)-a(Dv), D\eta_{2,\varepsilon}\rangle \, dy \\
&\quad = \int_{B_R} \langle a(Du)-a(Dv), D\eta_{1,\varepsilon}\rangle \frac{1}{(\alpha+(u-v)_\pm)^{\xi-1}}\, dy \\
&\qquad + (1-\xi)\int_{B_R} \langle a(Du)-a(Dv), D(u-v)_\pm\rangle \frac{\eta_{1,\varepsilon}}{(\alpha+(u-v)_\pm)^{\xi}}\, dy\,.
\end{aligned}$$

The first integral on the right can be majorized using (2.13) and the nonnegativity of the first integrand in (2.13), as, whenever $\varepsilon > 0$, it holds

$$\begin{aligned}
&\int_{B_R} \langle a(Du)-a(Dv), D\eta_{1,\varepsilon}\rangle \frac{1}{(\alpha+(u-v)_\pm)^{\xi-1}}\, dy \\
&\qquad \leq \alpha^{1-\xi}\int_{\Omega} \langle a(Du)-a(Dv), D\eta_{1,\varepsilon}\rangle \, dy \leq \alpha^{1-\xi}|\mu|(B_R)\,.
\end{aligned}$$

But since

$$\left|\int_{B_R} \eta_{2,\varepsilon}\, d\mu\right| \leq \alpha^{1-\xi}|\mu|(B_R)\,,$$

we obtain

$$(\xi-1)\int_{Q} \frac{\langle a(Du)-a(Dv), D(u-v)_\pm\rangle}{(\alpha+(u-v)_\pm)^{\xi}}\eta_{1,\varepsilon}\, dy \leq 2\alpha^{1-\xi}|\mu|(B_R)$$

and therefore, using (2.3) and the definition of $\eta_{1,\varepsilon}$, we conclude with

$$\int_{B_R} \frac{|V(Du)-V(Dv)|^2}{(\alpha+|u-v|)^{\xi}} \min\{1, |u-v|/\varepsilon\}\, dy \leq \frac{c\alpha^{1-\xi}}{\xi-1}|\mu|(B_R)\,.$$

Letting $\varepsilon \to 0$ yields (2.11).

Another comparison result is

Lemma 2.5. *Under the assumption* (1.11)*, let* $u \in W^{1,p}(\Omega)$ *be a solution to* (1.10) *as considered in Theorem 1.3, and* $v \in u + W^{1,p}_0(B_R)$ *be as in* (2.10)*. Then*

$$⨍_{B_R} |Du-Dv|^q\, dy \leq c_1 \left[\frac{|\mu|(B_R)}{R^{n-1}}\right]^{q/(p-1)} \tag{2.14}$$

holds whenever

$$0 < q < \min\left\{p, \frac{n(p-1)}{n-1}\right\} =: p_M\,, \qquad \textit{with } c_1 \equiv c_1(n,p,\nu,q)\,. \tag{2.15}$$

Proof. By approximation we can restrict to the case $\mu \in L^1$. Indeed, notice that with the assumption $u \in W^{1,p}(\Omega)$ it immediately follows that, as μ belongs to the dual $W^{-1,p'}$ (with dual norm depending on $\|u\|_{W^{1,p}(\Omega)}$). This in turn implies that if $w \in W^{1,p}(B_R)$ solves the Dirichlet problem

$$\begin{cases} \operatorname{div} a(Dw) = \mu & \text{in } B_R \\ \quad w = u & \text{on } \partial B_R \end{cases}$$

then $w = u$ (this is standard, just test by $u - w$ as in few lines below). Therefore, by approximation we can reduce to the case $\mu \in L^1$, as for instance explained in [1,2]. In turn, this allows to easily perform the scaling arguments used in [30, Lemma 4.1]. In turn, this allows to reduce to the case $B_R \equiv B_1$ and $|\mu|(B_1) \leq 1$, and to prove that

$$\fint_{B_1} |Du - Dv|^q \, dy \leq c(n, p, \nu, q) \,. \tag{2.16}$$

Next, notice that $p_M \leq p$ iff $p \leq n$, and this case has been proved in [30, Lemma 4.1]. The proof in the case $p > n$ is actually simpler, as μ belong to the dual $W^{-1,p'}$ (with dual norm depending on $|\mu|(B_1)$) and we provide a sketch here, by actually showing that (2.16) holds for $q = p$. By testing (2.12) with $u - v$ (recall that $p > n$ implies that $u - w \in L^\infty$) we easily get, via Morrey's embedding theorem and using $|\mu|(B_1) \leq 1$, that

$$\|Du - Dv\|_{L^p}^p \leq c\|u - v\|_{L^\infty} |\mu|(B_1) \leq c\|Du - Dv\|_{L^p} \,,$$

so that (2.16) with $q = p$ follows via Young's inequality.

Next we collect a number of purely technical lemmas that will be useful later. From now on, u and v will denote the functions appearing in Lemmas 2.5-2.4.

Lemma 2.6. *Assume that*

$$\left[\frac{|\mu|(B_R)}{R^{n-1}}\right]^{1/(p-1)} \leq \lambda \qquad \textit{and} \qquad \fint_{B_R} (|Du| + s) \, dy \leq \lambda$$

for some $\lambda \geq 0$. *Then*

$$\sup_{B_{R/2}} |Dv| + s \leq c_2 \lambda \qquad \textit{holds with } c_2 \equiv c_2(n, p, \nu, L) \,.$$

Proof. By using (2.6) and (2.14) we have

$$\sup_{B_{R/2}} |Dv| + s \leq c_l \fint_{B_R} |Dv - Du| \, dy + c_l \fint_{B_R} (|Du| + s) \, dy \leq 2c_1 c_l \lambda \,.$$

Lemma 2.7. *Let* $\delta, \theta \in (0, 1]$ *and* $\lambda \geq 0$. *Suppose that*

$$\left[\frac{|\mu|(B_R)}{R^{n-1}}\right]^{1/(p-1)} \leq \frac{\delta^n}{c_1} \theta \lambda \qquad \textit{and} \qquad \fint_{B_R} (|Du| + s) \, dy \leq \lambda \,,$$

where $c_1 \equiv c_1(n, p, \nu)$ *is as in Lemma 2.5. Then the following lower bound holds:*

$$\fint_{\delta B_R} (|Du| + s) \, dy - \theta \lambda \leq \fint_{\delta B_R} (|Dv| + s) \, dy \,. \tag{2.17}$$

Proof. Use of (2.14) yields

$$\begin{aligned}
\fint_{\delta B_R} |Du|\,dy &\leq \fint_{\delta B_R} |Du-Dv|\,dy + \fint_{\delta B_R} |Dv|\,dy \\
&\leq \delta^{-n} \fint_{B_R} |Du-Dv|\,dy + \fint_{\delta B_R} |Dv|\,dy \\
&\leq \theta\lambda + \fint_{\delta B_R} |Dv|\,dy\,. \qquad (2.18)
\end{aligned}$$

3. Proof of Theorem 1.3

In the following, given a ball $B \subset \Omega$, we define the excess functional

$$E(Du, B) := \fint_B |Du - (Du)_B|\,dy\,.$$

An elementary property of the functional above is given by the following:

$$\fint_B |Du - (Du)_B|\,dy \leq 2 \fint_B |Du - \gamma|\,dy \qquad \text{for every } \gamma \in \mathbb{R}^n\,. \qquad (3.1)$$

Finally, given a map $g\colon A \to \mathbb{R}^n$, we define its oscillation on A as

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

The proof of Theorem 1.3 begins with a preliminary lemma which will play an important role also in the proof of Theorem 1.5 in the next section. Given $\delta_1 \in (0, 1/2)$, we consider a ball $B(x, 2r) \subset \Omega$ and the shrinking balls

$$B_i := B(x, r_i)\,, \qquad r_i = \delta_1^i r\,, \qquad (3.2)$$

whenever $i \geq 0$ is an integer. The related comparison solutions $v_i \in u + W_0^{1,p}(B_i)$ are as in (2.10), that is

$$\begin{cases} \operatorname{div} a(Dv_i) = 0 & \text{in } B_i \\ v_i = u & \text{on } \partial B_i. \end{cases} \qquad (3.3)$$

Then we have

Lemma 3.1. *Assume that, for $i \geq 1$,*

$$\fint_{B_{i-1}} (|Du| + s)\,dy \leq \lambda \qquad \text{and} \qquad \fint_{B_i} (|Du| + s)\,dy \leq \lambda \qquad (3.4)$$

hold for a number $\lambda > 0$ satisfying

$$\left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{1/(p-1)} + \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{1/(p-1)} \leq \lambda\,, \qquad (3.5)$$

together with

$$\frac{\lambda}{H} \leq |Dv_{i-1}| \leq H\lambda \quad \textit{in } B_i \quad \textit{and} \quad \frac{\lambda}{H} \leq |Dv_i| \leq H\lambda \quad \textit{in } B_{i+1} \tag{3.6}$$

for some constant $H \geq 1$. Then there exists a constant $c_H \equiv c_H(n, p, \nu, L, \delta_1, H)$ such that the following inequality holds:

$$\fint_{B_{i+1}} |Du - Dv_i| \, dy \leq c_H \lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}} \right] . \tag{3.7}$$

Proof. We fix parameters γ and ξ as

$$\gamma := \frac{1}{4(n+1)(p+2)} \, , \qquad \xi = 1 + 2\gamma \, , \tag{3.8}$$

and introduce $\bar{v}_i = v_i/\lambda$ and $\bar{v}_{i-1} = v_{i-1}/\lambda$. In the case (3.7) does not trivialize (i.e. the left hand side is not zero) by (3.6) it will be sufficient to prove that

$$\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)} |Du - Dv_i| \, dy \leq c\lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}} \right] , \tag{3.9}$$

holds for a constant $c \equiv c(n, p, \nu, L, \delta_1, H)$. To this aim, applying Hölder's inequality and keeping (2.2) in mind, yields, for any $\alpha > 0$,

$$\begin{aligned}
&\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)} |Du - Dv_i| \, dy \\
&\leq c \fint_{B_i} \left[\lambda^{2-p} \frac{|V(Du) - V(Dv_i)|^2}{(\alpha + |u - v_i|)^\xi} \right]^{1/2} \\
&\qquad\qquad \cdot \left[|D\bar{v}_i|^{(p-2)(1+2\gamma)} (\alpha + |u - v_i|)^\xi \right]^{1/2} \, dy \\
&\leq c \left(\lambda^{2-p} \fint_{B_i} \frac{|V(Du) - V(Dv_i)|^2}{(\alpha + |u - v_i|)^\xi} \, dy \right)^{1/2} \\
&\qquad\qquad \cdot \left(\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+2\gamma)} (\alpha + |u - v_i|)^\xi \, dy \right)^{1/2} \qquad (3.10) \\
&\leq c \left(\lambda^{2-p} \frac{|\mu|(B_i)}{r_i^n} \alpha^{1-\xi} \right)^{1/2} \left(\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+2\gamma)} (\alpha + |u - v_i|)^\xi \, dy \right)^{1/2} .
\end{aligned}$$

Above c depends only on n, p, ν, because we have applied (2.11). As we are still free to choose α, we fix it as

$$\alpha := \left[\left(\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+2\gamma)} \, dy \right)^{-1} \fint_{B_i} |D\bar{v}_i|^{(p-2)(1+2\gamma)} |u - v_i|^\xi \, dy \right]^{1/\xi} .$$

We may assume $\alpha > 0$ since (3.6) would imply that (3.7) is trivial. Moreover, α is finite since in its definition the first integral is positive by (3.6), and, by (3.8), the second one is easily seen to be finite by Hölder's inequality and Sobolev embedding as both u and v_i belong to $W^{1,p}(B_i)$. Now (3.11) gives

$$
\begin{aligned}
&\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy \\
&\qquad \leq c\left(\lambda^{2-p}\frac{|\mu|(B_i)}{r_i^n}\alpha^{1-\xi}\right)^{1/2}\alpha^{\xi/2}\left(\fint_{B_i}|D\bar{v}_i|^{(p-2)(1+2\gamma)}\,dy\right)^{1/2} \qquad (3.11)
\end{aligned}
$$

To estimate the last integral, notice that

$$
\begin{aligned}
\fint_{B_i}|Dv_i - Dv_{i-1}|^q\,dy &\leq c\fint_{B_i}|Du - Dv_i|^q\,dy + c\fint_{B_i}|Du - Dv_{i-1}|^q\,dy \\
&\leq c\left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{q/(p-1)} + c\delta_1^{-n}\fint_{B_{i-1}}|Du - Dv_{i-1}|^q\,dy \\
&\leq c\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{q/(p-1)} \qquad (3.12)
\end{aligned}
$$

holds for all $1 \leq q \leq (p-1)(1+2\gamma) < p_M$, where $c \equiv c(n, p, \nu, \delta_1)$; here we have used (2.14) and kept (3.8) in mind. It then follows, by using (3.5) in (3.12), together with (3.6), that

$$
\begin{aligned}
\fint_{B_i}|D\bar{v}_i|^{(p-2)(1+2\gamma)}\,dy &\leq c\fint_{B_i}|D\bar{v}_i - D\bar{v}_{i-1}|^{(p-2)(1+2\gamma)}\,dy \\
&\quad + c\fint_{B_i}|D\bar{v}_{i-1}|^{(p-2)(1+2\gamma)}\,dy \leq c(n, p, \nu, L, H)\,.
\end{aligned}
$$

Using the last inequality in (3.11) then gives

$$
\begin{aligned}
\fint_{B_i}|D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy &\leq c\left(\lambda^{2-p}\frac{|\mu|(B_i)}{r_i^n}\alpha\right)^{1/2} \\
&\leq \varepsilon\frac{\alpha}{r_i} + \tilde{c}(\varepsilon)\lambda^{2-p}\left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]\,. \qquad (3.13)
\end{aligned}
$$

Notice that we have applied Young's inequality in the last estimate and, with $\varepsilon \in (0,1)$, the constant is $\tilde{c}(\varepsilon) \equiv \tilde{c}(n, p, \nu, L, \delta_1, H, \varepsilon)$.

We then start to estimate α. We first apply (3.6) to estimate

$$
\fint_{B_i}|D\bar{v}_i|^{(p-2)(1+2\gamma)}\,dy \geq \delta_1^n\fint_{B_{i+1}}|D\bar{v}_i|^{(p-2)(1+2\gamma)}\,dy \geq \frac{1}{c} \qquad (3.14)
$$

with $c \equiv c(n, p, \nu, L, \delta_1, H) > 1$. Therefore we obtain

$$
\alpha^{\xi} \leq c\fint_{B_i}|D\bar{v}_i|^{(p-2)(1+2\gamma)}|u - v_i|^{\xi}\,dy\,, \qquad (3.15)
$$

again for a constant $c \equiv c(n,p,\nu,L,\delta_1,H)$. In turn,

$$
\begin{aligned}
&\fint_{B_i} |D\bar v_i|^{(p-2)(1+2\gamma)}|u-v_i|^{\xi}\,dy \\
&\le c\fint_{B_i} |D\bar v_{i-1}|^{(p-2)(1+2\gamma)}|u-v_i|^{\xi}\,dy \\
&\qquad\qquad +c\fint_{B_i} |D\bar v_{i-1}-D\bar v_i|^{(p-2)(1+2\gamma)}|u-v_i|^{\xi}\,dy \\
&=: I_1+I_2\,. \qquad (3.16)
\end{aligned}
$$

To estimate I_2, we use (3.12) and hence deduce by Hölder's inequality, Sobolev embedding and (2.14) that

$$
\begin{aligned}
I_2 &\le c\lambda^{(2-p)(1+2\gamma)}\left[\fint_{B_i} |Dv_{i-1}-Dv_i|^{(p-1)(1+2\gamma)}\,dy\right]^{(p-2)/(p-1)} \\
&\qquad\cdot\left[\fint_{B_i} |u-v_i|^{\xi(p-1)}\,dy\right]^{1/(p-1)} \\
&\le c\lambda^{(2-p)(1+2\gamma)}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{(1+2\gamma)(p-2)/(p-1)} \\
&\qquad\cdot r_i^{\xi}\left[\fint_{B_i} |Du-Dv_i|^{p-1}\,dy\right]^{\xi/(p-1)} \\
&\le cr_i^{\xi}\lambda^{(2-p)(1+2\gamma)}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{[\xi+(1+2\gamma)(p-2)]/(p-1)}\,.
\end{aligned}
$$

The constant c depends on n,p,ν and δ_1; here we have again used (3.8) that ensures that $(p-1)(1+2\gamma)<p_M$ and that $\xi\le n/(n-1)$. By definitions of γ and ξ in (3.8) it readily follows that

$$
I_2^{1/\xi} \le cr_i\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right] \qquad (3.17)
$$

with $c\equiv c(n,p,\nu,\delta_1)$. Furthermore, appealing to (3.6) repeatedly, the term I_1 is first estimated by Sobolev embedding as

$$
\begin{aligned}
I_1 &\le cr_i^{\xi}\left[\fint_{B_i} |Du-Dv_i|\,dy\right]^{\xi} \\
&\le cr_i^{\xi}\left[\fint_{B_i} |D\bar v_{i-1}|^{(p-2)(1+\gamma)}|Du-Dv_i|\,dy\right]^{\xi}\,,
\end{aligned}
$$

and therefore, with $c\equiv c(n,p,H)$, we have

$$
I_1^{1/\xi} \le cr_i\fint_{B_i} |D\bar v_{i-1}|^{(p-2)(1+\gamma)}|Du-Dv_i|\,dy\,. \qquad (3.18)
$$

By using the triangle inequality we further deduce

$$\begin{aligned}\fint_{B_i} |D\bar{v}_{i-1}|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy &\leq c\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy \\ &\quad + c\lambda^{-(p-2)(1+\gamma)}\fint_{B_i} |Dv_i - Dv_{i-1}|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy\,.\end{aligned}$$

The last integral can be estimated similarly as in the case of I_2, that is by making use of (2.14) and (3.12); indeed

$$\begin{aligned}&\fint_{B_i} |Dv_i - Dv_{i-1}|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy \\ &\leq \left(\fint_{B_i} |Dv_i - Dv_{i-1}|^{(p-1)(1+\gamma)}\,dy\right)^{(p-2)/(p-1)} \\ &\qquad \cdot \left(\fint_{B_i} |Du - Dv_i|^{p-1}\,dy\right)^{1/(p-1)} \leq c\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{1+\gamma(p-2)/(p-1)}\,.\end{aligned}$$

Combining the last two inequalities and using (3.5) to estimate

$$\lambda^{-(p-2)(1+\gamma)}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^{1+\gamma(p-2)/(p-1)} \leq \lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,,$$

we obtain

$$\begin{aligned}&\fint_{B_i} |D\bar{v}_{i-1}|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy \\ &\qquad \leq c\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy + c\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,.\end{aligned}$$

Recalling (3.18) we thus end up with

$$I_1^{1/\xi} \leq cr_i\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy + cr_i\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,.$$

The last estimate with (3.17), and eventually with (3.15)-(3.16), gives

$$\begin{aligned}\alpha &\leq c\left(I_1^{1/\xi} + I_2^{1/\xi}\right) \\ &\leq cr_i\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy + cr_i\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,,\end{aligned}$$

for $c \equiv c(n,p,\nu,L,\delta_1,H)$. Inserting the last inequality into (3.13) and choosing $\varepsilon \equiv \varepsilon(n,p,\nu,L,\delta_1,H) = 1/(2c)$, we obtain

$$\begin{aligned}&\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy \\ &\qquad \leq \frac{1}{2}\fint_{B_i} |D\bar{v}_i|^{(p-2)(1+\gamma)}|Du - Dv_i|\,dy + c\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,,\end{aligned}$$

again with $c \equiv c(n, p, \nu, L, \delta_1, H)$, from which (3.9), and therefore (3.7), follows. Notice that by (3.8) the quantity we are reabsorbing here is finite.

The rest of the proof of Theorem 1.3 goes now in four steps.

Step 1: Main characters in the proof. Let $x \in \Omega$ be a Lebesgue point of Du and let $B(x, 2r) \subset \Omega$ be a ball; in the following all the balls considered will be centered at x. We start taking a positive number λ such that

$$\lambda := H_1 \fint_{B_r} (|Du| + s)\, dy + H_2 \left(\int_0^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-1}} \frac{d\varrho}{\varrho} \right)^{1/(p-1)}, \tag{3.19}$$

and fix the constants $H_1, H_2 \geq 1$ in a few lines (see (3.26) below), in a way that makes them depending only on n, p, ν, L; clearly, we may assume without loss of generality that $\lambda > 0$. In the end we shall simply prove that

$$|Du(x)| \leq \lambda \tag{3.20}$$

and, due to the fact that the ball $B(x, 2r) \subset \Omega$ is arbitrary, we shall conclude with (1.2) by taking $c := 2^{p-2} \max\{H_1, H_2\}^{p-1}$ (and changing a bit the radius r).

We begin by defining the number

$$\delta_1 := \left(\frac{1}{10^8 c_0} \right)^{1/\beta} \left(\frac{1}{10^8 c_h c_2^2} \right)^{1/\beta}, \tag{3.21}$$

where constants c_0, β, c_h and c_2 appear in Theorems 2.1-2.2 and in Lemma 2.6. It follows that δ_1 depends only on n, p, ν, L. With such a choice of δ_1 we consider the balls in (3.2) and the related comparison solutions in (3.3). Moreover, with δ_1 given in (3.21), we determine the new constant

$$c_3 := c_H(n, p, \nu, L, \delta_1, H) \qquad \text{with} \qquad H := 10^4 c_2 \tag{3.22}$$

according to Lemma 3.1. In this way we ultimately have that c_3 depends only on n, p, ν, L. We also define the composite quantity

$$C_i := \sum_{j=i-2}^{i} \fint_{B_j} (|Du| + s)\, dy + \delta_1^{-n} E(Du, B_i) \tag{3.23}$$

for every integer $i \geq 2$. We meanwhile record the following inclusions, that hold for every $i \geq 0$:

$$\delta_1 B_i = B_{i+1} \subset \left(\frac{1}{10^8 c_h c_2^2} \right)^{1/\beta} B_i \subset (1/4) B_i \subset B_i\,. \tag{3.24}$$

Next, we take k as the smallest integer satisfying

$$8 c_h (2\delta_1)^{k\beta} \leq \frac{\delta_1^n}{10^8 c_2^2} \qquad \text{and} \qquad k \geq 3\,. \tag{3.25}$$

In this way k depends only upon n, p, ν, L. We finally set

$$H_1 := 10^8 c_2 \delta_1^{-3n} \qquad \text{and} \qquad H_2 := 10^8 c_1 c_2 c_3 \delta_1^{-(k+3)n}\,, \tag{3.26}$$

where $c_1 \equiv c_1(n, p, \nu)$ has been fixed in Lemma 2.5. Notice that the choice of H_1 also implies

$$C_i \leq 5\delta_1^{-3n} \fint_{B_0} (|Du| + s)\, dy \leq \frac{\lambda}{10^7 c_2} \tag{3.27}$$

for $i = 2, 3$, together with

$$s \leq \frac{\lambda}{10^7 c_2}\,. \tag{3.28}$$

Now, observe that

$$\begin{aligned}
\int_0^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-1}} \frac{d\varrho}{\varrho} &= \sum_{i=0}^{\infty} \int_{r_{i+1}}^{r_i} \frac{|\mu|(B_\varrho)}{\varrho^{n-1}} \frac{d\varrho}{\varrho} + \int_r^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-1}} \frac{d\varrho}{\varrho} \\
&\geq \sum_{i=0}^{\infty} \frac{|\mu|(B_{i+1})}{r_i^{n-1}} \int_{r_{i+1}}^{r_i} \frac{d\varrho}{\varrho} + \frac{|\mu|(B_0)}{(2r)^{n-1}} \int_r^{2r} \frac{d\varrho}{\varrho} \\
&= \delta_1^{n-1} \log\left(\frac{1}{\delta_1}\right) \sum_{i=0}^{\infty} \frac{|\mu|(B_{i+1})}{r_{i+1}^{n-1}} + \frac{\log 2}{2^{n-1}} \left[\frac{|\mu|(B_0)}{r_0^{n-1}}\right] \\
&\geq \delta_1^{n} \sum_{i=0}^{\infty} \frac{|\mu|(B_i)}{r_i^{n-1}}\,.
\end{aligned} \tag{3.29}$$

Therefore, by (3.19) and the choice in (3.26) it follows that

$$10^8 c_1 c_2 c_3 \delta_1^{-4n} \left[\sum_{i=0}^{\infty} \frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{1/(p-1)} \leq \lambda \tag{3.30}$$

and

$$\left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{1/(p-1)} \leq \frac{\delta_1^{(k+2)n}}{10^8 c_1 c_2} \lambda \leq \lambda \qquad \text{for every } i \geq 0\,. \tag{3.31}$$

Of course we are using constants like 10^8 to emphasize the fact that in certain places of the proof what it matters is to take large/small quantities.

Step 2: The exit time index. Starting from (3.27), let us show that without loss of generality we may assume there exists an "exit time" index $i_e \geq 3$ such that

$$C_{i_e} \leq \frac{\lambda}{100}, \qquad C_j > \frac{\lambda}{100} \qquad \text{for every } j > i_e\,. \tag{3.32}$$

Indeed, on the contrary, we would have $C_{j_i} \leq \lambda/100$, for every $i \in \mathbb{N}$, for an increasing subsequence $\{j_i\}$, and then, as x is a Lebesgue point of Du, obviously

$$|Du(x)| \leq \lim_{i\to\infty} \fint_{B_{j_i}} (|Du| + s)\, dy \leq \frac{\lambda}{100}\,,$$

and the proof would be finished.

Step 3: Basic properties of the comparison functions. We here record a few basic properties of the functions $\{v_i\}$.

Lemma 3.2. *Assume that for a certain index $i \geq 0$ the following inequality holds:*

$$\fint_{B_i} (|Du| + s)\, dy \leq \lambda\,, \tag{3.33}$$

where $\lambda > 0$ is defined in (3.19). *Then*

$$|Dv_i| + s \leq c_2\lambda \qquad \text{in } B_i/2 \tag{3.34}$$

is satisfied with $c_2 \equiv c_2(n, p, \nu, L)$ appearing in Lemma 2.6. Moreover, if in addition it holds that $i \geq i_e - 2$, then we also have

$$\frac{\lambda}{10^4 c_2} \leq |Dv_i| \qquad \text{in } B_{i+1}\,. \tag{3.35}$$

Proof. The estimate (3.34) is a consequence of Lemma 2.6 applied with $B_R \equiv B_i$, thanks to (3.33) and (3.31). It remains to prove (3.35), and for this we want to apply Lemma 2.3. We start proving that

$$\frac{\lambda}{200} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} (|Dv_i| + s)\, dy\,, \tag{3.36}$$

where $k \equiv k(n, p, \nu, L)$ has been defined in (3.25). By (3.31) we apply Lemma 2.7 three times, with the choice $B_R \equiv B_i$, $\delta \in \{\delta_1^k, \delta_1^{k-1}, \delta_1^{k-2}\}$ and $\theta = \delta_1/(10^8 c_2)$; summing up the inequalities resulting from the application of (2.17) yields

$$\sum_{j=i-2+k}^{i+k} \fint_{B_j} (|Du| + s)\, dy - \frac{\lambda}{10^7 c_2} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} (|Dv_i| + s)\, dy\,.$$

The definition of C_i in (3.23) in turn gives

$$C_{i+k} - \delta_1^{-n} E(Du, B_{i+k}) - \frac{\lambda}{10^7 c_2} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} (|Dv_i| + s)\, dy\,.$$

As $i \geq i_e - 2$ and $k \geq 3$, we have $C_{i+k} \geq \lambda/100$ by (3.32); using this fact in the inequality appearing in the latest display leads to

$$\frac{\lambda}{100} - \delta_1^{-n} E(Du, B_{i+k}) - \frac{\lambda}{10^7 c_2} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} (|Dv_i| + s)\, dy\,. \tag{3.37}$$

To proceed in the estimation of the excess term we notice that

$$\fint_{B_{i+k}} |Du - Dv_i|\, dy \leq \frac{|B_i|}{|B_{i+k}|} \fint_{B_i} |Du - Dv_i|\, dy \leq c_1 \delta_1^{-kn} \left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{1/(p-1)}$$

where we have used (2.14) with $q = 1$. In turn, again using (3.31) provides us

$$\fint_{B_{i+k}} |Du - Dv_i|\, dy \leq \frac{\delta_1^n}{10^8 c_2} \lambda\,. \tag{3.38}$$

On the other hand, Theorem 2.2 (applied to v_i on B_i, choosing $B_R \equiv B_i/2$ in (2.7)), estimate (3.34), and finally (3.25), give

$$2\mathrm{osc}_{B_{i+k}} Dv_i \le 2c_h c_2 (2\delta_1)^{k\beta}\lambda \le \frac{\delta_1^n}{10^8 c_2}\lambda\,,$$

so that, thanks to (3.1) and (3.38), we can estimate

$$\begin{aligned}
E(Du, B_{i+k}) &\le 2 ⨍_{B_{i+k}} |Du - (Dv_i)_{B_{i+k}}|\, dy \\
&\le 2E(Dv_i, B_{i+k}) + 2 ⨍_{B_{i+k}} |Du - Dv_i|\, dy \\
&\le 2\mathrm{osc}_{B_{i+k}} Dv_i + \frac{\delta_1^n \lambda}{10^8 c_2} \\
&\le \frac{\delta_1^n \lambda}{10^7 c_2}\,.
\end{aligned}$$

Inserting the last estimate into (3.37) yields (3.36). In turn, (3.36) implies that there exists an index $j \in \{i-2+k, i-1+k, i+k\}$ such that

$$\frac{\lambda}{200} \le 3 ⨍_{B_j} (|Dv_i| + s)\, dy = 3 ⨍_{\delta_1^{j-i} B_i} (|Dv_i| + s)\, dy\,.$$

This inequality allows to apply Lemma 2.3 with parameters $B \equiv B_i$, $\sigma = \delta_1$, $m = j - i \ge 1$ (recall that $k \ge 3$ and observe that (3.34) is in force) and $\Gamma = 600c_2$ (keep (3.21) in mind to check (2.9)), and therefore we conclude with

$$\frac{\lambda}{2400c_2} \le |Dv_i| + s \qquad \text{in } B_{i+1} = \delta_1 B_i\,,$$

which, in turn, together with (3.28), implies (3.35).

Step 4: Final iteration and conclusion. First, a lemma.

Lemma 3.3. *Let* (3.4) *hold for a certain number* $i \ge i_e - 1$. *Then*

$$E(Du, B_{i+2}) \le \frac{1}{4} E(Du, B_{i+1}) + c_4 \lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]$$

holds for a constant c_4 depending only on n, p, ν, L, which is given by $c_4 = 4\delta_1^{-n} c_3$, and in turn $c_3 \equiv c_3(n, p, \nu, L)$ is the constant appearing in (3.22).

Proof. Let us first prove that

$$⨍_{B_{i+1}} |Du - Dv_i|\, dy \le c_3 \lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right] \tag{3.39}$$

holds, and for this we want to apply Lemma 3.1. To this aim, we repeatedly apply Lemma 3.2 to both v_i and v_{i-1}; notice that, since we are assuming $i \ge i_e - 1$, we obviously have $i - 1 \ge i_e - 2$. In conclusion, Lemma 3.2 gives

$$\frac{\lambda}{10^4 c_2} \le |Dv_{i-1}| \le c_2 \lambda \quad \text{in } B_i \quad \text{and} \quad \frac{\lambda}{10^4 c_2} \le |Dv_i| \le c_2 \lambda \quad \text{in } B_{i+1}\,.$$

Therefore assumptions (3.6) are satisfied with $H \equiv 10^4 c_2$, while (3.5) holds by (3.31), and hence (3.39) follows by Lemma 3.1.

Next, by using estimate (2.5) applied on v_i and recalling the choice of δ_1 in (3.21), we obtain

$$E(Dv_i, B_{i+2}) \le \frac{E(Dv_i, B_{i+1})}{2^6}\,.$$

This last inequality, (3.39) and (3.1) give

$$\begin{aligned}
E(Du, B_{i+2}) &\le 2 \fint_{B_{i+2}} |Du - (Dv_i)_{B_{i+2}}|\,dy \\
&\le 2E(Dv_i, B_{i+2}) + 2\fint_{B_{i+2}} |Du - Dv_i|\,dy \\
&\le \frac{E(Dv_i, B_{i+1})}{2^5} + 2\delta_1^{-n}\fint_{B_{i+1}} |Du - Dv_i|\,dy \\
&\le \frac{E(Du, B_{i+1})}{2^4} + 4\delta_1^{-n}\fint_{B_{i+1}} |Du - Dv_i|\,dy \\
&\le \frac{E(Du, B_{i+1})}{2^4} + c_4\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right].
\end{aligned}$$

To proceed with the proof of Theorem 1.3, denote in short

$$A_i := E(Du, B_i) \qquad \text{and} \qquad a_i = |(Du)_{B_i}|\,.$$

By the definition in (3.23) and (3.32) we have

$$s + \sum_{j=i_e-2}^{i_e} a_j + \delta_1^{-n} A_{i_e} \le C_{i_e} \le \frac{\lambda}{100}\,. \tag{3.40}$$

We now prove, by induction, that

$$s + a_j + A_j \le \lambda \tag{3.41}$$

holds whenever $j \ge i_e$. Indeed, by (3.40), the case $j = i_e$ of the previous inequality holds. Then, assume by induction that (3.41) holds whenever $j \in \{i_e, \ldots, i\}$. This, taking also (3.40) into account and recalling that the definition in (3.23) allows to control the averages of $|Du|$ on the three balls B_{i_e-2},B_{i_e-1} and B_{i_e}, implies that

$$\fint_{B_j} (|Du| + s)\,dy \le \lambda \qquad \text{for } j \in \{i_e - 2, \ldots, i\}\,.$$

We can thus apply Lemma 3.3, thereby getting

$$A_{j+2} \le \frac{1}{4}A_{j+1} + c_4\lambda^{2-p}\left[\frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}}\right] \tag{3.42}$$

for all $j \in \{i_e - 1, \ldots, i-1\}$. It immediately follows, by (3.41) inductively assumed for $i_e \leq j \leq i$ and (3.30), that

$$A_{i+1} \leq \frac{\lambda}{4} + c_4 \lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}} \right] \leq \frac{\lambda}{4} + \frac{\lambda}{10^3} \leq \frac{\lambda}{3}\,. \tag{3.43}$$

Furthermore, summing up (3.42) for $j \in \{i_e - 1, i_e, \ldots, i-1\}$ leads to

$$\sum_{j=i_e}^{i+1} A_j \leq A_{i_e} + \frac{1}{4} \sum_{j=i_e}^{i} A_j + c_4 \lambda^{2-p} \sum_{j=0}^{\infty} \left[\frac{|\mu|(B_j)}{r_j^{n-1}} \right] \tag{3.44}$$

in turn yielding

$$\sum_{j=i_e}^{i+1} A_j \leq 2A_{i_e} + 2c_4 \lambda^{2-p} \sum_{j=0}^{\infty} \left[\frac{|\mu|(B_j)}{r_j^{n-1}} \right]\,. \tag{3.45}$$

On the other hand, notice that

$$\begin{aligned}
a_{i+1} - a_{i_e} &= \sum_{j=i_e}^{i} (a_{j+1} - a_j) \\
&\leq \sum_{j=i_e}^{i} \fint_{B_{j+1}} |Du - (Du)_{B_j}|\, dy \\
&= \sum_{j=i_e}^{i} \frac{|B_j|}{|B_{j+1}|} E(Du, B_j)
\end{aligned}$$

and therefore (3.45) and eventually (3.30) give

$$a_{i+1} \leq a_{i_e} + 2\delta_1^{-n} A_{i_e} + 2\delta_1^{-n} c_4 \lambda^{2-p} \sum_{j=0}^{\infty} \left[\frac{|\mu|(B_j)}{r_j^{n-1}} \right] \leq 2C_{i_e} + \frac{\lambda}{10^3}\,.$$

In turn, by (3.40) the previous estimate yields $a_{i+1} \leq \lambda/3$. This inequality together with (3.28) and (3.43) allows to finally verify the induction step, i.e. $s + a_{i+1} + A_{i+1} \leq \lambda$. Therefore (3.41) holds for every $i \geq i_e$. Estimate (3.20) finally follows since x is Lebesgue point of Du and therefore

$$|Du(x)| = \lim_{i \to \infty} a_i \leq \lambda\,.$$

4. Proof of Theorem 1.5

By (1.13) and Corollary 1.4 we may assume that $Du \in L^\infty_{\rm loc}(\Omega, \mathbb{R}^n)$. In the rest of the Section we select open subsets $\Omega' \Subset \Omega'' \Subset \Omega$, and prove that Du is continuous in Ω'. Due to arbitrariness of the choice of the open subsets this is sufficient to prove Theorem 1.5. From now on we denote

$$\lambda := \|Du\|_{L^\infty(\Omega'')} + s\,. \tag{4.1}$$

We recall that we may assume that $\mu(\cdot)$ is defined on the whole $\mathbb{R}^n$ by eventually setting μ to be zero outside of Ω. In the following we shall also set $d := \operatorname{dist}(\Omega', \partial\Omega'') > 0$. We preliminary prove the VMO-regularity of Du.

Proposition 4.1. *Under the assumptions of Theorem 1.5, Du is locally* VMO*-regular in Ω. In particular, for every $\varepsilon \in (0,1)$, there exists a positive radius $r_\varepsilon \equiv r_\varepsilon(n,p,\nu,L,\lambda,\mu(\cdot),\varepsilon) < d$ such that*

$$\fint_{B(x,\varrho)} |Du - (Du)_{B(x,\varrho)}|\, dy < \lambda\varepsilon \tag{4.2}$$

holds whenever $\varrho \in (0, r_\varepsilon]$ and $x \in \Omega'$.

Proof. Let us set, for $\varepsilon \in (0,1)$ as in the statement,

$$\delta_1 := \left(\frac{\varepsilon}{10^8 c_0}\right)^{1/\beta} \left(\frac{\varepsilon^2}{10^8 c_h c_2^2}\right)^{1/\beta}\,, \tag{4.3}$$

where c_0, c_h, c_2 and β are the constants introduced in Theorems 2.1-2.2 and Lemma 2.6; they all depend only on n,p,ν,L and so the same dependence is inherited by δ_1. We then look at Lemma 3.1 and determine the constant

$$c_5 := c_H(n,p,\nu,L,\delta_1,H) \qquad \text{with} \qquad H = \frac{10^3 c_2}{\varepsilon}\,. \tag{4.4}$$

This ultimately yields $c_5 \equiv c_5(n,p,\nu,L,\varepsilon)$. We choose accordingly a positive radius $R \equiv R(n,p,\nu,L,\lambda,\mu(\cdot),\varepsilon) < d$ such that

$$\sup_{0<\varrho\le R}\ \sup_{y\in\Omega'} \left[\frac{|\mu|(B(y,\varrho))}{\varrho^{n-1}}\right]^{1/(p-1)} \le \frac{\delta_1^{3n}\lambda\varepsilon}{10^8 c_1 c_5}\,, \tag{4.5}$$

which is again possible by (1.13). Finally, we fix $x \in \Omega'$ and define the chain of shrinking intrinsic balls

$$B_i \equiv B(x, r_i)\,, \qquad r_i = \delta_1^i r, \qquad \text{where } r \in (\delta_1 R, R]\,, \tag{4.6}$$

for every integer $i \ge 0$, and define the related comparison solutions $v_i \in u + W^{1,p}_0(B_i)$ as in (3.3). Notice now that, thanks to (4.1), $B_i \subset \Omega'$ holds for every $i \in \mathbb{N}$, and therefore

$$\fint_{B_i} (|Du| + s)\, dy \le \lambda \qquad \text{for every integer } i \ge 0\,. \tag{4.7}$$

We shall prove that

$$E(Du, B_{i+2}) < \lambda\varepsilon \qquad \text{for every } i \in \mathbb{N} \cap [1, \infty)\,. \tag{4.8}$$

Let us single out an index $i \geq 1$ and let us distinguish two cases; the first is when

$$\fint_{B_{i+2}} |Du|\, dy < \frac{\lambda\varepsilon}{50}\,,$$

so that we trivially have $E(Du, B_{i+2}) < \lambda\varepsilon/25$ and (4.8) follows. The other case is fixed in the next lemma, so that (4.8) eventually follows via (4.1) and (4.5).

Lemma 4.2. *Assume that*

$$\fint_{B_{i+2}} |Du|\, dy \geq \frac{\lambda\varepsilon}{50}\,. \tag{4.9}$$

Then it holds that

$$E(Du, B_{i+2}) \leq \frac{\varepsilon}{2^4} E(Du, B_{i+1}) + 4c_5\delta_1^{-n}\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,. \tag{4.10}$$

Proof. We preliminary establish the validity of

$$\fint_{B_{i+1}} |Du - Dv_i|\, dy \leq c_5\lambda^{2-p}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]\,, \tag{4.11}$$

for $c_5 \equiv c_5(n, p, \nu, L, \varepsilon)$ as determined in (4.4). We use Lemma 3.1; as (3.4) and (3.5) are satisfied by (4.7) and (4.5), respectively, it remains to show that (3.6) is satisfied for the choice of H made above, that is $H = 10^3c_2/\varepsilon$. Specifically, we show that

$$\begin{cases} \dfrac{\lambda\varepsilon}{10^3c_2} \leq |Dv_{i-1}| \leq c_2\lambda & \text{in } B_i \\[2mm] \dfrac{\lambda\varepsilon}{10^3c_2} \leq |Dv_i| \leq c_2\lambda & \text{in } B_{i+1} \end{cases} \tag{4.12}$$

hold. By (4.5) and Lemma 2.7, applied to v_{i-1} and with $B \equiv B_{i-1}$, $\delta = \delta_1^3$ and $\theta = \varepsilon/100$, we gain

$$\fint_{B_{i+2}} (|Du| + s)\, dy - \frac{\lambda\varepsilon}{100} \leq \fint_{B_{i+2}} (|Dv_{i-1}| + s)\, dy\,,$$

and therefore (4.9) allows us to conclude with

$$\fint_{B_{i+2}} (|Dv_{i-1}| + s)\, dy \geq \frac{\lambda\varepsilon}{100}\,.$$

Observe also that by Lemma 2.6 we have $|Dv_{i-1}(x)| + s \leq c_2\lambda$ whenever $x \in B_{i-1}/2$, and in particular whenever $x \in B_i$. Keeping (4.3) in mind, we are now ready to apply the density Lemma 2.3 choosing $B \equiv B_{i-1}$, $\sigma = \delta_1$, $m = 3$ and $\Gamma = 100c_2/\varepsilon$. In this way the assertion about v_{i-1} in (4.12) is proved. Arguing in a completely similar way for v_i, and replacing B_{i-1} by B_i, we also obtain the assertions in (4.12) concerning v_i. Now, using Lemma 3.1 we gain (4.11).

Proceeding with the proof, by using estimate (2.5) applied to v_i and recalling the choice of δ_1 in (4.3), we obtain

$$E(Dv_i, B_{i+2}) \le \frac{\varepsilon}{2^6} E(Dv_i, B_{i+1}).$$

In turn, by using this inequality, together with (4.11) and (3.1), we get (4.10):

$$\begin{aligned}
E(Du, B_{i+2}) &\le 2 \fint_{B_{i+2}} |Du - (Dv_i)_{B_{i+2}}|\, dy \\
&\le 2E(Dv_i, B_{i+2}) + 2 \fint_{B_{i+2}} |Du - Dv_i|\, dy \\
&\le \frac{\varepsilon}{2^5} E(Dv_i, B_{i+1}) + 2\delta_1^{-n} \fint_{B_{i+1}} |Du - Dv_i|\, dy \\
&\le \frac{\varepsilon}{2^4} E(Du, B_{i+1}) + 4\delta_1^{-n} \fint_{B_{i+1}} |Du - Dv_i|\, dy \\
&\le \frac{\varepsilon}{2^4} E(Du, B_{i+1}) + 4c_5\delta_1^{-n} \lambda^{2-p} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right].
\end{aligned}$$

To conclude the proof of Proposition 4.1, since all the estimates above are uniform with respect to the choice of $x \in \Omega'$ and *of the initial radius* $r \in (\delta_1 R, R]$ chosen to build the chain in (4.6), we obtain (4.2) with $r_\varepsilon = \delta_1^3 R$. Indeed, let $\varrho \le \delta_1^3 R$; this means there exists an integer $m \ge 3$ such that $\delta_1^{m+1} R < \varrho \le \delta_1^m R$. Therefore we have $\varrho = \delta_1^m r$ for some $r \in (\delta_1 R, R]$ and (4.2) follows from (4.8). The proof is complete.

Proof (Proof of Theorem 1.5). We prove the continuity of Du by showing that it is the (local) uniform limit of a net of continuous maps. More precisely, we consider the family of maps, indexed in ϱ, defined by

$$x \to (Du)_{B(x,\varrho)} \qquad \text{for} \quad 0 < \varrho \le d := \operatorname{dist}(\Omega', \partial\Omega''),$$

which are obviously continuous. We shall then prove that, for every $\varepsilon > 0$, there exists a radius $r_\varepsilon \le d/2$, independent of the point $x \in \Omega'$ considered, such that

$$|(Du)_{B(x,\varrho)} - (Du)_{B(x,\rho)}| \le \lambda\varepsilon \qquad \text{for every choice of } \varrho, \rho \in (0, r_\varepsilon]. \tag{4.13}$$

The rest of the proof is indeed dedicated to show that (4.13) holds; a point worth stressing here, and that makes things different from Section 3, is that (4.13) will proved to be valid *for every* $x \in \Omega'$, and not only for Lebesgue points of Du. Therefore, all the inequalities will be derived by considering a fixed point $x \in \Omega'$ but will be uniform to this choice.

With $\varepsilon > 0$ being fixed in (4.13), we shall keep the choice of the constants δ_1 and c_5 made in (4.3) and (4.4), respectively. Next, we take a positive radius $R \le d/2$ such that (4.5) holds together with

$$\sup_y \left[\int_0^{2R} \frac{|\mu|(B(y,\varrho))}{\varrho^{n-1}} \frac{d\varrho}{\varrho}\right]^{1/(p-1)} \le \frac{\delta_1^{4n} \lambda\varepsilon}{10^8 c_1 c_5} \tag{4.14}$$

and

$$\sup_{0<\varrho\leq R}\ \sup_{y\in\Omega'} E(Du,B(y,\varrho)) \leq \frac{\delta_1^{4n}\lambda\varepsilon}{10^8}\,. \tag{4.15}$$

Notice that a computation similar to the one in (3.29) and (4.14) give

$$\left[\sum_{i=0}^{\infty}\frac{|\mu|(B_i)}{r_i^{n-1}}\right]^{1/(p-1)} \leq \frac{\delta_1^{3n}\lambda\varepsilon}{10^8 c_1 c_5}\,. \tag{4.16}$$

We shall eventually show that the radius $r_\varepsilon \equiv \delta_1^3 R$, determined via the smallness conditions (4.14)-(4.15), will work as r_ε in (4.13). With R just determined through (4.14)-(4.15), we consider the balls $\{B_i\}$ similarly to (4.6), that is $B_i \equiv B(x,r_i)$ and $r_i = \delta_1^i R$. Finally, the functions $v_i \in u + W_0^{1,p}(B_i)$ are accordingly defined exactly as in (3.3). We shall now prove that

$$|(Du)_{B_h} - (Du)_{B_k}| \leq \frac{\lambda\varepsilon}{12} \qquad \text{holds whenever } 3\leq k\leq h\,. \tag{4.17}$$

For this we consider the set $\mathcal{L}$ defined by

$$\mathcal{L} := \left\{ i\in\mathbb{N}\ :\ \fint_{B_i}|Du|\,dy < \frac{\lambda\varepsilon}{50}\right\}\,,$$

and, accordingly we then define the set

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

and call it *maximal iteration chain* of length m, starting at i. In other words, we have $\mathcal{C}_i^m = \{i,\ldots,i+m\}$ and each element of $\mathcal{C}_i^m$ but i lies outside of $\mathcal{L}$; $\mathcal{C}_i^m$ is maximal in the sense that there cannot be another set of the same type properly containing it. Obviously, such sets do not exist when $\mathcal{L}=\mathbb{N}$. In the same way we define $\mathcal{C}_i^\infty = \{j\in\mathbb{N}\ : i\leq j<\infty,\ i\in\mathcal{L},\ j\notin\mathcal{L} \text{ if } j>i\}$ as *the infinite maximal chain* starting at i. Notice that, in every case, the smallest element of such a chain always belongs to $\mathcal{L}$, being then the only one of the chain to have such a property. Moreover, we define $i_e := \min\,\mathcal{L}$. Note that we set $i_e=\infty$ if $\mathcal{L}=\emptyset$. We are now ready for the proof of (4.17); for this we need to distinguish three cases. We shall of course assume $3\leq k<h$, because otherwise (4.17) trivializes.

Case 1: $k<h\leq i_e$. This in particular applies when $i_e=\infty$, i.e when $\mathcal{L}$ is empty. Now, if $h-1>k$, then we can apply Lemma 4.2 repeatedly, and this yields

$$E(Du,B_{i+1}) \leq \frac{1}{2}E(Du,B_i) + 4c_5\delta_1^{-n}\lambda^{2-p}\left[\frac{|\mu|(B_{i-2})}{r_{i-2}^{n-1}}\right] \tag{4.18}$$

for every $i\in\{k,\ldots,h-2\}$. Summing up the previous inequalities, and making elementary manipulations - see (3.44)-(3.45) - we have

$$\sum_{i=k}^{h-1}E(Du,B_i) \leq 2E(Du,B_k) + 8c_5\delta_1^{-n}\lambda^{2-p}\sum_{i=0}^{\infty}\left[\frac{|\mu|(B_i)}{r_i^{n-1}}\right]\,.$$

Then, using (4.15)-(4.16) yields

$$\sum_{i=k}^{h-1} E(Du, B_i) \leq \frac{\delta_1^{2n}\lambda\varepsilon}{50}\,.$$

The previous inequality follows in any case as, when $h-1=k$, it is a direct consequence of (4.15). In turn, (4.17) follows since

$$\begin{aligned}
|(Du)_{B_h}-(Du)_{B_k}| &\leq \sum_{i=k}^{h-1} |(Du)_{B_{i+1}}-(Du)_{B_i}| \\
&\leq \sum_{i=k}^{h-1} -\!\!\!\!\!\!\int_{B_{i+1}} |Du-(Du)_{B_i}|\,dy \qquad (4.19)\\
&\leq \sum_{i=k}^{h-1} \frac{|B_i|}{|B_{i+1}|} E(Du, B_i) \\
&= \delta_1^{-n} \sum_{i=k}^{h-1} E(Du, B_i) \\
&\leq \frac{\lambda\varepsilon}{50}\,. \qquad (4.20)
\end{aligned}$$

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

$$|(Du)_{B_h}| \leq \frac{\lambda\varepsilon}{25} \qquad \text{and} \qquad |(Du)_{B_k}| \leq \frac{\lambda\varepsilon}{25}\,. \tag{4.21}$$

We prove the former inequality in (4.21), the proof of the latter being the same. If $h \in \mathcal{L}$, the first inequality in (4.21) follows immediately from the definition of $\mathcal{L}$. On the other hand, if $h \notin \mathcal{L}$, then, as $h \geq i_e$, it is possible to consider the maximal iteration chain $\mathcal{C}_{i_h}^{m_h}$ such that $h \in \mathcal{C}_{i_h}^{m_h}$; notice that $h > i_h$ as $h \notin \mathcal{L} \ni i_h$. Then, iterating Lemma 4.2 as done after (4.18) - i.e. replacing k by i_h - we gain the analogue of (4.20), that is $|(Du)_{B_h}-(Du)_{B_{i_h}}| \leq \lambda\varepsilon/50$. In turn, using that $|(Du)_{B_{i_h}}| \leq \lambda\varepsilon/50$ as $i_h \in \mathcal{L}$, we again obtain the first inequality in (4.21) and in any case (4.21) follows. Estimating as

$$|(Du)_{B_h}-(Du)_{B_k}| \leq |(Du)_{B_h}| + |(Du)_{B_k}| \leq \frac{\lambda\varepsilon}{25} + \frac{\lambda\varepsilon}{25} \leq \frac{\lambda\varepsilon}{12} \tag{4.22}$$

we have that (4.17) holds in the second case too.

Case 3: $k < i_e < h$. Here we prove that (4.21) still holds and then we conclude as in Step 2. Indeed, the first inequality in (4.21) follows as in Case 2. As for the second estimate in (4.21), let us remark that, as $i_e \in \mathcal{L}$, we have that

$$|(Du)_{B_{i_e}}| \leq \frac{\lambda\varepsilon}{50}\,. \tag{4.23}$$

On the other hand, we can argue exactly as in Case 1, i.e. this time replacing h by i_e, thereby obtaining $|(Du)_{B_{i_e}}-(Du)_{B_k}| \leq \lambda\varepsilon/50$ that together with (4.23) gives the second inequality in (4.21). In turn, (4.17) follows also in this case as in (4.22).

We are now ready to show (4.13), thereby concluding the proof. As already mentioned, we take $r_\varepsilon = \delta_1^3 R$, and then fix $0 < \rho < \varrho \le r_\varepsilon$. This means that there exist two integers, $3 \le k \le h$, such that

$$\delta_1^{k+1} R < \varrho \le \delta_1^k R \qquad \text{and} \qquad \delta_1^{h+1} R < \rho \le \delta_1^h R\,. \tag{4.24}$$

Applying (4.15) we get

$$\begin{aligned}
|(Du)_{B(x,\varrho)} - (Du)_{B_{k+1}}| &\le \fint_{B_{k+1}} |Du - (Du)_{B(x,\varrho)}|\, dy \\
&\le \frac{|B(x,\varrho)|}{|B_{k+1}|} \fint_{B(x,\varrho)} |Du - (Du)_{B(x,\varrho)}|\, dy \\
&\le \delta_1^{-n} E(Du, B(x,\varrho)) \\
&\le \frac{\lambda\varepsilon}{10}\,,
\end{aligned}$$

and in the same way

$$|(Du)_{B(x,\rho)} - (Du)_{B_{h+1}}| \le \frac{\lambda\varepsilon}{10}\,.$$

Using the last two inequalities together with (4.17) establishes (4.13).

Proof (Proof of Corollary 1.6). The proof follows by the arguments in [12, Theorem 1.3] and related references; it indeed follows that if $\mu \in L(n,1)$ holds locally in Ω, then (1.13) is satisfied.

By carefully inspecting the proof of Proposition 4.1 we gain the following:

Theorem 4.3. *Let $u \in W^{1,p}(\Omega)$ be a weak solution to the equation* (1.10) *under the assumptions of Theorem 1.3. If $\mathbf{I}_1^{|\mu|}(x,R)$ is locally bounded in Ω for some $R > 0$ and if*

$$\lim_{R\to 0} \frac{|\mu|(B(x,R))}{R^{n-1}} = 0 \ \textit{ locally uniformly in } \Omega \textit{ w.r.t. } x,$$

then Du is locally VMO*-regular in Ω.*

5. General measure data problems and extensions

Theorems 1.1 and 1.3 have been stated for classical energy distributional solutions, that is $u \in W^{1,p}(\Omega)$. They nevertheless hold for solutions to measure data problems, which in turn do not always belong to $W^{1,p}$. More precisely, Theorems 1.1 and 1.3 extend to the so-called SOLA (Solution Obtained as Limits of Approximations), which are solutions to the Dirichlet problems of the type

$$\begin{cases} -\mathrm{div}\; a(Du) = \mu & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{5.1}$$

obtained as limits of solutions u_k of similar problems, where this time $\mu \equiv \mu_k$ is a smooth function, typically obtained by a convolution. Here, however, we will not recall the whole construction, which is rather long and anyway available in several papers. The original approach has been introduced by Boccardo & Gallouet in [1, 2] to which we refer for the basic results and definitions (including that of SOLA) together with [7], while for applications of (1.2) in this setting we refer to [30, Section 4] and [9, Section 5]. We confine ourselves to report the main outcome, that is

Theorem 5.1. *Let $u \in W^{1,p-1}(\Omega)$ be a* SOLA *to* (5.1)*, under the assumptions* (1.11)*, with μ being a Borel measure with finite total mass and $\Omega \subset \mathbb{R}^n$ being a Lipschitz domain. Then there exists a constant c, depending only on n, p, ν, L such that the pointwise estimate* (1.12) *holds whenever $B(x,R) \subseteq \Omega$ and $x \in \Omega$ is a Lebesgue point of Du.*

We recall that SOLA are not known in general to be unique but under a few additional assumptions (for instance $p = 2$, or when $\mu \in L^1(\Omega)$). We refer to [15,24] for more results concerning the solvability of measure data problems.

We finally remark that Theorem 1.3 extends to more general equations with coefficients of the type

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

where the main assumption on the x-dependence is that the partial map

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

is Dini-continuous, uniformly with respect to $z \in \mathbb{R}^n$. This will be shown in a forthcoming paper.

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.

References

1. Boccardo L. & Gallouët T.: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* 87 (1989), 149–169.
2. Boccardo L. & Gallouët T.: Nonlinear elliptic equations with right-hand side measures. *Comm. PDE* 17 (1992), 641–655.
3. Cianchi A.: Maximizing the L^∞ norm of the gradient of solutions to the Poisson equation. *J. Geom. Anal.* 2 (1992), 499–515.
4. Cianchi A. & Maz'ya V.: Global Lipschitz regularity for a class of quasilinear elliptic equations. *Comm. PDE* 36 (2011), 100–133.
5. DiBenedetto E.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal. TMA* 7 (1983), 827–850.
6. DiBenedetto E. & Manfredi J.J.: On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems. *Amer. J. Math.* 115 (1993), 1107–1134.

7. 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.
8. Daskalopoulos P. & Kuusi T. & Mingione G.: Borderline estimates for fully nonlinear elliptic equations. arXiv:1205.4799 (2012).
9. Duzaar F. & Mingione G.: Gradient estimates via non-linear potentials. *Amer. J. Math.* 133 (2011), 1093–1149.
10. Duzaar F. & Mingione G.: Gradient estimates via linear and nonlinear potentials. *J. Funct. Anal.* 259 (2010), 2961–2998.
11. Duzaar F. & Mingione G.: Gradient continuity estimates. *Calc. Var. & PDE* 39 (2010), 379–418.
12. Duzaar F. & Mingione G.: Local Lipschitz regularity for degenerate elliptic systems. *Ann. Inst. H. Poincaré Anal. Non Linèaire* 27 (2010), 1361–1396.
13. Grafakos L.: *Classical and modern Fourier analysis.* Pearson Edu. Inc., Upper Saddle River, NJ, 2004.
14. Hedberg L. & Wolff Th. H.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* 33 (1983), 161–187.
15. Heinonen J. & Kilpeläinen T. & Martio O.: *Nonlinear potential theory of degenerate elliptic equations.* Oxford Mathematical Monographs., New York, 1993.
16. Iwaniec T.: Projections onto gradient fields and L^p-estimates for degenerated elliptic operators. *Studia Math.* 75 (1983), 293–312.
17. Jaye B. & Verbitsky I.: Local and Global Behaviour of Solutions to Nonlinear Equations with Natural Growth Terms. *Arch. Rat. Mech. Anal.* 204 (2012), 627–681.
18. Kilpeläinen T. & Malý J.: Degenerate elliptic equations with measure data and nonlinear potentials. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 19 (1992), 591–613.
19. Kilpeläinen T. & Malý J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* 172 (1994), 137–161.
20. Korte R. & Kuusi T.: A note on the Wolff potential estimate for solutions to elliptic equations involving measures. *Adv. Calc. Var.* 3 (2010), 99–113.
21. Kuusi T. & Mingione G.: Universal potential estimates. *J. Funct. Anal.* 262 (2012), 4205–4269.
22. 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.
23. Lieberman G. M.: Sharp forms of estimates for subsolutions and supersolutions of quasilinear elliptic equations involving measures. *Comm. PDE* 18 (1993), 1191-1212.
24. Lindqvist P.: On the definition and properties of p-superharmonic functions. *J. reine angew. Math. (Crelles J.)* 365 (1986), 67–79.
25. Lindqvist P.: Notes on the p-Laplace equation. *Univ. Jyväskylä, Report* 102, (2006).
26. Lindqvist P. & Manfredi J.J.: Note on a remarkable superposition for a nonlinear equation. *Proc. AMS* 136 (2008), 133–140.
27. Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Differential Equations* 76 (1988), 203–212.
28. Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis*. University of Washington, St. Louis.
29. Mikkonen P.: On the Wolff potential and quasilinear elliptic equations involving measures. *Ann. Acad. Sci. Fenn. Math. Diss.* 104 (1996).
30. Mingione G.: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)* 6 (2007), 195–261.
31. Mingione G.: Gradient estimates below the duality exponent. *Math. Ann.* 346 (2010), 571–627.
32. Mingione G.: Gradient potential estimates. *J. Europ. Math. Soc.* 13 (2011), 459–486.
33. Mingione G.: Nonlinear aspects of Calderón-Zygmund theory. *Jahres. Deut. Math. Verein.* 112 (2010), 159–191.
34. Phuc N. C. & Verbitsky I. E.: Quasilinear and Hessian equations of Lane-Emden type. *Ann. of Math. (II)* 168 (2008), 859–914.

35. Phuc N. C. & Verbitsky I. E.: Singular quasilinear and Hessian equations and inequalities. *J. Funct. Anal.* 256 (2009), 1875–1906.
36. Trudinger N.S. & Wang X.J.: Hessian measures. I. Dedicated to Olga Ladyzhenskaya. *Topol. Methods Nonlinear Anal.* 10 (1997), 225–239.
37. Trudinger N.S. & Wang X.J.: Hessian measures. II. *Ann. of Math. (2)* 150 (1999), 579–604.
38. 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.
39. Trudinger N.S. & Wang X.J.: Quasilinear elliptic equations with signed measure data. *Disc. Cont. Dyn. Systems A* 23 (2009), 477–494.

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

and

Dipartimento di Matematica
Università di Parma
Parco Area delle Scienze 53/a, I-43124, Parma
email:giuseppe.mingione@math.unipr.it