RIESZ POTENTIAL ESTIMATES FOR A GENERAL CLASS OF QUASILINEAR EQUATIONS

PAOLO BARONI

ABSTRACT. We consider solutions to nonlinear elliptic equations with measure data and general growth and ellipticity conditions as considered in Lieberman [Comm. PDE, 1991] and we prove pointwise gradient bounds for solutions in terms of linear Riesz potentials. As a direct consequence, we get optimal conditions for the continuity of the gradient.

1. INTRODUCTION AND STATEMENT OF RESULTS

Starting from the fundamental works of Kilpeläinen and Malý [21, 22], it is known that the local behavior of solutions to nonlinear elliptic equations with right-hand side measures can be controlled in terms of nonlinear Wolff potentials. In particular, for the classic p-Laplace equation

$$-\operatorname{div}(|Du|^{p-2}Du) = \mu \qquad \text{in } \mathbb{R}^n, \qquad p > 1, \tag{1.1}$$

pointwise estimates for solutions can be obtained via Wolff potentials, in the case the measure is positive, as follows:

$$|u(x_0)| \leq c(n,p)\, \mathbf{W}^{\mu}_{1,p}(x_0,\infty) = c \int_0^\infty \left[\frac{|\mu|(B_\varrho(x_0))}{\varrho^{n-p}}\right]^{\frac{1}{p-1}} \frac{d\varrho}{\varrho}\,. \tag{1.2}$$

The above estimate is sharp, in the sense that the Wolff potential appearing in the right-hand side cannot be replaced by any other potential whatsoever, and this is the consequence of the fact that the same potential bounds the pointwise values of u also from below.

In the recent paper [26], see also [29, 36], Kuusi and Mingione proved a pointwise estimate for gradient of solutions to (1.1) in terms of the linear, Riesz potential of the right-hand side, for $p \geq 2$:

$$|Du(x_0)|^{p-1} \leq c(n,p)\, \mathbf{I}^{|\mu|}_1(x_0,\infty) \qquad \text{for any } x_0 \in \mathbb{R}^n; \tag{1.3}$$

we recall that the linear Riesz potential of the measure $|\mu|$ is defined by

$$\mathbf{I}^{|\mu|}_1(x_0,\infty) := \int_0^\infty \frac{|\mu|(B_\varrho(x_0))}{\varrho^{n-1}} \frac{d\varrho}{\varrho}.$$

Estimate (1.3) surprisingly extends to the case of the p-Laplacian the classical gradient estimates that hold for the Poisson equations. These are, in turn, a straightforward consequence of the representation formula for the solution, a tool that is obviously unavailable in the nonlinear case. On the other hand, (1.3) tells us that, when considering pointwise gradient estimates for the p-Laplacian equation, different potentials come into the play. The reader might examine in this regard also the interpolation estimates in [23, 29] which clarify in which sense and at which extent linear and non-linear potentials are related to the different levels of regularity of solution to (1.1) in a bounded domain Ω.

Date: May 28, 2014.
2010 *Mathematics Subject Classification.* Primary: 35J62, Secondary: 35J70, 35B65.
Key words and phrases. Riesz potentials, general growth conditions, measure data, gradient continuity criterion.

This phenomenon is *formally related* to the operation of *inversion of the divergence operator*. Indeed, if for just a moment we imagine that a Riesz type potential can be used to invert the divergence operator, then (1.3) formally follows by applying the Riesz potential to both sides of (1.1); indeed, we would get

$$|v(x_0)| = |Du(x_0)|^{p-1} \lesssim \mathbf{I}_1^{|\mu|}(x_0,\infty) \qquad \text{where} \qquad v := |Du|^{p-2}Du. \tag{1.4}$$

Clearly this *rough heuristic* has to be made rigorous, and this has been accomplished in [26]. On the other hand, using the same argument, in order to get pointwise estimates for u we should again formally invert the differentiation operator:

$$|Du(x_0)| \lesssim \Big[\mathbf{I}_1^{|\mu|}(x_0,\infty)\Big]^{\frac{1}{p-1}} \quad \Rightarrow \quad |u(x_0)| \lesssim \mathbf{I}_1^{|\mu|}\Big(\big[\mathbf{I}_1^{|\mu|}(\cdot,\infty)\big]^{\frac{1}{p-1}}\Big)(x_0,\infty)$$

and by a classic result the latter potential, so called Havin-Maz'ya one [18], differs from $\mathbf{W}^{\mu}_{1,p}(x_0,\infty)$ just by constants.

In this paper we are interested in understanding for which classes of quasilinear equations this principle still works. We show that it does for equations like

$$-\operatorname{div}\left(\frac{g\big(|Du|\big)}{|Du|}Du\right) = \mu \qquad \text{in } \Omega, \tag{1.5}$$

where $\Omega \subset \mathbb{R}^n$, $n \geq 2$ is a bounded open set and μ is a Borel measure with finite total mass, $|\mu|(\Omega) < \infty$; for this equation, once considering appropriate but classic assumptions on the function g, pointwise gradient estimates in terms of linear Riesz potentials hold true. In order to give some more details, we introduce the positive function $g \in C^1(0,\infty)$, satisfying only the following bound:

$$\delta \leq \frac{tg'(t)}{g(t)} \leq g_0, \qquad \text{for } t > 0, \qquad \text{with } 1 \leq \delta \leq g_0. \tag{1.6}$$

Equations as in (1.5) have been introduced and studied by Lieberman [31] and they are, in his own words "the natural, and, in a sense, the best generalization of the p-Laplace equation" (where the function g takes the power-like form $g(t) = t^{p-1}$, with $p > 1$), seen as Euler equation for local minimizers of the functional

$$\mathcal{G}(Dw) := \int_\Omega G\big(|Dw|\big)\,dx, \tag{1.7}$$

with $G(t) = t^p$ (or, more in general, $G(t) = (s^2+t^2)^{p/2}$, $p > 1$, $s \geq 0$). The p-Dirichlet energy provides a model case for the growth and ellipticity conditions considered by Ladyzhenskaya and Ural'tseva; hence, it becomes natural to investigate about more general forms of convex densities $G(\cdot)$ than power-like ones, and related Euler equations. Notice that if G is smooth enough, then the Euler equation of minimizers of (1.7) is (1.5) with $g = G'$ and $\mu = 0$. In Lieberman's work [31] a full basic regularity theory (local boundedness and Hölder regularity both of solutions and their gradients, Harnack's inequalities and characterizations of De Giorgi classes) is proved for this class of equations, essentially only supposing (1.6), for different bounds on δ; we will exploit extensively his work. Natural examples of functions g satisfying (1.6) are the logarithmic perturbations of powers, i.e.

$$g(t) = t^{p-1}\big[\log(a+t)\big]^\alpha, \qquad p \geq 2, \quad a \geq 1, \quad \alpha \geq 0.$$

Another interesting examples, given in [31] and related to (p,q)-growth conditions, are given by appropriate gluing of the monomials $t^{\alpha_1}, t^{\alpha_1-\varepsilon}, t^{\alpha_2+\varepsilon}$ for $\varepsilon < \alpha_1 < \varepsilon_2$; it turns out that is this case $\delta = \alpha_1 - \varepsilon$ and $g_0 = \alpha_2 + \varepsilon$. In this paper we thus prove the following result:

Theorem 1.1. *Let $u \in W^{1,G}(\Omega)$ be a weak solution to*

$$-\operatorname{div} a(Du) = \mu, \tag{1.8}$$

where μ is a Radon measure with finite total mass or a function in $L^1(\Omega)$, where the vector field $a(\cdot)$ is modeled upon on (1.5), in the sense specified by (2.1), and where the function g satisfies (1.6) and the non-degeneracy condition (3.2). Then there exists a constant c, depending on n, ν, L, δ, g_0, such that the pointwise estimate

$$g\big(|Du(x_0)|\big) \le c\, \mathbf{I}_1^{|\mu|}(x_0, 2R) + c\, g\bigg(\fint_{B_R(x_0)} |Du|\, dx\bigg) \tag{1.9}$$

holds for every $x_0 \in \Omega$ Lebesgue's point of Du and for every ball $B_{2R}(x_0) \subset \Omega$.

We recall that the truncated, linear Riesz potential $\mathbf{I}_1^{|\mu|}(x, R)$ is naturally defined by

$$\mathbf{I}_1^{|\mu|}(x_0, R) := \int_0^R \frac{|\mu|(B_\varrho(x_0))}{\varrho^{n-1}} \frac{d\varrho}{\varrho};$$

in the case $\mu \in L^1(\Omega)$ we denote

$$|\mu|(B_\varrho(x_0)) := \int_{B_\varrho(x_0)} |\mu|\, dx.$$

Once having the *a priori* potential bound (1.9) at hand (see Section 7 for more about our choice of giving this form to the statements), the following Corollary follows in a straightforward way:

Corollary 1.2. *Let $u \in W^{1,G}(\Omega)$ be a weak solution to (1.8), where the vector field $a(\cdot)$ satisfies the assumptions (2.1) and where g satisfies (1.6). Then*

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

Moreover the following local estimate holds true:

$$\|Du\|_{L^\infty(B_{R/2})} \le c\, g^{-1}\bigg(\big\|\mathbf{I}_1^{|\mu|}(\cdot, R)\big\|_{L^\infty(B_R)}\bigg) + c \fint_{B_R} |Du|\, dx,$$

for every ball $B_R \subset \Omega$ and with constant depending upon n, ν, L, δ, g_0.

It is worth to remark that this shows that the classic, sharp Riesz potential criterium implying the Lipschitz continuity of solutions to the Poisson equations still remains valid when considering operators of the type (1.8), at least for $\delta \ge 1$. We stress however that global Lipschitz regularity for solution to (1.5) in the vectorial case has been proved by Cianchi and Maz'ya [9, 10], under weaker assumptions on the function g: they indeed prove, using our notation,

$$0 < \delta \le \frac{tg'(t)}{g(t)} \le g_0, \quad \text{and} \quad \mu \in L(n,1)(\Omega) \quad \Longrightarrow \quad Du \in L^\infty(\Omega)$$

under a very mild condition on $\partial\Omega$. $L(n,1)$ is the Lorentz space defined just few lines below, and the inclusion $\mu \in L(n,1)(\Omega)$ actually implies that $\mathbf{I}_1^{|\mu|}(\cdot, R) \in L^\infty(\Omega)$, see the proof of Corollary 1.4. Our local L^∞ bound, however, actually opens the way to gradient continuity statements. We indeed have the following general criterion:

Theorem 1.3. *Let $u \in W^{1,G}(\Omega)$ be as in Theorem 1.1 and suppose that*

$$\lim_{R \to 0} \mathbf{I}_1^{|\mu|}(\cdot, R) = 0 \qquad \textit{locally uniformly in } \Omega \textit{ with respect to } x; \tag{1.10}$$

then Du is continuous in Ω.

The previous theorem allows to extend to our setting a few classical results. The first is a classic theorem of Stein [40] that tells that

$$\triangle u \in L(n,1)(\Omega) \qquad \Longrightarrow \qquad Du \in C^0(\Omega). \tag{1.11}$$

We recall now that the Lorentz space $L(n,1)(\Omega)$ consists in all measurable functions μ such that

$$\int_0^\infty \big|\{x \in \Omega : |\mu(x)| > \lambda\}\big|^{1/n}\, d\lambda < \infty;$$

its local variant is defined in the usual way. Note that if μ belongs locally to $L(n,1)(\Omega)$, then not just $\mathbf{I}_1^{|\mu|}(\cdot, R) \in L^\infty(\Omega)$, but also (1.10) holds true, see again the proof of the forthcoming Corollary 1.4 in Section 7. The sharpness of (1.11) has been shown in [5].

The second one result we are going to cover has been proved by Lieberman in [32] and asserts that, when considering elliptic equations as (1.1) in Ω, the density condition

$$\mu(B_\rho(x)) \le c\,\rho^{n-1+\delta}, \qquad \text{for any } B_\rho(x) \subset \Omega \text{ and for some } \delta > 0,$$

implies the local Hölder continuity of the gradient (see Kilpeläinen, [20], for an analogous statement concerning u rather than Du). We have the following

Corollary 1.4. *Let $u \in W^{1,G}(\Omega)$ be as in Theorem 1.1. If one of the following two conditions holds:*

1. *$\mu \in L(n,1)$ locally in Ω,*
2. *$|\mu|(B_R(x)) \le c\,R^{n-1}h(R)$, for some constant $c \ge 1$ and for any ball $B_R(x) \subset \Omega$, being h Dini continuous in the sense that $\displaystyle\int_0 h(\rho)\frac{d\rho}{\rho} < \infty$,*

then Du is continuous in Ω.

This Corollary (and the previous Theorem 1.3) shows in particular that the gradient continuity result in (1.11) still holds true if we replace the linear, Laplace operator with a much more general one, and, quite surprisingly, the structure of the operator does not influence the sharp condition on the right-hand side, which is actually independent on the form of the operator itself. This fact is essentially encoded in the following observation, again formal and similar to the one in (1.4): the correct quantity to consider, once treating equations as (1.5), is

$$\tilde v := \frac{g\big(|Du|\big)}{|Du|}Du.$$

Therefore, if in order to prove continuity of Du we instead prove the continuity of $\tilde v$, it then becomes clear that the condition to be imposed on μ has to be independent of the vector field. A similar argument has been developed in detail in particular in the recent papers [28] by Kuusi and Mingione, where it is proved that the condition $\mu \in L(n,1)(\Omega)$ is a sufficient one for the continuity of the gradient of solutions to p-Laplacian type systems of quasi-diagonal type.

To conclude, we first refer the reader to the classic book [19] for more on the relation between Wolff potentials and local behavior of solution to (1.1) and we mention the paper [33] where a zero-order estimate, analogue to (1.2), has been proved for minimizers of functionals satisfying growth conditions related to g; the papers [24, 27] contain parabolic potential estimates in terms of Riesz potentials and sharp continuity results similar to those just described, but in the evolutionary setting; we also mention [23] where, as well as the aforementioned interpolation potential estimates in terms of linear and nonlinear potentials together with sharp criteria for Hölder regularity for solutions, the reader can find also gradients Hölder estimates for solutions to measure data problems. We finally highlight the paper [35], where a low-order regularity theory in term of Riesz potential for nonlinear equations is developed, in the sense that integrability properties analogue to those implied by (1.3) are studied when the regularity of the vector field is so weak (for instance, when the p-Laplacian equation is endowed with merely bounded and measurable coefficients) to just allow for gradient estimates in terms of Riesz potentials at the level of measures of super-level sets.

Finally note that Theorem 1.1, which is given in the form of *a priori estimate*, actually extends to the case when u is a particular *very weak solution*, which does not necessarily belong to the energy space $W^{1,G}(\Omega)$; lack of integrability is indeed typical when dealing with measure data problems, see [2, 3]. This extension goes via a standard approximation argument briefly described in the last Section.

2. Assumptions and notation

We shall consider a differentiable vector field satisfying the ellipticity and growth conditions

$$\begin{cases} \langle \partial a(z)\lambda, \lambda \rangle \geq \nu \, \dfrac{g(|z|)}{|z|} |\lambda|^2 \\ |a(z)| + |\partial a(z)||z| \leq L g(|z|) \end{cases} \tag{2.1}$$

for all $z, \lambda \in \mathbb{R}^n$ and with $0 < \nu \leq 1 \leq L < \infty$. ∂a denotes the gradient of a with respect the variable z and g is the function considered in the preceding Section and satisfying (1.6) and the forthcoming non-degeneracy conditions (3.2). By defining

$$V_g(z) := \left[\frac{g(|z|)}{|z|} \right]^{1/2} z \tag{2.2}$$

we have the analog of a well-known quantity in the study of the p-Laplacian operator, and also in our case the following relation holds:

$$\left| V_g(z_1) - V_g(z_2) \right|^2 \approx \frac{g(|z_1| + |z_2|)}{|z_1| + |z_2|} |z_1 - z_2|^2 \approx g'\big(|z_1| + |z_2|\big) |z_1 - z_2|^2. \tag{2.3}$$

We introduce here a notation which we find pretty convenient; we shall use it many times through the whole paper. By writing $A \lesssim B$ we will mean that there exists a positive constant $\tilde{c}$, depending only upon δ and/or g_0, such that $A \leq \tilde{c} B$. With the expression $A \approx B$ we will mean that both $A \lesssim B$ and $B \lesssim A$ hold. Moreover in the case the constant $\tilde{c}$ will depend also on other quantities, we will write them below these signs. For example, if $A \leq \tilde{c}(n, \delta, g_0)\, B$, we shall write $A \lesssim_n B$. This notation will show to be very useful, besides lightening notation, since it will also highlight how (1.6) plays a fundamental role in our proofs; therefore it will be used mainly for equivalences of functions. For example, using (1.6), we have

$$g(t) \approx f(t) := \int_0^t \frac{g(s)}{s} \, ds, \tag{2.4}$$

being $f(\cdot)$ convex (see (3.1)) while $g(\cdot)$ is not; note that also $f(\cdot)$ satisfies (1.6). Using $(2.1)_1$ it is easy to prove, or see [12, Lemma 20], the following monotonicity inequality

$$\begin{aligned} \langle a(z_1) - a(z_2), z_1 - z_2 \rangle &\gtrsim c(\nu) g'\big(|z_1| + |z_2|\big) |z_1 - z_2|^2 \\ &\gtrsim c \, |V_g(z_1) - V_g(z_2)|^2 \end{aligned} \tag{2.5}$$

and the Lipschitz continuity

$$|a(z_1) - a(z_2)| \lesssim c(L) g\big(|z_1| + |z_2|\big) |z_1 - z_2|.$$

More on notation. In the following we shall adopt the customary convention of denoting by c a constant, *always larger than one*, that may vary from line to line; peculiar dependencies on parameters will be properly emphasized in parentheses when needed, sometimes just at the end of the chains of equations, for the sake of readability. Special occurrences will be denoted by special symbols, such as $c_1, c_2, \tilde{c}, c_*$. We again stress that we will try to use the "$\lesssim$" notation mainly to highlight equivalence of functions.

$B_R(x_0)$ will be the open ball with center x_0 and radius R. We shall often avoid to write the center of the balls when no ambiguity will arise: often the reader will read $B_R \equiv B_R(x_0)$ or the like. Being $C \in \mathbb{R}^n$ a measurable set with positive measure and $\ell : C \to \mathbb{R}^k$ an integrable map, $k \in \mathbb{N}$, we denote with $(\ell)_C$ the averaged integral

$$(\ell)_C := \fint_C \ell(x)\, dx := \frac{1}{|C|} \int_C \ell(x)\, dx.$$

Again for ℓ and C as above, the (L^1-)excess functional $E(\ell, C)$ is defined as

$$E(\ell, C) := \fint_C \big|\ell - (\ell)_C\big|\, dx. \tag{2.6}$$

Note that

$$E(\ell, C) \le 2 \fint_C |\ell - \xi|\, dx \qquad \text{for all } \xi \in \mathbb{R}^k \tag{2.7}$$

is a useful property which will be often used, despite not always being made explicit. We use the notation

$$\int_0 l(s)\, ds < \infty \qquad \text{for } l : (0, \infty) \to (0, \infty) \text{ continuous function}$$

to mean $\int_0^r l(s)\, ds < \infty$ for some (and then for all) $r > 0$. Similarly for $\int_0 l(s)\, ds = \infty$; the same for improper integrals at infinity $\int^\infty$. By $\mathbb{R}^+$ we will mean the open half-line $(0, \infty)$, by $\mathbb{N}$ the set $\{1, 2, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

3. Basic properties of the g function and Orlicz-Sobolev spaces.

First of all we note that we can suppose without loss of generality $g_1 > 1$ and that the lower bound $\mathcal{O}_G(t) \ge 1$ implies not only that g is monotone, but also that

$$t \to \frac{g(t)}{t} \qquad \text{is increasing;} \tag{3.1}$$

the proof is a simple computation of its derivative and this implies that $g(0) = 0$. Moreover we shall assume the degeneracy conditions

$$\lim_{t \to 0^+} \frac{g(t)}{t} = 0, \qquad \lim_{t \to \infty} \frac{g(t)}{t} = \infty; \tag{3.2}$$

this is to say that as the gradient vanishes, the modulus of ellipticity of the equation becomes zero and that the equation is not asymptotically non-degenerate. We make this choice in order to simplify the (already technically heavy) presentation, still considering a case that in many respects can be considered as the most interesting one. Indeed, if the second assumption fails, for instance, results and techniques for asymptotically elliptic operators may be applied, see for instance [4, 25, 39].

We can also assume the normalization condition

$$\int_0^1 g(s)\, ds = 1. \tag{3.3}$$

At this point elementary calculus shows that $g(1) \approx 1$. Note now that under the only assumption (1.6) a simple computation of derivatives gives that $t \mapsto t^{-\delta} g(t)$ is increasing and $t \mapsto t^{-g_0} g(t)$ is decreasing, so we have

$$\min\{\alpha^\delta, \alpha^{g_0}\}\, g(t) \le g(\alpha t) \le \max\{\alpha^\delta, \alpha^{g_0}\}\, g(t) \qquad \text{for all } t \ge 0, \alpha \ge 0. \tag{3.4}$$

Define now the function $G \in C^2(0,+\infty)$ as the primitive of g:

$$G(t) := \int_0^t g(s)\,ds.$$

It is straightforward to see that $G(1) = 1$ from (3.3), G is convex and there holds

$$\frac{tg(t)}{1+g_0} \le G(t) \le \frac{tg(t)}{1+\delta} \qquad \text{if } t \ge 0. \tag{3.5}$$

Inequality (3.5) can be rewritten in a more expressive way as

$$1+\delta \le \frac{tG'(t)}{G(t)} \le 1+g_0 :$$

as above, this implies that the function $t^{-(1+\delta)}G(t)$ is increasing and that $t^{-(1+g_0)}G(t)$ is decreasing, so we have

$$\min\{\alpha^{1+\delta},\alpha^{1+g_0}\}\,G(t) \le G(\alpha t) \le \max\{\alpha^{1+\delta},\alpha^{1+g_0}\}\,G(t) \tag{3.6}$$

for all $t \ge 0, \alpha \ge 0$. In the customary terminology, the right-hand side inequalities of (3.4) and (3.6) mean that g and G satisfy a global Δ_2-condition. We recall that a function $A : \mathbb{R}^+ \to \mathbb{R}^+$ is said to satisfy a *global* Δ_2*- (or doubling) condition* if there exists a constant $k \ge 1$ such that

$$A(2t) \le k\,A(t) \quad \text{for all } t \ge 0.$$

Note the peculiar form of (3.6) when $t = 1$ (and then $G(1) = 1$). Being G strictly increasing and with infinite limit, then G^{-1} exists, is defined for all $t \in \mathbb{R}$ and it is strictly increasing. Replacing t with $G^{-1}(t)$ and α with $\alpha^{1+\delta}$ (respectively, α^{1+g_0}), (3.6) implies

$$\min\{\alpha^{\frac{1}{1+\delta}},\alpha^{\frac{1}{1+g_0}}\}\,G^{-1}(t) \le G^{-1}(\alpha t) \le \max\{\alpha^{\frac{1}{1+\delta}},\alpha^{\frac{1}{1+g_0}}\}\,G^{-1}(t) \tag{3.7}$$

for $\alpha, t \ge 0$; something similar holds for g. We shall use this estimate, as also (3.6) and (3.4), mainly with the purpose of confining constants outside the Young functions we are going to consider.

Remark 3.1. Note that for an increasing function $f : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying a doubling Δ_2 condition $f(2t) \lesssim f(t)$ for $t \ge 0$, it is easy to prove that $f(t+s) \lesssim f(t) + f(s)$ holds for every $t, s \ge 0$. Indeed $f(t+s) \le f(2t) + f(2s)$. Analogue estimates hold, e.g., if $f(2t) \lesssim_n f(t)$ or similar conditions.

The Remark above shows that both g and G satisfy the following subadditivity property:

$$G(t+s) \lesssim G(t) + G(s), \qquad g(t+s) \lesssim g(t) + g(s). \tag{3.8}$$

Finally note that, since we have at hand monotonicity (3.1), then

$$\begin{aligned} G\big(|Du-Dv|\big) &\lesssim \frac{g\big(|Du-Dv|\big)}{|Du-Dv|}|Du-Dv|^2 \lesssim \frac{g\big(|Du|+|Dv|\big)}{|Du|+|Dv|}|Du-Dv|^2 \\ &\lesssim c\,\big|V_g(Du)-V_g(Dv)\big|^2. \end{aligned} \tag{3.9}$$

3.1. **Young functions, N-functions and Young's inequality.** We call Young function a left-continuous convex function $A : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$ such that $A(0) = 0$. An N-function is a finite valued (therefore continuous) Young function C such that

$$\lim_{t\to 0^+}\frac{C(t)}{t} = 0, \qquad \lim_{t\to\infty}\frac{C(t)}{t} = \infty.$$

A good reference for such functions, as for all the results stated in the following lines, is the book [37]. The *Young's conjugate* of an N-function is defined by

$$\widetilde{C}(t) := \sup_{s>0}\big\{st - C(s)\big\}$$

and throughout the whole paper when using the *tilde* notation over a function *we shall always mean its Young's conjugate*. If C is an N-function, then also $\widetilde{C}$ is an N-function and for these functions Young's inequality holds; moreover, in the case a condition of the type $C(\alpha t) \leq \alpha^q C(t)$ for $\alpha \in (0,1)$ and with some positive exponent q, the choice of an appropriate power of $\varepsilon \in (0,1)$, $\varepsilon^{1/q}$, leads to the following improved form:

$$ts \leq \varepsilon\, C(t) + c(\varepsilon, q)\widetilde{C}(s) \tag{3.10}$$

for all $t, s \geq 0$. Another important feature of Young's conjugate function is the following inequality, which can be found in [1, Chapter 8, (6)]:

$$\widetilde{C}\left(\frac{C(t)}{t}\right) \leq C(t) \tag{3.11}$$

for $t > 0$. The previous inequality can be inferred from the similar one:

$$t \leq C^{-1}(t)\widetilde{C}^{-1}(t) \leq 2t \qquad \text{for all } t \geq 0. \tag{3.12}$$

3.2. **Orlicz and Orlicz-Sobolev spaces.** Given a Young function A satisfying a global Δ_2-condition, the Orlicz space $L^A(\Omega)$ is the Banach space of all measurable functions $f : \Omega \to \mathbb{R}$ such that $\int_\Omega A\big(|f|\big)\, dx < \infty$, endowed with the Luxemburg norm

$$\|f\|_{L^A(\Omega)} := \inf\left\{\lambda > 0 : \int_\Omega A\left(\frac{|f|}{\lambda}\right) dx \leq 1\right\}.$$

Note that for the above norm there holds the inequality

$$\|f\|_{L^A(\Omega)} \leq \int_\Omega A(|f|)\, dx + 1, \tag{3.13}$$

see for example [37, Chapter III]. The Orlicz-Sobolev space $W^{1,A}(\Omega)$ is just made up of the functions $f \in L^A(\Omega) \cap W^{1,1}(\Omega)$ such that $Df \in L^A(\Omega)$. By $W_0^{1,A}(\Omega)$ we mean the subspace of $W^{1,A}(\Omega)$ made up of the functions whose continuation by zero outside Ω belongs to $W^{1,A}(\mathbb{R}^n)$. Note that for $\partial\Omega$ smooth enough, say Lipschitz regular, this space coincides with the closure of $C_c^\infty(\Omega)$ in $W^{1,A}(\Omega)$, at least when A satisfies a Δ_2 condition, that is our case; we shall need this observation in particular for Ω a ball. Finally we remark that in the following we shall mention the space $W^{1,g}(\Omega)$ that, in view of the previous lines, cannot even be defined, since $g(\cdot)$ is not necessarily a convex function. However we can define the Orlicz-Sobolev space $W^{1,f}(\Omega)$, $f(\cdot)$ defined in (2.4), and set $W^{1,g}(\Omega) := W^{1,f}(\Omega)$. This definition makes sense since $f \approx g$ and therefore there is no qualitative difference in between these spaces, i.e., $\int_\Omega g(|v|)\, dx < \infty$ if and only if $\int_\Omega f(|v|)\, dx < \infty$.

3.3. **Sobolev's embedding.** For the Sobolev-Orlicz spaces usual embedding theorems hold true. In particular, we can still find what can be roughly distinguished as the two different behaviors of function belonging to $W^{1,A}(\Omega)$ depending on the growth of the Young function A at infinity. If the function A grows "slowly" at infinity, then we get that functions in $W_0^{1,A}(\Omega)$ are more integrable in the Orlicz setting, as in the standard case when we have $p \leq n$. Note that the borderline case $p = n$ can be "embedded" in this case, due to the general structure of Orlicz spaces (Trudinger's Theorem [42] is nothing else than the embedding of $W^{1,n}(\Omega)$ into the Orlicz space $L^B(\Omega)$, where $B(t) = e^{t^{n'}} - 1$). In order to be more precise, let us suppose that the Young function A satisfies the following bounds:

$$\int_0 \left(\frac{s}{A(s)}\right)^{\frac{1}{n-1}} ds < \infty \qquad \text{and} \qquad \int^\infty \left(\frac{s}{A(s)}\right)^{\frac{1}{n-1}} ds = \infty. \tag{3.14}$$

In this case we have the space $W^{1,A}(\Omega)$ embeds into $L^{A_n}(\Omega)$, where we define the Young function A_n in the following line:

$$H_n(t) := \left(\int_0^t \left[\frac{s}{A(s)}\right]^{\frac{1}{n-1}} ds\right)^{\frac{n-1}{n}}, \qquad A_n(t) := (A \circ H_n^{-1})(t). \tag{3.15}$$

Note that the function $H_n(\cdot)$ depends on the starting function A, but we don't explicit this dependence for ease of notation. We will however recall this fact often in order to avoid misunderstandings. Moreover observe that the first condition in (3.14), call it $(3.14)_1$, is not really restrictive: given a Young function satisfying $(3.14)_2$, we can appropriately modify it near zero in order to satisfy $(3.14)_1$. This does not invalidate the function as belonging to the Orlicz-Sobolev space, and also in our context will lead to minor changes, see Section 5. The following (sharp) integral form of Sobolev's embedding can be found in this form in [7, Theorem 3] by Cianchi.

Proposition 3.2 (Sobolev's embedding)**.** *Let $\Omega \subset \mathbb{R}^n$, $n \geq 2$ be a bounded open set and let A be a Young function satisfying (3.14). Then there exists a constant c_S depending only on n such that for every weakly differentiable function $u \in W_0^{1,A}(\Omega)$ there holds*

$$\int_\Omega A_n\left(\frac{u}{c_S(n)\big(\int_\Omega A(|Du|)\,dx\big)^{1/n}}\right) dx \leq \int_\Omega A(|Du|)\,dx, \tag{3.16}$$

where $A_n(t) := A\big(H_n^{-1}(t)\big)$ is the function defined in (3.15).

If indeed A grows "quickly" at infinity, i.e. if

$$\int^\infty \left(\frac{s}{A(s)}\right)^{\frac{1}{n-1}} ds < \infty, \tag{3.17}$$

we have the embedding into L^∞ by Talenti [41]. The more transparent version we propose here can be found in the paper [6] by Cianchi.

Proposition 3.3 (Sobolev's embedding - II)**.** *Let Ω as in the previous proposition and let A be a Young function satisfying (3.17). Then there exists a constant depending on $n, \delta, g_0, |\Omega|$ such that for every function $u \in W_0^{1,A}(\Omega)$*

$$\sup_\Omega |u| \leq c\,\|Du\|_{L^A(\Omega)}. \tag{3.18}$$

Finally an easy Sobolev-type embedding for the function g. We state it explicitly here since in the following we shall often need to refer to it.

Proposition 3.4. *Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a positive increasing $C^1(0,\infty)$ function satisfying (1.6) and let B_R be a ball. Then there exists a constant $c \equiv c(n, \delta, g_0)$ such that*

$$\fint_{B_R} \left[g\left(\frac{|u|}{R}\right)\right]^{\frac{n}{n-1}} dx \leq c\left(\fint_{B_R} g(|Du|)\,dx\right)^{\frac{n}{n-1}}$$

for every weakly differentiable function $u \in W_0^{1,g}(B_R)$.

Proof. Define f as in (2.4) and note that it is convex and satisfies (1.6). By Sobolev's embedding in $W^{1,1}$, using (1.6) for f we know

$$\begin{aligned}\left(\fint_{B_R}\left[f\left(\frac{|u|}{R}\right)\right]^{\frac{n}{n-1}} dx\right)^{\frac{n-1}{n}} &\leq c(n)\,R \fint_{B_R} f'\left(\frac{|u|}{R}\right)\frac{|Du|}{R}\,dx \\ &\lesssim c \fint_{B_R} f\left(\frac{|u|}{R}\right)\frac{R}{|u|}|Du|\,dx.\end{aligned}$$

Now we use Young's inequality (3.10) with conjugate functions f and $\tilde{f}$ – note that f is convex and recall (3.2); (3.11) and Hölder's inequality then yield

$$\fint_{B_R} f\left(\frac{|u|}{R}\right)\frac{R}{|u|}|Du|\,dx\,dx \le c\fint_{B_R} f(|Du|)\,dx + \frac{1}{2}\fint_{B_R} \tilde f\left(f\left(\frac{|u|}{R}\right)\frac{R}{|u|}\right)dx$$
$$\le c\fint_{B_R} f(|Du|)\,dx + \frac{1}{2}\left(\fint_{B_R}\left[f\left(\frac{|u|}{R}\right)\right]^{\frac{n}{n-1}}dx\right)^{\frac{n-1}{n}}.$$

To conclude we reabsorb the latter term in the left-hand side and we recall that $f \approx g$. □

4. Homogeneous equations

In this section we collect some results for homogeneous equations of the form

$$-\operatorname{div} a(Dv) = 0 \qquad \text{on } A \subset \mathbb{R}^n \text{ bounded open set.} \tag{4.1}$$

We will assume that the vector field $a : \mathbb{R}^n \to \mathbb{R}^n$ satisfies the ellipticity and growth conditions (2.1) and in the following we will propose some variations on classical themes of Lieberman [30, 31]. The following Lemma is indeed essentially a little variation of [31, Lemma 5.1]:

Lemma 4.1. *Let $v \in W^{1,G}(A)$ be a solution to (4.1) under the assumptions (2.1)–(1.6). Then for every ball $B_R \equiv B_R(x_0) \subset A$ the following De Giorgi type estimate holds:*

$$\sup_{B_{R/4}} |Dv| \le c\fint_{B_R} |Dv|\,dx. \tag{4.2}$$

Moreover $v \in C^{1,\alpha}(A)$ for some $\alpha \in (0,1)$ and the following estimate for the excess decay holds:

$$\fint_{B_r} \left|Dv - (Dv)_{B_r}\right| dx \le c_h \left(\frac{r}{R}\right)^{\alpha} \fint_{B_R} \left|Dv - (Dv)_{B_R}\right| dx, \tag{4.3}$$

for $0 < r < R$ and B_r having the same center of B_R. Finally we have

$$|Dv(x_1) - Dv(x_2)| \le c_o \left(\frac{r}{R}\right)^{\alpha} \fint_{B_R} |Dv|\,dx \tag{4.4}$$

for every $x_1, x_2 \in B_{r/2}$. The three constants and the exponent α share the same dependence on n, ν, L, δ, g_0.

Proof. For (4.2) we merge the De Giorgi estimate present in [31, Lemmata 5.1 & 5.2] with the following Lemma 4.2, which allows to reduce the integrability of Dv on the right-hand side. We have

$$\sup_{B_{R/4}} G(|Dv|) \le c\fint_{B_{R/2}} G(|Dv|)\,dx \le c\,G\left(\fint_{B_R} |Dv|\,dx\right).$$

For (4.3) and (4.4) we take inspiration from [15, Theorem 3.1], which is in turn a revisitation of [31, Lemma 5.1] for the case of Hölder estimates below the natural growth exponent in the standard super-quadratic case. At some points we therefore shall only sketch the proof, leaving to the reader the task of completing the missing details with the help of [15, Theorem 3.1]. First note that by an appropriate argument, described in [31, Lemma 5.2], we could consider approximating vector fields a_ε satisfying (2.1) with g_ε in place of g and satisfying

$$\lim_{t\to 0} \frac{g_\varepsilon(t)}{t} = \varepsilon > 0; \tag{4.5}$$

accordingly, we could consider approximate solutions v_ε solving (4.1) with a_ε in place of a; condition (4.5) will allow to differentiate the equation for v_ε. Once proved (4.3) and (4.4) for v_ε, then we shall pass to the limit, exactly as shown in [31, Lemma 5.2], to infer the result for our original solution v. For ease of notation we shall omit the subscript ε in the proof. Take a ball $B_{\tilde r}(x_0) \subset B_R$, recall the definition of the excess in (2.6) and set

$$M(\tilde r) := \max_{k\in\{1,\dots,n\}} \sup_{B_{\tilde r}} |D_k v|.$$

It is a well known regularity fact, see [11, 30], that there exists constant $\mu_0, \eta \in (0,1)$ depending only on n, ν, L, δ, g_0 such that if one of the following two alternatives holds

$$|\{D_k v < M(\tilde r)/2\} \cap B_{\tilde r}| \le \mu_0 |B_{\tilde r}| \qquad \text{for some } k \in \{1,\dots,n\}, \tag{4.6}$$

$$|\{D_k v > -M(\tilde r)/2\} \cap B_v| \le \mu_0 |B_{\tilde r}| \qquad \text{for some } k \in \{1,\dots,n\}, \tag{4.7}$$

then

$$|Dv| \ge \frac{1}{4} M(\tilde r) \qquad \text{in } B_{\tilde r/2},$$

while if neither (4.6) nor (4.7) holds for any k, then

$$M(\tilde r/2) \le \eta M(\tilde r). \tag{4.8}$$

Now we define first the constant $H_1 \in \mathbb{N}$, then $K_1 \in \mathbb{N}$, both depending on n, ν, L, δ, g_0, such that they satisfy

$$8c_* \sqrt{n}\, \eta^{H_1} \le 1, \qquad\qquad 2^{n(H_1+2)+2} \eta^{K_1 - 1} \le 1, \tag{4.9}$$

where $\eta \in (0,1)$ is the quantity appearing in (4.8) and c_* the constant appearing in (4.2). Define moreover

$$j_0 := H_1 + K_1.$$

We distinguish now two situations, which may seem unrelated; at the end of the proof we shall show how to merge all these together.

The first alternative. Consider first the case where, for some $\bar\jmath \in \mathbb{N}$, we have that (4.6) or (4.7) holds for some k in for $\tilde r = R/2^{\bar\jmath - 1}$. Hence we have

$$|Dv| \ge \frac{1}{4} M(R/2^{\bar\jmath-1}) \qquad \text{in } B_{R/2^{\bar\jmath}}.$$

Note now that for $k = 1, \dots, n$, $\tilde v := D_k v$ is a weak solution to a *uniformly elliptic* linear equation:

$$\operatorname{div}\big(\tilde a(x) D\tilde v\big) = 0, \qquad \text{where } \tilde a(x) = \partial_\xi a(Dv(x)), \qquad \text{in } B_{\rho/2};$$

notice that the differentiation of the equation is possible, see [30, Lemma 1] since we are in the nondegenerate regime (4.5). We can now use (2.1), the monotonicity (3.1) and to get

$$\langle \tilde a(x)\lambda, \lambda\rangle = \langle \partial_\xi a(Dv)\lambda, \lambda\rangle \ge \nu \frac{g(|Dv|)}{|Dv|}|\lambda|^2 \ge \nu \frac{2^{-2g_0}}{\sqrt{n}} \frac{g(M(\tilde r))}{M(\tilde r)}|\lambda|^2,$$

$$|\tilde a(x)| \le L \frac{g(|Dv|)}{|Dv|} \le 4L\sqrt{n^{g_0}} \frac{g(M(\tilde r))}{M(\tilde r)},$$

both the inequalities being valid for $x \in B_{R/2^{\bar\jmath}}$; we shortened here $\tilde r = R/2^{\bar\jmath-1}$. Hence here $\tilde v = D_k v$ satisfies a uniformly elliptic linear equation in $B_{\rho/2}$ and hence classic theory, see e.g. [15, Lemma 3.2], gives in particular for $\tilde\varrho \le R/2^{\bar\jmath}$

$$\fint_{B_{\tilde\varrho}} \big|Dv - (Dv)_{B_{\tilde\varrho}}\big|\,dx \le c\left(\frac{\tilde\varrho}{R/2^{\bar\jmath}}\right)^{\alpha_1} \fint_{B_{R/2^{\bar\jmath}}} \big|Dv - (Dv)_{B_{R/2^{\bar\jmath}}}\big|\,dx, \tag{4.10}$$

with α_1 and c depending on n, ν, L, δ, g_0. The important point here (and also later) is that the dependence of the Hölder exponent and the constant is upon the *ellipticity ratio*, and therefore they do not depend on $M(\rho/2)$.

The second alternative. Suppose here that there exists $\rho \le R$ such that neither (4.6) nor (4.7) holds for every $r \in \{\rho, \rho/2, \dots \rho/2^{j_0}\}$ and for any k. This implies

$$M(2^{-j}\rho) \le \eta M(2^{-(j-1)}\rho) \qquad j \in \{1, \dots, j_0+1\};$$

(4.6) and (4.7) fail in for $\tilde r = 2^{-j}\rho$ for $j = 1, \dots, j_0$ and for every k. Then, iterating the previous inequality we get in particular

$$M(2^{-(H_1+2)}\rho) \le \eta^{H_1} M(\rho/4),$$
$$M(2^{-(j_0+1)}\rho) \le \eta^{K_1-1} M(2^{-(H_1+2)}\rho). \tag{4.11}$$

Note moreover that there holds $E(Dv, B_{\tilde r}) \le 2\sqrt{n} M(\tilde\varrho)$ for any $\tilde\varrho \le R$. Now we consider two different cases. In the case $|(Dv)_{B_\rho}| \le 2\sqrt{n} M(2^{-(H_1+2)}\rho)$ we have, using estimate (4.2), that in particular

$$M(\rho/4) \le c_* \fint_{B_\rho} |Dv|\,dx \le c_*\, E(Dv, B_\rho) + 2c_*\sqrt{n}\, M(2^{-(H_1+2)}\rho).$$

Combining the last two estimates and taking into account the definition of H_1 in (4.9) we can reabsorb the second term on the right-hand side obtaining

$$M(\rho/4) \le 2c_* E(Dv, B_\rho) \quad \Rightarrow \quad M(2^{-(H_1+2)}\rho) \le 2c_* \eta^{H_1} E(Dv, B_\rho)$$

and consequently, using $(4.11)_2$

$$\begin{aligned} E(Dv, B_{2^{-(j_o+1)}\rho}) &\le 2\sqrt{n} M(2^{-(j_0+1)}\rho) \le 2\sqrt{n} M(2^{-(H_1+2)}\rho) \\ &\le 4c_* \sqrt{n}\, \eta^{H_1} E(Dv, B_\rho) \le \frac{1}{2} E(Dv, B_\rho) \end{aligned}$$

by (4.9) again. In *the case* $|(Dv)_{B_\rho}| > 2\sqrt{n}\, M(2^{-(H_1+2)}\rho)$ we have

$$|Dv - (Dv)_{B_\rho}| > \sqrt{n}\, M(2^{-(H_1+2)}\rho) \qquad \text{in } B_{2^{-(H_1+2)}\rho}.$$

Taking averages of the previous relation and using $(4.11)_2$ yields

$$\begin{aligned} 2\sqrt{n} M(2^{-(H_1+K_1+1)}\rho) &\le 2\sqrt{n}\, \eta^{K_1-1} M(2^{-(H_1+2)}\rho) \\ &\le 2\eta^{K_1-1} \fint_{B_{2^{-(H_1+2)}\rho}} |Dv - (Dv)_{B_\rho}|\,dx \\ &\le 2^{n(H_1+2)+1} \eta^{K_1-1} \fint_{B_\rho} |Dv - (Dv)_{B_\rho}|\,dx \\ &\le \frac{1}{2} E(Dv, B_\rho) \end{aligned}$$

by the choice of K_1 in (4.9). Hence also in this case $E(Dv, B_{2^{-(j_o+1)}\rho}) \le \frac{1}{2} E(Dv, B_\rho)$.

Hence we proved that in the case neither (4.6) nor (4.7) holds in $B_{2^{-j}\rho}$ for some $\rho \le R$ and all $j = 1, \dots, j_0+1$, then there exists $\tau = 2^{-(j_0+1)} \in (0, 1/2]$ depending only on n, ν, L, δ, g_0 such that $E(Dv, B_{\tau\rho}) \le \frac{1}{2} E(Dv, B_\rho)$. The way this previous inequality together with (4.10) leads to (4.3) is quite standard and we refer to [15] for its proof; we just sketch the main argument.

For r as in (4.3) we choose $k \in \mathbb{N}$ such that $\tau^{k+1}R < r \le \tau^k R$ and we define

$$\mathcal{S} := \{i \in \mathbb{N} : (4.8) \text{ holds in } B_{\tau^i R}\}.$$

Now if $\mathcal{S} = \mathbb{N}$ the conclusion is easy, since we can apply the result of *the second alternative* choosing as ρ every dyadic radius $\tau^i R$ and hence we have

$$E(Dv, B_{\tau^i R}) \le \frac{1}{2^i} E(Dv, B_R) = \tau^{i\alpha} E(Dv, B_R)$$

for an appropriate $\alpha \in (0,1)$; this gives $E(Dv, B_r) \le c\,(r/R)^\alpha E(Dv, B_R)$ which is (4.3). Otherwise we set $m := \min(\mathbb{N} \setminus \mathcal{S})$ and we can apply *the first alternative* with $\bar{\jmath} = m$ and $\tilde{\varrho} = \tau^{\gamma+\ell}R$, for $\gamma, \ell \in \mathbb{N}_0$ to be chosen, such that $\gamma + \ell \ge m$, which yields

$$E(Dv, B_{\tau^{\gamma+\ell}R}) \le c\left(2^m \tau^{\gamma+\ell}\right)^{\alpha_1} E(Dv, B_{R/2^m}) \tag{4.12}$$

for all $\tilde{\varrho} \le R/2^m$; recall that $\tau = 2^{-(j_0+1)}$. Then we choose $\gamma \in \mathbb{N}_0$ such that $\gamma(j_0 + 1) < m \le (\gamma+1)(j_0+1)$ and this gives in the estimate above $E(Dv, B_{\tau^{\gamma+\ell}R}) \le c\,\tau^{\ell\alpha_1} E(Dv, B_{R/2^m})$. Now we apply the second alternative exactly γ times, with γ which could even be zero. Together with the estimate $E(Dv, B_{R/2^m}) \le 2^{n(j_0+1)+1} E(Dv, \tau^\gamma R)$, this yields

$$E(Dv, B_{R/2^m}) \le c\,\tau^{\gamma\alpha_2} E(Dv, B_R), \quad \alpha_2(n, \delta, g_0, \nu, L) \in (0,1)$$
$$\Longrightarrow \quad E(Dv, B_{\tau^{\gamma+\ell}R}) \le c\,\tau^{(\gamma+\ell)\min\{\alpha_1,\alpha_2\}} E(Dv, B_{R/2^m}),$$

and this estimate gives (4.3) exactly as after (4.12). We have been quite sloppy in this proof, but we refer to [15, Theorem 3.1], where it is performed in full detail.

(4.4) follows now by a Campanato type argument, see [17, Theorem 2.9]. □

Now we give the proof of the Reverse Hölder's inequality we used to deduce (4.2):

Lemma 4.2 (Reverse Hölder's inequality)**.** *Let $v \in W^{1,G}(A)$ be a solution to (4.1) under the conditions (2.1)–(1.6). Then for every ball $B_R(x_0) \equiv B_R \subset A$ there holds*

$$\fint_{B_{R/2}} G\big(|Dv|\big)\,dx \le c\,G\bigg(\fint_{B_R} |Dv|\,dx\bigg). \tag{4.13}$$

for a constant depending on n, ν, L, δ, g_0.

Proof. In order to lower the integrability level on the right-hand side we first consider a preliminary reverse Hölder inequality and then we exploit the self-improving character of such kind of inequalities, see [38], with an approach which wants to mimic [17, Remark 6.12]. In particular we have the following inequality, which can be found in [8, Equation (1.11)]:

$$\fint_{B_{\rho/2}(y)} G\big(|Dv|\big)\,dx \le c\left(G \circ S^{-1}\right)\bigg(\fint_{B_\rho(y)} S\big(|Dv|\big)\,dx\bigg) \tag{4.14}$$

for a constant depending upon n, ν, L, δ, g_0, valid for balls $B_\rho(y) \subset B_R$, and where the Young's function S, which grows essentially slower than G at infinity, is given by

$$S(t) := G(t)\left[\frac{G(t)}{t}\right]^{-\frac{1}{n}}.$$

Note that (4.14) is proved for minimizers of functionals like (1.7) combining a Caccioppoli's inequality with an appropriate Sobolev's type inequality involving the function S; its proof for our case of equations requires however only slight modifications. See also [38] for a Caccioppoli's inequality for minimizers satisfying hypotheses similar than ours.

Suppose now $R = 1$; we will prove the general case with the help of a scaling argument. Moreover take $r \le 1$, $\alpha \in (0,1)$ and a point $y \in B_{\alpha r}(x_0)$. Apply inequality (4.14) for $\rho = (1-\alpha)r$, i.e. over $B_{(1-\alpha)r}(y)$. Note that we have $B_{(1-\alpha)r}(y) \subset B_r$. We have

$$\fint_{B_{(1-\alpha)r/2}(y)} G\big(|Dv|\big)\,dx \le c\left(G \circ S^{-1}\right)\bigg(\fint_{B_{r(1-\alpha)}(y)} S\big(|Dv|\big)\,dx\bigg).$$

Now we come to a bit of algebra. By its definition, with $B_\rho \equiv B_{(1-\alpha)r}(y)$, we have

$$\fint_{B_\rho} S\big(|Dv|\big)\,dx = \fint_{B_\rho} \big[G(|Dv|)\big]^{\frac{n-1}{n}} |Dv|^{\frac{1}{n}}\,dx \tag{4.15}$$

$$\leq \left(\fint_{B_\rho} G(|Dv|)\,dx \right)^{\frac{n-1}{n}} \left(\fint_{B_\rho} |Dv|\,dx \right)^{\frac{1}{n}},$$

using Hölder's inequality. Now we want to use Young's inequality with conjugate functions $C(t) := S(t^n)$ and $\widetilde{C}(t)$. Using an argument we will use also elsewhere in the paper, changing variable ($s = \sigma^{\frac{1}{n}}$) in the definition of the Young's conjugate function, for $\alpha > 0$,

$$\begin{aligned} \widetilde{C}(\alpha^{\frac{n-1}{n}}) &= \sup_{s>0} \alpha^{\frac{n-1}{n}} s - S(s^n) \qquad (4.16)\\ &\lesssim_n \left[\sup_{\sigma>0} \alpha^{n-1}\sigma - \big[S(\sigma)\big]^n \right]^{\frac{1}{n}} =: \left[\widetilde{T}\big(\alpha^{n-1}\big) \right]^{\frac{1}{n}} \end{aligned}$$

where the function T is defined by $T(t) := [S(t)]^n$. Notice that we can clearly restrict our attention to the set where $\alpha^{\frac{n-1}{n}} s \geq S(s^n)$. At this point the reader might recall that

$$\widetilde{T}\big([G(\tau)]^{n-1}\big) = \widetilde{T}\left(\frac{[S(\tau)]^n}{\tau} \right) = \widetilde{T}\left(\frac{T(\tau)}{\tau} \right) \leq T(\tau) = \big[S(\tau)\big]^n$$

from the definition of S; the choice $\tau = G^{-1}(\alpha)$ leads now to

$$\widetilde{T}\big(\alpha^{n-1}\big) \leq \big[S\big(G^{-1}(\alpha)\big)\big]^n,$$

and plugging the latter estimate into (4.16) and choosing $\alpha = \fint_{B_r} G(|Dv|)\,dx$ gives the bound

$$\widetilde{C}\left(\left(\fint_{B_\rho} G(|Dv|)\,dx \right)^{\frac{n-1}{n}} \right) \lesssim_n \big(S \circ G^{-1}\big)\left(\fint_{B_\rho} G(|Dv|)\,dx \right). \qquad (4.17)$$

Note now that $t \mapsto (G \circ S^{-1})(t)$ is increasing and therefore the subadditivity property (3.8) holds. Using this fact and Young's inequality with appropriate $\varepsilon \in (0,1)$, together with (4.17) into (4.15), and recalling that $\rho = (1-\alpha)r$ gives

$$\fint_{B_{(1-\alpha)r/2}} G(|Dv|)\,dx \leq \frac{1}{2} \fint_{B_{(1-\alpha)r}} G(|Dv|)\,dx + c\,G\left(\fint_{B_{(1-\alpha)r}} |Dv|\,dx \right)$$

with $c \equiv c(n, \delta, g_0)$; in turn

$$\begin{aligned} \int_{B_{(1-\alpha)r/2}} G(|Dv|)\,dx &\leq \frac{1}{2} \int_{B_{(1-\alpha)r}} G(|Dv|)\,dx \\ &\quad + c\big[(1-\alpha)r\big]^{-ng_0} G\left(\int_{B_{(1-\alpha)r}} |Dv|\,dx \right). \end{aligned}$$

Note now that the ball $B_{\alpha r}$ can be covered by balls of this kind in such a way that only a finite and independent of α number of balls of double radius intersect, all included in B_r. We then have calling $\alpha r =: s < r$

$$\int_{B_s} G(|Dv|)\,dx \leq \frac{1}{2} \int_{B_r} G(|Dv|)\,dx + \frac{c}{(r-s)^{-ng_0}} G\left(\int_{B_r} |Dv|\,dx \right);$$

at this point a standard iteration Lemma (see [17, Lemma 6.1]) gives (4.13) for the case $R = 1$. For the general case rescale in the following way: define $\tilde{v}(x) := v(x_0 + Rx)/R$. $\tilde{v}$ solves $-\operatorname{div} a(D\tilde{v}) = 0$ on $B_1(0)$ and therefore we can apply (4.13) to $\tilde{v}$. Rescaling back gives the reverse Hölder's inequality in the general case. □

Finally, a so-called "density improvement Lemma":

Lemma 4.3. *Suppose that the two conditions*

$$\frac{\lambda}{C} \leq \fint_{\sigma^m B} |Dv|\,dx \qquad \textit{and} \qquad \sup_{B/4} |Dv| \leq C\lambda, \qquad (4.18)$$

hold for some $m \in \mathbb{N}$*, some numbers* $C \geq 1$ *and* $\lambda \geq 0$ *and with*

$$0 < \sigma^\alpha \leq \frac{1}{2^{3\alpha+2} c_o C^2} < \frac{1}{8^\alpha}, \tag{4.19}$$

where $\alpha \in (0,1)$ *and* c_o *appear in Lemma 4.1. Then*

$$\frac{\lambda}{4C} \leq |Dv| \qquad \text{in } \sigma B.$$

Proof. From $(4.18)_1$ we deduce that there exists a point $x_0 \in \sigma^m B$ such that $|Dv(x_0)| > \lambda/2C$. On the other hand, (4.4) and $(4.18)_2$ give $|Dv(x) - Dv(x_0)| \leq c_o(2\sigma)^\alpha C\lambda$ whenever $x \in B_{r/2} \equiv \sigma B \subset B/8$. The choice above for σ together with the last two inequalities gives

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

for all $x \in \sigma B$. □

5. Comparison estimates

In this Section we want to derive comparison estimates between the solution to equation (1.8) and to a suitable homogeneous Cauchy problem. In particular, given a ball $B_R \equiv B_R(x_0) \subset \Omega$, we consider the solution $v \in u + W_0^{1,G}(B_R)$ to the Cauchy problem

$$\begin{cases} -\operatorname{div} a(Dv) = 0 & \text{in } B_R, \\ v = u & \text{on } \partial B_R. \end{cases} \tag{5.1}$$

Existence and uniqueness of such functions are given with approximation and monotonicity arguments, see [31, Lemma 5.2].

The technique permitting to obtain L^q estimates, being q "close" to one, for the difference $Du - Dv$ via scaling arguments has appeared several time before this paper, see [14, 15, 34]. However the proof for the general setting we are considering in this paper appears to be more technical and involved. We point out that the principal point we want to avoid here is to get estimates degenerating when either the difference $g_0 - \delta$ or the ratio g_0/δ becomes too large.

In his first part of the Section we introduce and study separately some auxiliary functions we shall use in the proofs. First we introduce two functions directly depending on g:

$$f_\chi(t) := \int_0^t \left[\frac{g(s)}{s}\right]^{1+\chi} ds, \qquad g_\chi(t) := \left[\frac{g(t)}{t}\right]^{1+\chi} t, \tag{5.2}$$

for $\chi \geq -1$. Note that functions similar to g_χ have been already used for example in [16]. We immediately stress that, by a simple computation of derivatives, the use of (1.6) and integration over $(0,t)$ we have

$$\left[\delta(1+\chi) - \chi\right] f_\chi(t) \leq g_\chi(t) \leq \left[g_0(1+\chi) - \chi\right] f_\chi(t) \tag{5.3}$$

and therefore $f_\chi(t) \approx_\chi g_\chi(t)$. Note also that

$$f_\chi(\alpha t) \lesssim \max\left\{\alpha^{(g_o-1)(1+\chi)+1}, \alpha^{(\delta-1)(1+\chi)+1}\right\} f_\chi(t)$$

for $\alpha \geq 0$. Moreover we need to introduce the function $H_\chi(t)$ defined *through* the following formula

$$H_\chi^{-1}(t) := t^{-\chi}\left[G^{-1}\left(\frac{1}{t}\right)\right]^{-(2\chi+1)};$$

here (and in the sequel) we eventually use the conventions that $1/0 = \infty$, $1/\infty = 0$ and $G^{-1}(\infty) = \infty$, so that H_χ^{-1} (and other functions) are defined in zero in a direct way. Note that a computation shows that

$$\begin{aligned}\frac{d}{dt}H_\chi^{-1}(t) &= \big[G(\tau)\big]^{\chi+1}\tau^{-(2\chi+2)}\bigg[(2\chi+1)\frac{G(\tau)}{g(\tau)} - \chi\tau\bigg]\\ &\geq \bigg[\frac{2\chi+1}{1+g_0} - \chi\bigg]\big[G(\tau)\big]^{\chi+1}\tau^{-(2\chi+1)} \geq 0\end{aligned}$$

with $\tau := G^{-1}(\frac{1}{t})$, if $\chi \leq \frac{1}{g_o-1}$. $t \mapsto H_\chi^{-1}(t)$ is hence increasing and it makes sense to define its inverse, namely $H_\chi(t)$. Note moreover that by (5.3) we have

$$H_\chi^{-1}(t) \approx_\chi t\, g_\chi\Big(G^{-1}\Big(\frac{1}{t}\Big)\Big) \approx_\chi t\, f_\chi\Big(G^{-1}\Big(\frac{1}{t}\Big)\Big) \tag{5.4}$$

and by (3.12) we deduce

$$[\widetilde{H_\chi}]^{-1}\bigg(\frac{1}{G(t)}\bigg) \approx \frac{1}{G(t)}\frac{1}{H_\chi^{-1}(\frac{1}{G(t)})} \approx \frac{1}{G(t)}\frac{t^{2\chi+1}}{\big[G(t)\big]^\chi} \approx \frac{1}{f_\chi(t)}.$$

Matching this estimate together with the one inferred from the left-hand side inequality of (3.12) and (5.4) we deduce

$$H_\chi^{-1}\bigg(\widetilde{H_\chi}\bigg(\frac{1}{f_\chi(t)}\bigg)\bigg) \approx H_\chi^{-1}\bigg(\frac{1}{G(t)}\bigg) = \frac{\big[G(t)\big]^\chi}{t^{2\chi+1}} \approx \frac{\big[g(t)\big]^\chi}{t^{\chi+1}} = \frac{g_\chi(t)}{g(t)t}. \tag{5.5}$$

Finally, being the function $H_\chi^{-1}(\cdot)$ increasing and having the doubling property $H_\chi^{-1}(2t) \lesssim_\chi H_\chi^{-1}(t)$ at hand, by Remark 3.1 we have that the following inequality

$$H_\chi^{-1}(t+s) \lesssim_\chi H_\chi^{-1}(t) + H_\chi^{-1}(s) \tag{5.6}$$

for $t, s \geq 0$. Finally we come to the proof of the comparison estimate:

Lemma 5.1. *Let $u \in W^{1,G}(\Omega)$ be the solutions to the equation (1.5) and $v \in u + W_0^{1,G}(B_R)$ the solution to the problem (5.1) on B_R. Then the following estimate holds true:*

$$\fint_{B_R} g_\chi\big(|Du - Dv|\big)\, dx \leq c_1 g_\chi(A) \quad \textit{where} \quad A := g^{-1}\bigg(\frac{|\mu|(B_R)}{R^{n-1}}\bigg), \tag{5.7}$$

where g_χ is the functions defined in (5.2), for

$$\chi \in \Big[-1, \min\Big\{\frac{1}{g_0-1}, \frac{g_0}{(g_0-1)(n-1)}\Big\}\Big) \tag{5.8}$$

and with a constant c_1 depending on $n, \nu, \delta, g_0, \chi$.

Proof. *Step 1: rescaling.*

Define A as in (5.7); we can suppose without loss of generality that $A > 0$, since in the case $|\mu|(B_R) = 0$ the monotonicity of the vector field ensures $u = v$ on B_R and then (5.7) is trivially true. Consequently we define

$$\begin{aligned} \bar u(x) &:= \frac{u(x_0+Rx)}{AR}, & \bar v(x) &:= \frac{v(x_0+Rx)}{AR},\\ \bar a(z) &:= \frac{a(Az)}{g(A)}, & \bar\mu(x) &:= R\,\frac{\mu(x_0+Rx)}{g(A)}, \end{aligned} \tag{5.9}$$

then, subtracting the weak formulations of (1.8) and (5.1) and rescaling we have

$$-\operatorname{div}\big[\bar a(D\bar u) - \bar a(D\bar v)\big] = \bar\mu \qquad \text{in } B_1; \tag{5.10}$$

note that the growth function $\bar g$ of the vector field $\bar a$ is given by

$$\bar g(t) := \frac{g(At)}{g(A)}:$$

indeed

$$\langle \partial\tilde{a}(z)\lambda,\lambda\rangle = \frac{A}{g(A)}\langle \partial a(Az)\lambda,\lambda\rangle \geq \nu\,\frac{A}{g(A)}\frac{g(A|z|)}{A|z|}|\lambda|^2 = \nu\,\frac{\bar{g}(|z|)}{|z|}|\lambda|^2$$

for all $z,\lambda\in\mathbb{R}^n$. Since we are treating measure data problems, this is enough since we will use only the ellipticity of the vector field. However a similar estimate holds true for the growth of the vector field. Moreover note that

$$\frac{t\bar{g}'(t)}{\bar{g}(t)} = \frac{At\,g'(At)}{g(At)} \in [\delta, g_0]$$

for all $t>0$. The aim of this substitution is twofold: we can restrict ourselves to prove the Lemma in the case $B_R(x_0)=B_1$; moreover we can exploit the following estimate

$$|\tilde{\mu}|(B_1) = \frac{1}{g(A)}\frac{|\mu|(B_R)}{R^{n-1}} = 1. \tag{5.11}$$

In this case what we want to prove is simply

$$\int_{B_1}\bar{g}_\chi\big(|D\bar{u}-D\bar{v}|\big)\,dx \leq c(n,\nu,\delta,g_0), \tag{5.12}$$

where $\bar{g}_\chi$ is obtained starting from $\bar{g}$ instead of g in the expression appearing in (5.2). At the very end of the proof we will show how to recover the full result from (5.12).

Step 2: measure data estimates.

From now on we will drop the tilde notation, recovering it only in Step 3, equation (5.31). We recall we are working under the assumptions $B_R=B_1$ and $|\mu|(B_1)=1$. Since we want estimates involving only the mass of the measure μ, we shall at least initially follow the standard truncation method for which the unavoidable references are the works of Boccardo and Gallouët [2, 3]. Some changes are however needed in order to handle the growth condition we are considering. Moreover, two different approach are needed to treat the two different kind of growth G could have at infinity. In the standard case, this correspond to consider the two cases $p\leq n$ and $p>n$. In both cases we will need to consider the weak formulation of (5.10)

$$\int_{B_1}\langle a(Du)-a(Dv),D\varphi\rangle\,dx = \int_{B_1}\varphi\,d\mu \tag{5.13}$$

holding true for bounded functions $\varphi\in W_0^{1,G}\cap L^\infty(B_1)$.

Step 2.1: The slow growth case. With this expression we want to suggest the case where

$$\int^\infty\left(\frac{s}{G(s)}\right)^{\frac{1}{n-1}}ds = \infty. \tag{5.14}$$

In order to use Sobolev's embedding, we need to introduce a slightly modified function in order to have the integrability property $(3.14)_1$. We therefore define the continuous function

$$\mathfrak{f}_\chi(t) := \begin{cases} 0 & t=0,\\ f_\chi(1)t & \text{for } t\in(0,1),\\ f_\chi(t) & \text{for } t\in[1,\infty). \end{cases} \tag{5.15}$$

Let's begin putting into (5.13) the test function

$$\varphi := T_k\left(\frac{u-v}{c_S(n)\big(\int_{B_1}\mathfrak{f}_\chi(|Du-Dv|)\,dx\big)^{\frac{1}{n}}}\right) =: T_k\left(\frac{u-v}{c_S(n)\,\mathcal{F}}\right),$$

for any $k \in \mathbb{N}_0$, being $c_S(n)$ the constant appearing in (3.16) and $\mathfrak{f}_\chi(\cdot)$ the function defined in (5.15). The classical truncation operators are defined as

$$T_k(\sigma) := \max\{-k, \min\{k, \sigma\}\}, \qquad \Phi_k(\sigma) := T_1\big(\sigma - T_k(\sigma)\big) \tag{5.16}$$

for $k \in \mathbb{N}_0$ and $\sigma \in \mathbb{R}$. Note that we can clearly suppose $\mathcal{F} \geq 1$ and that $\varphi \in W_0^{1,G}(B_1) \cap L^\infty(B_1)$, since $\sigma \to T_k(\sigma)$ is Lipschitz; then φ is allowed as test function. We moreover have $D\varphi = \frac{D(u-v)}{c_S(n)\mathcal{F}}\chi_{C_k}$, being χ_{C_k} the characteristic function of the set C_k, where

$$C_k := \Big\{x \in B_1 : \frac{|u(x) - v(x)|}{c_S(n)\big(\int_{B_1} \mathfrak{f}_\chi(|Du - Dv|)\,dx\big)^{\frac{1}{n}}} \leq k\Big\}.$$

Using (3.9) we have

$$\begin{aligned}\int_{B_1} \langle a(Du) - a(Dv), D\varphi\rangle\,dx &= \frac{1}{c_S(n)\,\mathcal{F}}\int_{C_k} \langle a(Du) - a(Dv), Du - Dv\rangle\,dx \\ &\geq \frac{c}{c_S(n)\,\mathcal{F}}\int_{C_k} G\big(|Du - Dv|\big)\,dx,\end{aligned}$$

Estimating the right-hand side in the trivial way

$$\left|\int_{B_1} T_k\Big(\frac{u-v}{c_S(n)\,\mathcal{F}}\Big)\,d\mu\right| \leq \int_{B_1} k\,d|\mu| = k|\mu|(B_1) = k$$

by (5.11), we deduce the estimate

$$\int_{C_k} G\big(|Du - Dv|\big)\,dx \leq c\,k\Big(\int_{B_1} \mathfrak{f}_\chi\big(|Du - Dv|\big)\,dx\Big)^{\frac{1}{n}} = c\,k\mathcal{F}, \tag{5.17}$$

for all $k \in \mathbb{N}_0$, where $c \equiv c(n, \nu, \delta, g_0)$. Reasoning in an analogous way, using as test function $\Phi_k\big((u-v)/(c_S(n)\mathcal{F})\big) \in W_0^{1,G}(B_1) \cap L^\infty(B_1)$, we infer

$$\int_{D_k} G\big(|Du - Dv|\big)\,dx \leq c(n,\nu)\Big(\int_{B_1} \mathfrak{f}_\chi\big(|Du - Dv|\big)\,dx\Big)^{\frac{1}{n}} = c\,\mathcal{F}$$

since $\Phi_k \leq 1$, where we have denoted

$$D_k := \Big\{x \in B_1 : k < \frac{|u(x) - v(x)|}{c_S(n)\big(\int_{B_1} \mathfrak{f}_\chi(|Du - Dv|)\,dx\big)^{\frac{1}{n}}} \leq k+1\Big\}.$$

Now we come back to f_χ defined in (5.2) and we note that $t \mapsto f_\chi\big(G^{-1}(t)\big)$ is increasing and concave: indeed a computation of derivatives, denoting $\tau := G^{-1}(t)$, gives

$$\begin{aligned}\frac{d}{dt} f_\chi\big(G^{-1}(t)\big) &= \frac{f'_\chi(\tau)}{g(\tau)} = \frac{[g(\tau)]^\chi}{\tau^{1+\chi}}, \\ \frac{d^2}{dt^2} f_\chi\big(G^{-1}(t)\big) &= \frac{\chi\tau^{1+\chi}[g(\tau)]^{\chi-1}g'(\tau) - [g(\tau)]^\chi(1+\chi)\tau^\chi}{g(\tau)\tau^{2(1+\chi)}} \\ &= \frac{\big[g(\tau)\big]^{\chi-2}}{\tau^{\chi+2}}\big[\chi\tau g'(\tau) - (1+\chi)g(\tau)\big] \\ &\leq \frac{\big[g(\tau)\big]^{\chi-1}}{\tau^{\chi+2}}\big[\chi g_0 - (1+\chi)\big] < 0\end{aligned} \tag{5.18}$$

by (1.6) and the fact that $\chi < \frac{1}{g_0-1}$. Therefore using Jensen's inequality and (5.17) we get

$$\begin{aligned}\fint_{C_k} f_\chi\big(|Du - Dv|\big)\,dx &\leq \big(f_\chi \circ G^{-1}\big)\Big(\fint_{C_k} G\big(|Du - Dv|\big)\,dx\Big) \\ &\lesssim c\,\big(f_\chi \circ G^{-1}\big)\Big(\frac{k\mathcal{F}}{|C_k|}\Big).\end{aligned} \tag{5.19}$$

So using (5.4) and doing an easy algebraic manipulation we infer

$$\int_{C_k} f_\chi\big(|Du-Dv|\big)\,dx \lesssim c\,k\mathcal{F}H_\chi^{-1}\Big(\frac{|C_k|}{k\mathcal{F}}\Big). \tag{5.20}$$

with $c \equiv c(n,\nu,\delta,g_0,\chi)$. By a similar argument we have for the integrals over D_k

$$\int_{D_k} f_\chi\big(|Du-Dv|\big)\,dx \lesssim c\,\mathcal{F}H_\chi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big). \tag{5.21}$$

We hence have, using (5.20) and (5.21)

$$\begin{aligned}\int_{B_1} f_\chi\big(|Du-Dv|\big)\,dx &= \int_{C_1} f_\chi\big(|Du-Dv|\big)\,dx + \sum_{k=1}^\infty \int_{D_k} f_\chi\big(|Du-Dv|\big)\,dx \\ &\le \tilde c\,\mathcal{F}\Big[H_\chi^{-1}\Big(\frac{|B_1|}{\mathcal{F}}\Big) + \sum_{k=1}^\infty H_\chi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big)\Big] \end{aligned} \tag{5.22}$$

with $\tilde c \equiv \tilde c(n,\nu,\delta,g_0,\chi)$. Here to estimate the summation appearing on the right-hand side, we have to work on the modified function $\mathfrak{f}_\chi$ defined in (5.15). Note that $\mathfrak{f}_\chi(\cdot)$ is a Young function and

$$\int_0 \Big(\frac{s}{\mathfrak{f}_\chi(s)}\Big)^{\frac{1}{n-1}} ds < \infty \quad \text{and} \quad \int^\infty \Big(\frac{s}{\mathfrak{f}_\chi(s)}\Big)^{\frac{1}{n-1}} ds = +\infty; \tag{5.23}$$

the first by construction and the second by (5.14), since for $s \ge 1$

$$\mathfrak{f}_\chi(s) = f_\chi(s) \approx \Big[\frac{g(s)}{s}\Big]^{1+\chi} s \lesssim G(s)\frac{[g(s)]^\chi}{s^{1+\chi}} \lesssim_\chi G(s)\, s^{\chi g_0-(1+\chi)} \le G(s)$$

being $\chi \le \frac{1}{g_0-1}$. We can therefore define the Sobolev's conjugate function $(\mathfrak{f}_\chi)_n := \mathfrak{f}_\chi \circ H_n^{-1}$, where in this case H_n is given by $(3.15)_1$ with the choice $A \equiv \mathfrak{f}_\chi$. Moreover, since

$$f_\chi(1) = \int_0^1 \Big[\frac{g(s)}{s}\Big]^{1+\chi} ds \le \int_0^1 s^{(\delta-1)(1+\chi)}\,ds = \frac{1}{(\delta-1)(1+\chi)+1} \le 1,$$

then we have, for $t \ge 1$

$$H_n^{-1}(t) \ge \Big[\int_0^1 \Big(\frac{s}{\mathfrak{f}_\chi(s)}\Big)^{\frac{1}{n-1}} ds\Big]^{\frac{n-1}{n}} = \Big(\frac{1}{f_\chi(1)}\Big)^{\frac{1}{n}} \ge 1. \tag{5.24}$$

At this point we have

$$|D_k| \le \frac{1}{(\mathfrak{f}_\chi)_n(k)} \int_{D_k} (\mathfrak{f}_\chi)_n\Big(\frac{|u-v|}{c_S(n)\mathcal{F}}\Big)\,dx \tag{5.25}$$

for every $k \in \mathbb{N}$ by the definition of the set D_k. Having now both assumptions (3.14) at hand, we can deduce, using Sobolev's embedding (3.16), the following estimate for the summation: taking into account Young's inequality with conjugate functions H_χ, $\widetilde{H_\chi}$ with $\varepsilon \in (0,1)$ to be chosen, (3.16) and subadditivity (5.6)

$$\begin{aligned}\sum_{k=1}^\infty H_\chi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big) &\le \sum_{k=1}^\infty H_\chi^{-1}\Big(\frac{1}{\mathcal{F}(f_\chi)_n(k)}\int_{D_k}(\mathfrak{f}_\chi)_n\Big(\frac{|u-v|}{c_S(n)\mathcal{F}}\Big)\,dx\Big) \\ &\le \frac{\varepsilon}{\mathcal{F}}\int_{B_1}(\mathfrak{f}_\chi)_n\Big(\frac{|u-v|}{c_S(n)\mathcal{F}}\Big)\,dx + c(\delta,g_0,\chi,\varepsilon)\sum_{k=1}^\infty H_\chi^{-1}\Big(\widetilde{H_\chi}\Big(\frac{1}{(f_\chi)_n(k)}\Big)\Big) \\ &\le \frac{\varepsilon}{\mathcal{F}}\int_{B_1}\mathfrak{f}_\chi\big(|Du-Dv|\big)\,dx + c_\varepsilon\sum_{k=1}^\infty H_\chi^{-1}\Big(\widetilde{H_\chi}\Big(\frac{1}{(f_\chi)_n(k)}\Big)\Big) \\ &\le \frac{\varepsilon}{\mathcal{F}}\int_{B_1} f_\chi\big(|Du-Dv|\big)\,dx + c(n,\delta,g_0,\chi) + c_\varepsilon\sum_{k=1}^\infty \frac{\big[g\big(H_n^{-1}(k)\big)\big]^\chi}{\big[H_n^{-1}(k)\big]^{1+\chi}} \end{aligned} \tag{5.26}$$

by (5.5), having

$$(\mathfrak{f}_\chi)_n(k) = \mathfrak{f}_\chi\big(H_n^{-1}(k)\big) = f_\chi\big(H_n^{-1}(k)\big) \qquad \text{since } H_n^{-1}(k) \geq 1 \quad \text{by (5.24);}$$

here $c_\varepsilon \equiv c_\varepsilon(\delta, g_0, \chi, \varepsilon)$. Moreover in the last line we also replaced $\mathfrak{f}_\chi$ with f_χ in the first term, since $\mathfrak{f}_\chi(t) \leq f_\chi(1) + f_\chi(t) \lesssim_\chi 1 + f_\chi(t)$ and $\varepsilon/\mathcal{F} \leq 1$. Now we inquire the convergence of the series on the right-hand side. A quite long but elementary calculation of its derivative, similar to (5.18), shows that $\sigma \mapsto [g(H_n^{-1}(\sigma))]^\chi/[H_n^{-1}(\sigma)]^{1+\chi}$ is decreasing, since $\chi < \frac{1}{g_0-1}$, and hence it is easily seen that the series is dominated by the quantity

$$\frac{[g(H_n^{-1}(1))]^\chi}{[H_n^{-1}(1)]^{1+\chi}} + \int_1^\infty \left[\frac{g\big(H_n^{-1}(\sigma)\big)}{H_n^{-1}(\sigma)}\right]^\chi \frac{d\sigma}{H_n^{-1}(\sigma)}.$$

Therefore now we want to show that

$$\int_1^\infty \left[\frac{g\big(H_n^{-1}(\sigma)\big)}{H_n^{-1}(\sigma)}\right]^\chi \frac{d\sigma}{H_n^{-1}(\sigma)} = c(n, \delta, g_0, \chi) < \infty. \tag{5.27}$$

We use the change of variable $s = H_n^{-1}(\sigma)$; this is allowed by (5.23) and the fact that $\sigma \mapsto H_n^{-1}(\sigma)$ is increasing. We note that

$$d\sigma = H_n'(s)\, ds = c(n)\big[H_n(s)\big]^{\frac{1}{1-n}} \left[\frac{s}{\mathfrak{f}_\chi(s)}\right]^{\frac{1}{n-1}} ds \tag{5.28}$$

and by monotonicity

$$H_n^{-1}(s) = \left(\int_0^s \left[\frac{\tau}{\mathfrak{f}_\chi(\tau)}\right]^{\frac{1}{n-1}} d\tau\right)^{\frac{n-1}{n}} \geq s^{\frac{n-1}{n}} \left[\frac{s}{\mathfrak{f}_\chi(s)}\right]^{\frac{1}{n}}.$$

Hence for $s \geq H_n^{-1}(1) \geq 1$ we have

$$d\sigma \leq c(n)\left[\frac{s}{\mathfrak{f}_\chi(s)}\right]^{\frac{1}{n}} s^{-\frac{1}{n}} \leq c(n)\left[\frac{s}{f_\chi(s)}\right]^{\frac{1}{n}} s^{-\frac{1}{n}} \lesssim_n \left[\frac{s}{g_\chi(s)}\right]^{\frac{1}{n}} s^{-\frac{1}{n}},$$

since we have $f_\chi(s) \leq \mathfrak{f}_\chi(s)$. We therefore have by (5.3) and $H_n^{-1}(1) \geq 1$

$$\int_1^\infty \left[\frac{g\big(H_n^{-1}(\sigma)\big)}{H_n^{-1}(\sigma)}\right]^\chi \frac{d\sigma}{H_n^{-1}(\sigma)} \lesssim_{n,\chi} \int_1^\infty \big[g_\chi(s)\big]^{1-\frac{1}{n}} \frac{ds}{sg(s)} < \infty.$$

The latter integral is finite since in the case $\chi \geq 1/(n-1)$ (i.e. in the case we can use the estimate from above (3.4) in the second inequality of the next line)

$$\frac{\big[g_\chi(s)\big]^{1-\frac{1}{n}}}{sg(s)} = \big[g(s)\big]^{(1+\chi)(1-\frac{1}{n})-1} s^{-1-\chi(1-\frac{1}{n})} \lesssim_{n,\chi} s^{\mathfrak{e}(g_0)},$$

where $\mathfrak{e}(\alpha) := \alpha(1+\chi)\big(1-\dfrac{1}{n}\big) - \alpha - 1 - \chi\big(1-\dfrac{1}{n}\big) < -1$ and $\mathfrak{e}(g_0) < -1$ by the fact that $\chi < \frac{g_0}{(g_0-1)(n-1)}$. In the case $\chi \in [-1, 1/(n-1))$ we instead have

$$\frac{\big[g_\chi(s)\big]^{1-\frac{1}{n}}}{sg(s)} \lesssim_\chi \big[g(t)\big]^{(1+\chi)(1-\frac{1}{n})-1} s^{-1-\chi(1-\frac{1}{n})} \lesssim_{n,\chi} s^{\mathfrak{e}(\delta)}$$

and $\mathfrak{e}(\delta) < -1$ since $\chi < \frac{g_0}{(g_0-1)(n-1)} < \frac{\delta}{(\delta-1)(n-1)}$. Therefore in both cases (5.27) holds. Coming then back to (5.26) and (5.22)

$$\int_{B_1} f_\chi\big(|Du - Dv|\big)\, dx \leq \tilde{c}\,\mathcal{F} H_\chi^{-1}\left(\frac{|B_1|}{\mathcal{F}}\right) + \varepsilon\tilde{c}\int_{B_1} f_\chi\big(|Du - Dv|\big)\, dx + \tilde{c}_\varepsilon \mathcal{F}.$$

First we choose ε, depending on $n, \nu, \delta, g_0, \chi$, so small that we can reabsorb the second term of the right-hand side into the left-hand side, i.e. $\varepsilon = 1/(4\tilde{c})$. This fixes the value of $\tilde{c}_\varepsilon$ as a constant depending on $n, \nu, \delta, g_0, \chi$. Then we recall the definition of $\mathcal{F}$ and we estimate

$$\mathcal{F} = \left(\int_{B_1} \mathfrak{f}_\chi\big(|Du - Dv|\big)\, dx \right)^{\frac{1}{n}} \leq \tilde{\varepsilon} \int_{B_1} f_\chi\big(|Du - Dv|\big)\, dx + c(n, \delta, g_0, \tilde{\varepsilon})$$

with $\tilde{\varepsilon}$ small in order to reabsorb also this term, i.e. $\tilde{\varepsilon} := 1/(4\tilde{c}_\varepsilon)$. To conclude note that (5.4) gives

$$\mathcal{F} H_\chi^{-1}\left(\frac{|B_1|}{\mathcal{F}}\right) \lesssim_\chi |B_1|\, g_\chi\left(G^{-1}\left(\frac{\mathcal{F}}{|B_1|}\right)\right) \lesssim_{n,\chi} \mathcal{F}^{1+\chi}\Big[G^{-1}(\mathcal{F})\Big]^{-(2\chi+1)}.$$

from the definition of g_χ and the fact that $g(t) \approx G(t)/t$. Recall again we are supposing $\mathcal{F} \geq 1$, and therefore using (3.7) we infer

$$\mathcal{F}^{1+\chi}\Big[G^{-1}(\mathcal{F})\Big]^{-(2\chi+1)} \lesssim_\chi \mathcal{F}^{1+\chi-\frac{1+2\chi}{1+g_0}}.$$

Since the exponent of $\mathcal{F}$ reveals to be strictly smaller than one by $\chi < \frac{1}{g_0-1}$, we use for the third time Young's inequality together with $f_\chi \approx g_\chi$ to finally get (5.12).

Step 2.2: The fast growth case. We here approach the simpler case where

$$\int^\infty \left(\frac{s}{G(s)}\right)^{\frac{1}{n-1}} ds < \infty.$$

In this case, since both u and v belong to $W^{1,G}(B_1)$, their difference is bounded and we can directly choose $\varphi = u - v \in W_0^{1,G}(B_1) \cap L^\infty$ as a test function in (5.13). Therefore using (2.5) we infer

$$\begin{aligned} \int_{B_1} G\big(|Du - Dv|\big)\, dx &\leq \int_{B_1} (u - v)\, d\mu \\ &\leq \sup_{B_1} |u - v| |\mu|(B_1) \leq c\, \|Du - Dv\|_{L^G(B_1)} \end{aligned} \tag{5.29}$$

by (3.18) and the fact that $|\mu|(B_1) = 1$, with $c \equiv c(n, \delta, g_0)$. Let's apply inequality (3.13) to the function εf with $\varepsilon \in (0,1)$. We have

$$\varepsilon \|f\|_{L^G(B_1)} = \|\varepsilon f\|_{L^G(B_1)} \leq \int_{B_1} G(\varepsilon |f|)\, dx + 1 \leq \varepsilon^{1+\delta} \int_{B_1} G(|f|)\, dx + 1$$

by (3.6) and therefore we can use the following version of Young's inequality:

$$\|f\|_{L^G(B_1)} \leq \varepsilon^\delta \int_{B_1} G(|f|)\, dx + \varepsilon^{-1}. \tag{5.30}$$

Let's make use of it with $f = Du - Dv$ into (5.29): choosing ε small enough and reabsorbing the right-hand side term gives

$$\fint_{B_1} G\big(|Du - Dv|\big)\, dx \leq c(n, \delta, g_0).$$

Arguing as in (5.19) we infer

$$\begin{aligned} \int_{B_1} g_\chi\big(|Du - Dv|\big)\, dx &\lesssim c(n) \fint_{B_1} f_\chi\big(|Du - Dv|\big)\, dx \\ &\leq c\,\big(f_\chi \circ G^{-1}\big)\left(\fint_{B_1} G\big(|Du - Dv|\big)\, dx\right) \leq c \end{aligned}$$

with $c \equiv c(n, \delta, g_0, \chi)$, from the fact that $t \mapsto \big(f_\chi \circ G^{-1}\big)(t)$ is concave.

Step 3: Recovering the situation.

Now we recover the tilde notation: recalling the definitions given in (5.9), denoting in short $y = x_0 + Rx$, (5.12) can be rephrased as

$$\left[\frac{A}{g(A)}\right]^{1+\chi} \frac{1}{A} \fint_{B_1} \left[\frac{g\big(|Du(y) - Dv(y)|\big)}{|Du(y) - Dv(y)|}\right]^{1+\chi} |Du(y) - Dv(y)|\, dx$$
$$= \fint_{B_1} \left[\frac{\bar g\big(|D\bar u - D\bar v|\big)}{|D\bar u - D\bar v|}\right]^{1+\chi} |D\bar u - D\bar v|\, dx \le c, \tag{5.31}$$

that is (5.7), once performing a simple change of variable on the left-hand side and recalling the definition of g_χ in (5.2). □

Once having the previous Lemma at hand, a minor modification of the proof allows to get the following similar result which, despite being surely not optimal, it is therefore sufficient for our purposes. We introduce the further following function for the sake of shortness:

$$h_\chi(t) := \left[\frac{g(t)}{t}\right]^{1+\chi} = \frac{g_\chi(t)}{t}.$$

Corollary 5.2. *Let $u \in W^{1,G}(\Omega)$ and $v \in u + W^{1,G}_0(B_R)$ as in Lemma 5.1. Then the following comparison estimate holds true:*

$$\fint_{B_R} h_\chi\big(|Du - Dv|\big)\, dx \le c\, h_\chi(A) \quad \textit{with} \quad A := g^{-1}\left(\frac{|\mu|(B_R)}{R^{n-1}}\right),$$

for χ as in (5.8) and with a constant c depending on $n, \nu, \delta, g_0, \chi$.

Proof. We rescale both the function as in *Step 1* of the proof of Lemma 5.1. Having after *Step 2* estimate (5.12) at hand, we can estimate

$$\fint_{B_1} h_\chi\big(|Du - Dv|\big)\, dx = c(n) \int_{B_1} \left[\frac{g\big(|Du - Dv|\big)}{|Du - Dv|}\right]^{1+\chi} dx$$
$$= c \int_{B_1 \cap \{|Du - Dv| \le 1\}} \dots\, dx + c \int_{B_1 \cap \{|Du - Dv| > 1\}} \dots\, dx$$
$$\le c(n, \delta, g_0, \chi) + c \int_{B_1} \left[\frac{g\big(|Du - Dv|\big)}{|Du - Dv|}\right]^{1+\chi} |Du - Dv|\, dx \le c.$$

At this point performing a rescaling similar to that in (5.31) gives (5.32). □

Now another, albeit similar, proof of this kind:

Lemma 5.3. *Let $u \in W^{1,G}(\Omega)$ be the solutions to the equation (1.8) and $v \in u + W^{1,G}_0(B_R)$ the solution to the problem (5.1) on B_R. Then the following estimate holds true:*

$$\fint_{B_R} \big[g\big(|Du - Dv|\big)\big]^\xi\, dx \le c \left[\frac{|\mu|(B_R)}{R^{n-1}}\right]^\xi \tag{5.32}$$

for

$$\xi \in \left[1, \min\left\{\frac{g_0 + 1}{g_0}, \frac{n}{n-1}\right\}\right) \tag{5.33}$$

and with a constant c depending on n, ν, δ, g_0, ξ.

Proof. Since the proof is very similar to that of Lemma 5.1, we will only highlight the main points. First of all we perform a scaling as in (5.9) and subsequent lines; therefore from now on we can suppose $B_R = B_1$ and moreover $|\mu|(B_1) = 1$. Introducing the auxiliary function

$$f_\xi(t) := \xi \int_0^t \frac{[g(s)]^\xi}{s}\, ds;$$

for ξ as in (5.33) and noting that we have $f_\xi(t) \approx [g(t)]^\xi$ (and $f_\xi(1) \le 1$, for later use), all that we want to prove now is

$$\int_{B_1} f_\xi\big(|Du - Dv|\big)\,dx \le c(n,\nu,\delta,g_0). \tag{5.34}$$

Exactly as in *Step 3* of the proof of Lemma 5.1, (5.32) will follow simply by coming back to $[g(Du - Dv)]^\xi$ and using the scaling of the equation.

The slow growth case. We first consider the case where

$$\int^\infty \left(\frac{s}{G(s)}\right)^{\frac{1}{n-1}} ds = \infty. \tag{5.35}$$

We moreover define

$$\mathcal{F} := \left(\int_{B_1} \mathfrak{f}_\xi\big(|Du - Dv|\big)\,dx\right)^{\frac{1}{n}}.$$

We choose in (5.13) the test function

$$\varphi \equiv T_k\left(\frac{u - v}{c_S(n)\,\mathcal{F}}\right) \in W^{1,G}_0(B_1)$$

for $k \in \mathbb{N}_0$, being $c_S(n)$ being the constant appearing in (3.16) and recalling the definition of the truncation operator T_k in (5.16). Also here we can suppose $\mathcal{F} \ge 1$ and we have $D\varphi = \frac{D(u-v)}{c_S(n)\,\mathcal{F}}\chi_{C_k}$; this time $C_k = B_1 \cap \{|u - v|/(c_S(n)\mathcal{F}) \le k\}$. Therefore we deduce

$$\int_{C_k} G\big(|Du - Dv|\big)\,dx \lesssim c(n)\frac{1}{\mathcal{F}}\int_{B_1} \langle a(Du) - a(Dv), D\varphi\rangle\,dx \tag{5.36}$$

$$\le c\,\mathcal{F}\left|\int_{B_1} \varphi\,dx\right| \le c\,k\mathcal{F} \tag{5.37}$$

for $k \in \mathbb{N}_0$, with $c \equiv c(n,\nu,\delta,g_0)$. Similarly

$$\int_{D_k} G\big(|Du - Dv|\big)\,dx \le c\,\mathcal{F},$$

where $D_k := B_1 \cap \{k < |u-v|/(c_S(n)\mathcal{F}) \le k+1\}$. Now we compute the first derivatives of the function $G\big(f_\xi^{-1}(\cdot)\big)$: note that

$$f'_\xi(t) = \xi\frac{[g(t)]^\xi}{t} \quad \text{and then} \quad \frac{d}{dt} f_\xi\big(G^{-1}(t)\big) = \xi\frac{\big[g(\tau)\big]^{\xi-1}}{\tau}$$

with $\tau := G^{-1}(t)$; moreover

$$\begin{aligned}\frac{d^2}{dt^2} f_\xi\big(G^{-1}(t)\big) &= \xi\frac{\tau(\xi-1)[g(\tau)]^{\xi-2}g'(\tau) - [g(\tau)]^{\xi-1}}{g(\tau)\tau^2}\\ &= \xi\frac{\big[g(\tau)\big]^{\xi-1}}{\tau}\left[(\xi-1)\frac{\tau g'(\tau)}{g(\tau)} - 1\right] < 0\end{aligned}$$

by (1.6) and since $\xi < \frac{1+g_0}{g_0}$. Hence here $t \mapsto f_\xi\big(G^{-1}(t)\big)$ is increasing and concave. Jensen's inequality and (5.36) yield

$$\begin{aligned}\int_{C_k} f_\xi\big(|Du - Dv|\big)\,dx &\le |C_k|\big(f_\xi \circ G^{-1}\big)\left(\fint_{C_k} G\big(|Du - Dv|\big)\,dx\right) \\ &\lesssim c\,|C_k|\big(f_\xi \circ G^{-1}\big)\left(\frac{k\mathcal{F}}{|C_k|}\right) = c\,k\mathcal{F}H_\xi^{-1}\left(\frac{|C_k|}{k\mathcal{F}}\right),\end{aligned} \tag{5.38}$$

where

$$H_\xi^{-1}(t) := t^{1-\xi}\Big[G^{-1}\Big(\frac{1}{t}\Big)\Big]^{-\xi} \approx_\xi t f_\xi\Big(G^{-1}\Big(\frac{1}{t}\Big)\Big). \tag{5.39}$$

Also here computation of derivatives yields $t \mapsto H_\xi^{-1}(t)$ increasing: indeed using (3.5)

$$\frac{d}{dt}H_\xi^{-1}(t) \ge \Big(1-\xi+\frac{\xi}{1+g_0}\Big)t^{-\xi}\Big[G^{-1}\Big(\frac{1}{t}\Big)\Big]^{\xi} > 0$$

since $\xi < \frac{1+g_0}{g_0}$. Hence there holds the subadditivity property similar as that in (5.6). By arguments we have for the integrals over D_k

$$\int_{D_k} f_\xi\big(|Du-Dv|\big)\,dx \lesssim \mathcal{F}H_\xi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big).$$

We have as in (5.22):

$$\int_{B_1} f_\xi\big(|Du-Dv|\big)\,dx \le c\,\mathcal{F}\Big[H_\xi^{-1}\Big(\frac{|B_1|}{\mathcal{F}}\Big)+\sum_{k=1}^\infty H_\xi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big)\Big].$$

As in (5.15) we modify the function f_ξ linearly near zero in order to use Sobolev's embedding. Define

$$\mathfrak{f}_\xi(t) := \begin{cases} 0 & t=0,\\ f_\xi(1)t & \text{for } t\in(0,1),\\ f_\xi(t) & \text{for } t\in[1,\infty).\end{cases}$$

We denote also here the Sobolev's conjugate function $(\mathfrak{f}_\xi)_n := \mathfrak{f}_\xi \circ H_n^{-1}$, with H_n given by $(3.15)_1$ with the choice $A \equiv \mathfrak{f}_\xi$. We infer as in (5.25)

$$|D_k| \le \frac{1}{(\mathfrak{f}_\xi)_n(k)}\int_{D_k}(\mathfrak{f}_\xi)_n\Big(\frac{|u-v|}{c_S(n)\mathcal{F}}\Big)\,dx$$

Note that

$$\int_0\Big(\frac{s}{\mathfrak{f}_\xi(s)}\Big)^{\frac{1}{n-1}}ds<\infty \quad\text{and}\quad \int^\infty\Big(\frac{s}{\mathfrak{f}_\xi(s)}\Big)^{\frac{1}{n-1}}ds=+\infty;$$

the first since $\mathfrak{f}_\xi$ is linear near zero and the second by (5.35), since

$$\mathfrak{f}_\chi(s) \lesssim G(s)\frac{[g(s)]^{\xi-1}}{s} \lesssim_\chi G(s)\,s^{g_0(\xi-1)-1} \le G(s)$$

when $s \ge 1$, since $\chi < \frac{1+g_0}{g_0}$. At this point, using Young's inequality with conjugate functions H_ξ, $\widetilde{H}_\xi$ and with $\varepsilon \in (0,1)$ to be chosen, estimating exactly as in (5.26)

$$\sum_{k=1}^\infty H_\xi^{-1}\Big(\frac{|D_k|}{\mathcal{F}}\Big) \le \frac{\varepsilon}{\mathcal{F}}\int_{B_1} f_\xi\big(|Du-Dv|\big)\,dx + \frac{c(\delta,g_0)\varepsilon}{\mathcal{F}}|B_1| \\ + c(\delta,\varepsilon)\sum_{k=1}^\infty H_\xi^{-1}\Big(\widetilde{H}_\xi\Big(\frac{1}{(\mathfrak{f}_\xi)_n(k)}\Big)\Big).$$

At this point in order to estimate the summation on the right-hand side we deduce the following chain of up-to-constants equivalences: for $\alpha > 0$

$$[\widetilde{H}_\xi]^{-1}\Big(\frac{1}{G(\alpha)}\Big) \approx \frac{1}{G(\alpha)}\Big[H_\xi^{-1}\Big(\frac{1}{G(\alpha)}\Big)\Big]^{-1} = [f_\xi(\alpha)]^{-1};$$

for the first one we used (3.12) and for the second one just (5.39). At this point with $\alpha = H_n^{-1}(k)$, using again the definition of H_ξ^{-1} and the fact that $\mathfrak{f}_\xi(H_n^{-1}(k)) = f_\xi(H_n^{-1}(k))$, $k \in \mathbb{N}$

$$H_\xi^{-1}\Big(\widetilde{H_\xi}\Big(\frac{1}{(\mathfrak{f}_\xi)_n(k)}\Big)\Big) \lesssim_\xi \frac{\big[g\big(H_n^{-1}(k)\big)\big]^{\xi-1}}{H_n^{-1}(k)}. \tag{5.40}$$

The convergence of the series is hence equivalent to fact

$$\int_1^\infty \frac{\big[g\big(H_n^{-1}(\sigma)\big)\big]^{\xi-1}}{H_n^{-1}(\sigma)}\,d\sigma < \infty$$

– again a calculation shows that the function in (5.40) is decreasing. Again the change of variable $s = H_n^{-1}(\sigma)$ is allowed; estimating in a completely similar way as in (5.28) and subsequent line we have

$$\int_1^\infty \frac{\big[g\big(H_n^{-1}(\sigma)\big)\big]^{\xi-1}}{H_n^{-1}(\sigma)}\,d\sigma \lesssim_{n,\xi} \int_1^\infty \big[g(s)\big]^{\xi(1-\frac{1}{n})}\frac{ds}{sg(s)}.$$

and the integral is finite since the exponent of $g(s)$ is negative, i.e.

$$\xi\big(1-\frac{1}{n}\big)-1<0$$

since $\xi < \frac{n}{n-1}$. Coming then back to (5.26) and then to (5.22)

$$\begin{aligned}\int_{B_1} f_\xi\big(|Du-Dv|\big)\,dx &\le \tilde{c}\,\mathcal{F}H_\xi^{-1}\Big(\frac{|B_1|}{\mathcal{F}}\Big)\\ &\qquad + \varepsilon\tilde{c}\int_{B_1} f_\xi\big(|Du-Dv|\big)\,dx + c\,\varepsilon + c(n,\delta,g_0,\varepsilon,\chi)\mathcal{F}.\end{aligned}$$

After reabsorbing the second and the fourth term first by an appropriate choice of ε and then by the use of Young's inequality, we estimate

$$\mathcal{F}H_\xi^{-1}\Big(\frac{|B_1|}{\mathcal{F}}\Big) \lesssim_{n,\xi} \mathcal{F}^\xi\Big[G^{-1}(\mathcal{F})\Big]^{-\xi}$$

by the definition of ${H_\xi}^{-1}$. Again using (3.7) and recalling that $\mathcal{F} \ge 1$ gives

$$\mathcal{F}H_\xi^{-1}\Big(\frac{|B_1|}{\mathcal{F}}\Big) \lesssim_{n,\xi} \mathcal{F}^{\xi(1-\frac{1}{1+g_0})} = \mathcal{F}^{\xi\frac{g_0}{1+g_0}}$$

and since the exponent is strictly smaller than one by $\xi < \frac{1+g_0}{g_0}$, we use again Young's inequality to finally get (5.34).

The fast growth case. The case where

$$\int^\infty \Big(\frac{s}{G(s)}\Big)^{\frac{1}{n-1}}\,ds < \infty.$$

is again much simpler. Directly testing (5.13) with $\varphi = u - v \in W_0^{1,G}(B_1)\cap L^\infty$ yields

$$\begin{aligned}\int_{B_1} G\big(|Du-Dv|\big)\,dx &\le c(n,\delta,g_0)\,\|Du-Dv\|_{L^G(B_1)}\\ &\le \frac{1}{2}\int_{B_1} G(|Du-Dv|)\,dx + c(n,\delta,g_0),\end{aligned}$$

by the fact that $|\mu|(B_1) = 1$ and using Young's inequality (5.30). Finally, using Jensen's inequality as in (5.38) and the monotonicity of $t \mapsto \big(f_\chi \circ G^{-1}\big)(t)$

$$\fint_{B_1} f_\xi\big(|Du-Dv|\big)\,dx \le c\,\big(f_\xi\circ G^{-1}\big)\Big(\fint_{B_1} G\big(|Du-Dv|\big)\,dx\Big) \le c$$

and the proof is concluded. □

Lemma 5.4. *Let u and v as above. Then there exists a constant $c \equiv c(n,\delta,g_0,\nu)$ such that*

$$\fint_{B_R} \frac{|V_g(Du)-V_g(Dv)|^2}{(\alpha+|u-v|)^{\xi}}\,dx \le \xi c\,\frac{\alpha^{1-\xi}}{\xi-1}\,\frac{|\mu|(B_R)}{R^n} \tag{5.41}$$

holds whenever $\alpha>0$ and $\xi>1$.

Proof. The proof is exactly the same given in [26] for the standard case, once we replace the monotonicity condition therein considered with (2.5). Note that our monotonicity condition (2.5) reads exactly as the one in [26], once we replace the standard V_s function with the one defined in (2.2). □

Now a list of technical Lemmata. Their proof is only sketched, since they are very similar to those in [26], which are already almost trivial once having at hand the previous results. From now on, u and v will be the functions of Lemma 5.1 and B_R the ball therein appearing.

Lemma 5.5. *Suppose that*

$$g^{-1}\left(\frac{|\mu|(B_R)}{R^{n-1}}\right) \le \lambda \qquad \textit{and} \qquad \fint_{B_R}|Du|\,dx \le \lambda$$

hold; then for a constant c_2 depending on n,ν,L,δ,g_0 there holds

$$\sup_{B_{R/4}}|Dv| \le c_2\lambda. \tag{5.42}$$

Proof. Use Lemma 4.1 and then Lemma 5.1 with $\chi=-1$ to estimate the left-hand side of (5.42):

$$\begin{aligned}\sup_{B_{R/4}}|Dv| &\le \fint_{B_R}|Du-Dv|\,dx + \fint_{B_R}|Du|\,dx\\ &\le c_1 g^{-1}\left(\frac{|\mu|(B_R)}{R^{n-1}}\right) + \fint_{B_R}|Du|\,dx \le (c_1+1)\lambda.\end{aligned}$$

□

Lemma 5.6. *Let $\tilde\eta,\vartheta\in(0,1]$, and suppose that*

$$g^{-1}\left(\frac{|\mu|(B_R)}{R^{n-1}}\right) \le \frac{\tilde\eta^n}{c_1}\vartheta\lambda,$$

where c_1 is the constant appearing in Lemma 5.1 for $\chi=-1$. Then the lower bound

$$\fint_{\tilde\eta B_R}|Du|\,dx - \vartheta\lambda \le \fint_{\tilde\eta B_R}|Dv|\,dx$$

holds.

Proof. Use triangle's inequality and Lemma 5.1 for $\chi=-1$. □

6. Proof of Theorem 1.1

Define the scaling parameter $\eta\in(0,\frac12)$ in the following way:

$$\eta := \left(\frac{1}{10\,2^{3\alpha+10}c_o c_2^2 c_h}\right)^{\frac{1}{\alpha}} \le \min\left\{\left(\frac{1}{2^4 c_h}\right)^{\frac{1}{\alpha}}, \left(\frac{1}{2^{3\alpha+2}c_o\,(48c_2)^2}\right)^{\frac{1}{\alpha}}\right\}. \tag{6.1}$$

Here α is the exponent and c_o, c_h are the constants appearing in Lemma 4.1 and c_2 appears in Lemma 5.5. All these quantities are a priori defined, depending only on n,ν,L,δ,g_0 and therefore also η is a universal constant depending only on n,ν,L,δ,g_0.

For a fixed ball $B_R \equiv B_R(x)$ such that $B_{2R} \subset \Omega$ as in the statement of Theorem 1.1, build the sequence of shrinking balls $\{B_i\}_{i=0,1,\dots}$ defined by

$$B_i := B_{R_i}(x) \qquad \text{where} \qquad R_i := \eta^i R, \tag{6.2}$$

and subsequently the sequence of functions v_i solutions to the homogeneous problem (5.1) in the ball $B_R \equiv B_i$:

$$\begin{cases} \operatorname{div} a(Dv_i) = 0 & \text{in } B_i, \\ v_i = u & \text{on } \partial B_i. \end{cases} \tag{6.3}$$

Lemma 6.1. *Suppose that for a certain index $i \in \mathbb{N}$ and for a number $\lambda > 0$ there holds*

$$g^{-1}\left(\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right) + g^{-1}\left(\frac{|\mu|(B_i)}{r_i^{n-1}}\right) \le \lambda \tag{6.4}$$

and

$$\frac{\lambda}{H} \le |Dv_i| \le H\lambda \quad \text{in } B_{i+1}, \qquad \frac{\lambda}{H} \le |Dv_{i-1}| \le H\lambda \quad \text{in } B_i \tag{6.5}$$

for a constant $H \ge 1$. Then there exists a constant $c_H \equiv c_H(n,\nu,L,\delta,g_0,H)$ such that

$$\fint_{B_{i+1}} |Du - Dv_i|\, dx \le c_H \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]. \tag{6.6}$$

Proof. Define the parameter $\chi > 0$ in the following way:

$$2\chi := \frac{1}{2}\min\left\{\frac{1}{g_0 - 1}, \frac{g_0}{(g_0-1)(n-1)}, \frac{1}{n-1}\right\} \tag{6.7}$$

and let $\xi := 1 + 2\chi$. Note that $\xi < 1^* = n/(n-1)$ and $\chi, \xi \equiv \chi, \xi(n, g_0)$. By (2.3) and by monotonicity (3.1) it follows that

$$\left[\frac{g(|Dv_i|)}{|Dv_i|}\right]^{1+\chi} |Du - Dv_i| \lesssim \left[\frac{g(|Dv_i|)}{|Dv_i|}\right]^{\frac{1+2\chi}{2}} \big|V_g(Du) - V_g(Dv_i)\big|.$$

Recalling the definition of h_χ, taking averages over B_i and using Schwarz-Hölder's inequality yields, for $\alpha > 0$:

$$\begin{aligned} &\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\, dx \qquad\qquad (6.8)\\ &\quad \lesssim \fint_{B_i} \left[\frac{\big|V_g(Du) - V_g(Dv_i)\big|^2}{\big(\alpha + |u - v_i|\big)^\xi}\right]^{\frac{1}{2}} \left[h_{2\chi}\big(|Dv_i|\big)\big(\alpha + |u - v_i|\big)^\xi\right]^{\frac{1}{2}} dx \\ &\quad \lesssim \left[\fint_{B_i} \frac{\big|V_g(Du) - V_g(Dv_i)\big|^2}{\big(\alpha + |u - v_i|\big)^\xi}\, dx\right]^{\frac{1}{2}} \left[\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\big(\alpha + |u - v_i|\big)^\xi\, dx\right]^{\frac{1}{2}}. \end{aligned}$$

Now we use (5.41) to bound the first term of the right-hand side and we choose α such that

$$\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)|u - v_i|^\xi\, dx = \alpha^\xi \fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\, dx.$$

Note that this definition of α makes sense, see (6.11); moreover the integral on the left-hand side is finite, see the calculations after (6.12). With these actions (6.8) takes the form

$$\begin{aligned} \fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\, dx &\le c\left[\alpha^{1-\xi}\frac{|\mu|(B_i)}{r_i^n}\right]^{\frac{1}{2}} \left[\alpha^\xi \fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\, dx\right]^{\frac{1}{2}} \\ &= c\left[\frac{\alpha}{r_i}\frac{|\mu|(B_i)}{r_i^{n-1}} \fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\, dx\right]^{\frac{1}{2}} \end{aligned}$$

with $c \equiv c(n, \nu, \delta, g_0)$. Note that since $t \mapsto g(t)/t$ is increasing and satisfies a doubling property, then Remark 3.1 applies; therefore for h_χ a subadditivity property similar to (3.8), holds true, with a constant depending on n, g_0. Hence

$$\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\, dx \lesssim_n \fint_{B_i} h_{2\chi}\big(|Du - Dv_{i-1}|\big)\, dx + \fint_{B_i} h_{2\chi}\big(|Du - Dv_i|\big)\, dx + \fint_{B_i} h_{2\chi}\big(|Dv_{i-1}|\big)\, dx =: I_1 + I_2 + I_3.$$

Before estimating term by term, we introduce for ease of notation the following quantities:

$$A_{i-1} := g^{-1}\left(\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right), \qquad A_i := g^{-1}\left(\frac{|\mu|(B_i)}{r_i^{n-1}}\right); \tag{6.9}$$

note that $A_i + A_{i-1} \le \lambda$ by (6.4) and therefore by monotonicity $h_\chi(A_i) \le h_\chi(\lambda)$ and analogously for A_{i-1}. Now by the pointwise estimate $(6.5)_2$ and the definition of $h_{2\chi}$ we have $I_3 \lesssim c(n, g_0, H)\big[g(\lambda)/\lambda\big]^{\xi}$. We estimate I_1 using Corollary 5.2, due to (6.7):

$$I_1 \le \eta^{-n} \fint_{B_{i-1}} h_{2\chi}\big(|Du - Dv_{i-1}|\big)\, dx \le c\, h_{2\chi}(A_{i-1}) \le c\, h_{2\chi}(\lambda),$$

with $c \equiv c(n, \nu, L, \delta, g_0, H)$. The estimate for I_2 is analogous and even more direct. Hence we have, using Young's inequality with $\varepsilon \in (0,1)$ to be chosen and the definition of $h_{2\chi}(\lambda)$

$$\begin{aligned}\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\, dx &\le c\left[\frac{\alpha}{r_i}\frac{|\mu|(B_i)}{r_i^{n-1}}\left[\frac{g(\lambda)}{\lambda}\right]^{1+2\chi}\right]^{\frac12}\\ &\le \varepsilon\frac{\alpha}{r_i}\left[\frac{g(\lambda)}{\lambda}\right]^{1+\chi} + c(\varepsilon)\frac{|\mu|(B_i)}{r_i^{n-1}}\left[\frac{g(\lambda)}{\lambda}\right]^{\chi}.\end{aligned} \tag{6.10}$$

To conclude the proof, we need to estimate α. As a first step we bound from below

$$\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)\, dx \ge \eta^n \fint_{B_{i+1}} h_{2\chi}\big(|Dv_i|\big)\, dx \ge c\left[\frac{g(\lambda)}{\lambda}\right]^{1+2\chi} \tag{6.11}$$

and therefore

$$\alpha^\xi \le c\left[\frac{\lambda}{g(\lambda)}\right]^{\xi} \fint_{B_i} h_{2\chi}\big(|Dv_i|\big)|u - v_i|^\xi\, dx, \tag{6.12}$$

with $c \equiv c(n, \nu, L, \delta_0, H)$. We split the latter averaged integral in the following way

$$\begin{aligned}\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)|u - v_i|^\xi\, dx &\le c\fint_{B_i} h_{2\chi}\big(|Dv_i - Dv_{i-1}|\big)|u - v_i|^\xi\, dx\\ &\quad + c\fint_{B_i} h_{2\chi}\big(|Dv_{i-1}|\big)|u - v_i|^\xi\, dx =: c\,(II_1) + c\,(II_2).\end{aligned}$$

We begin with the easier (II_2): since we have the pointwise estimate $h_{2\chi}(|Dv_{i-1}|) \approx_{n,H} h_{2\chi}(\lambda) = \big[g(\lambda)/\lambda\big]^{1+2\chi}$ on B_i, using standard Sobolev's embedding by (6.7), subadditivity for h_χ and recalling that $\xi = 1 + 2\chi$, we infer

$$\begin{aligned}\frac{(II_2)^{\frac{1}{\xi}}}{r_i} &\le c\,\frac{g(\lambda)}{\lambda}\fint_{B_i} |Du - Dv_i|\, dx = c\left[\frac{\lambda}{g(\lambda)}\right]^{\chi} h_\chi(\lambda)\fint_{B_i} |Du - Dv_i|\, dx\\ &\le c\left[\frac{\lambda}{g(\lambda)}\right]^{\chi}\fint_{B_i} h_\chi\big(|Dv_{i-1}|\big)|Du - Dv_i|\, dx\\ &\le c\left[\frac{\lambda}{g(\lambda)}\right]^{\chi}\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\, dx + c\left[\frac{\lambda}{g(\lambda)}\right]^{\chi}(III)\end{aligned} \tag{6.13}$$

see next estimate for the definition of (III), with $c \equiv c(n, \delta, g_0, H)$. Moreover, again subadditivity for h_χ gives

$$
\begin{aligned}
(III) &:= \fint_{B_i} h_\chi\big(|Dv_i - Dv_{i-1}|\big)|Du - Dv_i|\,dx \\
&\lesssim_n \fint_{B_i} g_\chi\big(|Du - Dv_i|\big)\,dx + \fint_{B_i} h_\chi\big(|Du - Dv_{i-1}|\big)|Du - Dv_i|\,dx \qquad (6.14)
\end{aligned}
$$

While the first integral is less or equal than $c_1\, g_\chi(A_{i-1})$, with c_1 depending on $data$, by (5.7), for the second one we need the pointwise estimate

$$
\widetilde{g_\chi}\big(h_\chi(t)\big) = \widetilde{g_\chi}\left(\frac{g_\chi(t)}{t}\right) \le g_\chi(t)
$$

see (3.11). Therefore Young's inequality with conjugate functions g_χ and $\widetilde{g_\chi}$ gives

$$
\begin{aligned}
\fint_{B_i} h_\chi\big(|Du - Dv_{i-1}|\big)|Du - Dv_i|\,dx &\le c\fint_{B_i} g_\chi\big(|Du - Dv_i|\big)\,dx \\
&\quad + c\fint_{B_i} g_\chi\big(|Du - Dv_{i-1}|\big)\,dx \le c\, g_\chi(A_{i-1}),
\end{aligned}
$$

as for the first term in the second line of (6.14). Note that here we used $g_\chi(A_i) + g_\chi(A_{i-1}) \le c(n, \nu, L, \delta, g_0)\, g_\chi(A_{i-1})$, following from (6.2) and the monotonicities of both the measure $|\mu|$ and g_χ. We here have the following algebraic manipulation:

$$
\left[\frac{\lambda}{g(\lambda)}\right]^\chi g_\chi\big(A_{i-1}\big) = \left[\frac{\lambda}{g(\lambda)}\right]^\chi \left[\frac{g\big(A_{i-1}\big)}{A_{i-1}}\right]^\chi g\big(A_{i-1}\big) \le \frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}, \qquad (6.15)
$$

since $t \mapsto g(t)/t$ is monotone and $A_{i-1} \le \lambda$ by (6.4). The reader here may need to recall also the definition of A_{i-1} in (6.9). Therefore, taking into account (6.15), we have proved

$$
\left[\frac{\lambda}{g(\lambda)}\right]^\chi (III) \le c\left[\frac{\lambda}{g(\lambda)}\right]^\chi g_\chi\big(A_{i-1}\big) \le c\,\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}.
$$

Hence, putting the last estimate into (6.13), all in all we have

$$
(II_2) \le c\, r_i^\xi \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^\xi + c\, r_i^\xi \left[\frac{\lambda}{g(\lambda)}\right]^{\chi\xi} \left[\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\,dx\right]^\xi. \qquad (6.16)
$$

Now we come to the estimate of (II_1): we use Young's inequality with $k(t) := [2g(t^{\frac{1}{\xi}})]^\xi$ and $\tilde{k}(t)$ and then we estimate the first term with Hölder's inequality and Proposition 3.4:

$$
\begin{aligned}
(II_1) &= \fint_{B_i} \left[\frac{g\big(|Dv_i - Dv_{i-1}|\big)}{|Dv_i - Dv_{i-1}|}\right]^\xi |u - v_i|^\xi\,dx \qquad (6.17) \\
&\le (2r_i)^\xi \fint_{B_i} \left[g\left(\frac{|u - v_i|}{r_i}\right)\right]^\xi dx + r_i^\xi \fint_{B_i} \tilde{k}\left(\left[\frac{g\big(|Dv_i - Dv_{i-1}|\big)}{|Dv_i - Dv_{i-1}|}\right]^\xi\right) dx \\
&\le c\, r_i^\xi \left[\fint_{B_i} g\big(|Du - Dv_i|\big)\,dx\right]^\xi + r_i^\xi \fint_{B_i} \tilde{k}\left(\left[\frac{g\big(|Dv_i - Dv_{i-1}|\big)}{|Dv_i - Dv_{i-1}|}\right]^\xi\right) dx.
\end{aligned}
$$

While for the first term we then have

$$
r_i^\xi \left[\fint_{B_i} g\big(|Du - Dv_i|\big)\,dx\right]^\xi \le c\, r_i^\xi \big[g\big(A_i\big)\big]^\xi \le c\, r_i^\xi \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^\xi
$$

from Lemma 5.1 or 5.3, for the second one we need, for $\alpha \ge 0$, the following estimate:

$$
\tilde{k}(\alpha^\xi) = \sup_{s>0}\left\{\alpha^\xi s - \big[2g(s^{\frac{1}{\xi}})\big]^\xi\right\} = \sup_{\sigma>0}\left\{\alpha^\xi \sigma^\xi - 2^\xi\big[g(\sigma)\big]^\xi\right\}
$$

$$\leq 2^\xi \left[\sup_{\sigma>0}\Big\{\alpha\sigma - g(\sigma)\Big\}\right]^\xi = c(n,g_0)\,\big[\tilde g(\alpha)\big]^\xi$$

since $\xi \geq 1$. Therefore with $\alpha = g\big(|Dv_i - Dv_{i-1}|\big)/|Dv_i - Dv_{i-1}|$, using (3.11) and Lemma 5.3

$$\begin{aligned}
\fint_{B_i} \tilde k\bigg(\bigg[\frac{g\big(|Dv_i - Dv_{i-1}|\big)}{|Dv_i - Dv_{i-1}|}\bigg]^\xi\bigg)\,dx &\leq \fint_{B_i}\bigg[\tilde g\bigg(\frac{g\big(|Dv_i - Dv_{i-1}|\big)}{|Dv_i - Dv_{i-1}|}\bigg)\bigg]^\xi dx\\
&\leq \fint_{B_i}\big[g\big(|Dv_i - Dv_{i-1}|\big)\big]^\xi\,dx\\
&\lesssim_n \fint_{B_i}\big[g\big(|Du - Dv_i|\big)\big]^\xi\,dx + \eta^{-n}\fint_{B_{i-1}}\big[g\big(|Du - Dv_{i-1}|\big)\big]^\xi\,dx\\
&\leq c\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^\xi.
\end{aligned}$$

Plugging these two estimates into (6.17) and also taking into account (6.16) yields

$$(II_1)+(II_2) \leq c\,r_i^\xi\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]^\xi + c\,r_i^\xi\left[\frac{\lambda}{g(\lambda)}\right]^{\chi\xi}\left[\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du-Dv_i|\,dx\right]^\xi;$$

therefore finally we estimate α as follows: from (6.12)

$$\begin{aligned}
\frac{\alpha}{r_i} &\leq \frac{c}{r_i}\,\frac{\lambda}{g(\lambda)}\left[\fint_{B_i} h_{2\chi}\big(|Dv_i|\big)|u-v_i|^\xi\,dx\right]^{\frac{1}{\xi}}\\
&\leq c\,\frac{\lambda}{g(\lambda)}\,\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}} + \tilde c\left[\frac{\lambda}{g(\lambda)}\right]^{1+\chi}\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\,dx.
\end{aligned}$$

Inserting this estimate into (6.10) and choosing $\varepsilon \equiv \varepsilon(n,\nu,L,\delta,g_0,H) \in (0,1)$ small enough - i.e. $\varepsilon = 1/(2\tilde c)$ leads to

$$\frac12\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\,dx \leq c\,\frac{|\mu|(B_i)}{r_i^{n-1}}\left[\frac{g(\lambda)}{\lambda}\right]^\chi.$$

Now (6.6) plainly follows taking into account that

$$\fint_{B_i} h_\chi\big(|Dv_i|\big)|Du - Dv_i|\,dx \gtrsim_{n,H} \eta^n\left[\frac{g(\lambda)}{\lambda}\right]^{1+\chi}\fint_{B_{i+1}}|Du - Dv_i|\,dx.$$

□

Let $x\in\Omega$ be a Lebesgue's point of Du and let $B_{2R}(x)\subset\Omega$. Define the quantity

$$\lambda := g^{-1}\bigg(H_1\,g\bigg(\fint_{B_R}|Du|\,dx\bigg) + H_2\,\mathbf{I}_1^{|\mu|}(x,2R)\bigg), \tag{6.18}$$

where the constants H_1, H_2 will be fixed in a few lines, in a way making them depending only on n,ν,L,δ,g_0. We want to prove that

$$|Du(x)| \leq \lambda, \tag{6.19}$$

and (1.9) will follow simply taking $c := \max\{H_1,H_2\}$. Without loss of generality we can clearly assume $\lambda > 0$, whether this were not the case (6.19) would trivially follow by the monotonicity of the vector field.

Step 1: the choice of the constants. With $i \in \mathbb{N}+1$ define the quantity

$$C_i := \sum_{j=i-2}^{i} \fint_{B_j} |Du|\,dx + \eta^{-n} E(Du, B_i) \le 5\eta^{-3n} \fint_{B_{i-2}} |Du|\,dx. \tag{6.20}$$

Note that the inclusions $B_{i+1} = \eta B_i \subset \frac{1}{4} B_i \subset B_i$ hold. Take $k \in \mathbb{N}$, $k \ge 3$ as the smallest integer such that

$$(8\eta^k)^\alpha \le \eta^n \frac{1}{128 c_o c_2}; \tag{6.21}$$

here c_o is the constant of Lemma 4.1 and c_2 is the one appearing in Lemma 5.5. Once fixed $k \equiv k(n, \nu, L, \delta, g_0)$ in such way, fix the constant H_1 and H_2 as follows

$$H_1 = \left(10\eta^{-4n}\right)^{g_0}, \qquad H_2 := 2^{7g_0} c_1^{g_0} \eta^{-ng_0(k+1+\frac{1}{\delta})} c_{200c_2}.$$

here c_1 is the constant appearing in Lemma 5.1 for $\chi = -1$ and c_{200c_2} is the constant c_H appearing in Lemma 6.1 for the choice $H = 200c_2$. Note that the dependences of k and η yield that both H_1 and H_2 are *a priori* constants depending only on n, ν, L, δ, g_0; moreover with this choice there holds – recall that $k \ge 3$

$$\eta^{-\frac{n}{\delta}} H_2^{-\frac{1}{g_0}} \le \frac{1}{2}, \qquad \eta^{-\frac{n}{\delta}} H_2^{-\frac{1}{g_0}} \le \frac{\eta^{nk}}{96 c_1}, \qquad c_1 \eta^{-n(k+\frac{1}{\delta})} H_2^{-\frac{1}{g_0}} \le \frac{\eta^n}{128},$$
$$4c_{200c_2} \frac{\eta^{-2n}}{H_2} \le \frac{1}{4}, \qquad 8c_{200c_2} \frac{\eta^{-3n}}{H_2} \le \frac{1}{4}, \tag{6.22}$$

which we shall use in this order. Moreover the choice of H_1 implies

$$C_2 + C_3 \le 10\eta^{-4n} \fint_{B_0} |Du|\,dx \le 10\eta^{-4n} H_1^{-\frac{1}{g_0}} \lambda \le \frac{\lambda}{8} \le \lambda. \tag{6.23}$$

We recall the dyadic decomposition

$$\eta^n \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{R_j^{n-1}} \le \int_0^{2R} \frac{|\mu|(B_\rho(x))}{\rho^{n-1}} \frac{d\rho}{\rho} = \mathbf{I}_1^{|\mu|}(x, 2R),$$

see [26]. Hence for every $i \in \mathbb{N}_0$ using (6.22) we have

$$g^{-1}\left(\frac{|\mu|(B_i)}{R_i^{n-1}}\right) \le g^{-1}\left(\sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{R_j^{n-1}}\right) \le g^{-1}\left(\eta^{-n} \mathbf{I}_1^{|\mu|}(x, 2R)\right)$$
$$\le \eta^{-\frac{n}{\delta}} H_2^{-\frac{1}{g_0}} \lambda \le \frac{\lambda}{2}. \tag{6.24}$$

Step 2: the exit time and after the exit time. Now we state that we can suppose that there exists an “exit time” index $i_e \ge 3$, see (6.23), such that

$$C_{i_e} \le \frac{\lambda}{8} \qquad \text{but} \qquad C_j > \frac{\lambda}{8} \qquad \text{for every } j > i_e. \tag{6.25}$$

Indeed, on the contrary, we would have $C_{i_h} \le \lambda/8$ for an increasing subsequence $\{i_h\}$ and then, being x a Lebesgue point for Du,

$$|Du(x)| \le \lim_{h\to\infty} \fint_{B_{i_h}} |Du|\,dx \le \frac{\lambda}{8}$$

and the proof would be finished. Now an important Lemma which asserts that after the exit time the gradient Dv_i is far away from zero; this finally gives meaning to the assumption (6.5) we imposed on the Dv_is.

Lemma 6.2. *Suppose that*

$$\fint_{B_i} |Du|\,dx \leq \lambda \tag{6.26}$$

holds for a certain index $i \in \mathbb{N}$, $i \geq i_e - 2$, for $\lambda > 0$ defined in (6.18). Then

$$\frac{\lambda}{200c_2} \leq |Dv_i| \leq c_2\lambda \qquad \text{in } B_{i+1}, \tag{6.27}$$

where c_2 is the constant appearing in Lemma 5.5.

Proof. The right-hand side estimate in (6.27) is a consequence of Lemma 5.5 applied with $B_R \equiv B_i$, which gives

$$\sup_{B_i/4} |Dv_i| \leq c_2\lambda. \tag{6.28}$$

Note that the assumptions of the Lemma are satisfied since (6.26) and (6.24) hold, and moreover $B_{i+1} \subset B_i/4$. In order to prove the left-hand side inequality we want to use Lemma 4.3 with $\sigma \equiv \eta$, $B \equiv B_i$ so that $\sigma B = B_{i+1}$; since we already have (6.28) we just need to prove

$$\frac{\lambda}{C} \leq \fint_{B_{i+m}} |Dv_i|\,dx$$

for some $m \in \mathbb{N}$ and some $C \geq 1$. We start by proving that

$$\frac{\lambda}{16} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} |Dv_i|\,dx, \tag{6.29}$$

where $k \geq 3$ is the number defined in (6.21). By (6.24) we can apply Lemma 5.6 three times, with $B_R \equiv B_i$, respectively $\tilde\eta \equiv \eta^k, \eta^{k-1}, \eta^{k-2}$ and $\vartheta \equiv 1/96$. Note indeed that from (6.24) and condition (6.22) follows

$$g^{-1}\left(\frac{|\mu|(B_i)}{R_i^{n-1}}\right) \leq \eta^{-\frac{n}{\delta}} H_2^{-\frac{1}{90}} \lambda \leq \frac{\eta^{nk}}{c_1}\frac{1}{96}\lambda \leq \frac{\eta^{nj}}{c_1}\vartheta\lambda$$

for $j = k-2, k-1, k$. Summing up the resulting inequalities gives

$$\begin{aligned} C_{i+k} - \eta^{-n}E(Du, B_{i+k}) - 3\lambda\vartheta &= \sum_{j=i-2+k}^{i+k} \fint_{B_j} |Du|\,dx - 3\lambda\vartheta \\ &\leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} |Dv_j|\,dx. \end{aligned}$$

Since $i \geq i_e - 2$ and $k \geq 3$, we have $i + k > i_e$ and subsequently by the definition of the exit time index $C_{i+k} \geq \lambda/8$. Using this fact and the value of ϑ in the inequality above gives

$$\frac{\lambda}{8} - \eta^{-n}E(Du, B_{i+k}) - \frac{\lambda}{32} \leq \sum_{j=i-2+k}^{i+k} \fint_{B_j} |Dv_j|\,dx. \tag{6.30}$$

In order to estimate the excess term first we note that enlarging the domain of integration from B_{i+k} to B_i and using (6.24) gives

$$\begin{aligned} \fint_{B_{i+k}} |Du - Dv_i|\,dx &\leq \frac{|B_i|}{|B_{i+k}|} c_1 g^{-1}\left(\frac{|\mu|(B_i)}{R_i^{n-1}}\right) \\ &\leq c_1\eta^{-kn-\frac{n}{\delta}} H_2^{-\frac{1}{90}} \lambda \leq \eta^n \frac{\lambda}{128} \end{aligned} \tag{6.31}$$

where we used (5.7) with $\chi = -1$. Lemma 4.1 applied with $B_R \equiv B_i/4$, $B_{r/2} \equiv B_{i+k}$ using the just proved right-hand side inequality of (6.27) and the definition of k gives

$$2 \operatorname*{osc}_{B_{i+k}} Dv_i \leq 2c_o(8\eta^k)^\alpha c_2\,\lambda \leq \eta^n \frac{\lambda}{64}$$

using (6.21), so that, with the help of (2.7) and then of (6.31) we infer

$$\begin{aligned} E(Du, B_{i+k}) &\leq 2 \fint_{B_{i+k}} \left|Du - (Dv_i)_{B_{i+k}}\right| dx \\ &\leq 2 \fint_{B_{i+k}} \left|Dv_i - (Dv_i)_{B_{i+k}}\right| dx + 2 \fint_{B_{i+k}} |Du - Dv_i|\, dx \\ &\leq 2 \operatorname*{osc}_{B_{i+k}} Dv_i + 2\eta^n \frac{\lambda}{128} \leq \eta^n \frac{\lambda}{32}. \end{aligned} \tag{6.32}$$

Inserting this last estimate into (6.30) gives (6.29). (6.29) in turn implies that there exists an index $m \in \{k-2, k-1, k\}$ such that

$$\fint_{B_m} |Dv_i|\, dx \geq \frac{1}{3}\frac{\lambda}{16} \geq \frac{\lambda}{48c_2}.$$

Hence now we can apply Lemma 4.3 with the choices listed just above and with $C \equiv 48c_2$. Note that the condition on σ (4.19) holds true by (6.1); then the choice $\sigma = \eta$ is allowed. Therefore we can conclude with

$$\frac{\lambda}{200c_2} \leq |Dv_i| \qquad \text{in } B_{i+1}$$

which coupled with (6.28) gives (6.27). □

Step 3: Iteration.

Lemma 6.3. *Suppose that for some $i \in \mathbb{N}$, $i \geq i_e - 1$ there holds*

$$\fint_{B_{i-1}} |Du|\, dx \leq \lambda \qquad \text{and} \qquad \fint_{B_i} |Du|\, dx \leq \lambda. \tag{6.33}$$

Then there exists a constant c_3 depending on n, ν, L, δ, g_0 such that

$$E(Du, B_{i+2}) \leq \frac{1}{4} E(Du, B_{i+1}) + c_3 \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{i-1})}{R_{i-1}^{n-1}}\right]$$

holds true. c_3 has the expression $4\eta^{-n} c_{200c_2}$, where c_{200c_2} is the constant of Lemma 6.1 for $H = 200c_2$.

Proof. We clearly want to apply Lemma 6.1. Assumption (6.4) is satisfied as a consequence of (6.24), while for (6.5) we appeal to Lemma 6.2: since $i \geq i_e - 1$, obviously $i - 1 \geq i_e - 2$ and then we can use estimate (6.27) both for Dv_i in B_{i+1} and Dv_{i-1} in B_i, i.e.

$$\frac{\lambda}{200c_2} \leq |Dv_{i-1}| \leq c_2\lambda \qquad \text{in } B_i, \qquad \frac{\lambda}{200c_2} \leq |Dv_i| \leq c_2\lambda \qquad \text{in } B_{i+1}.$$

Hence assumptions (6.5) are satisfied with $H \equiv 200c_2$, so we have

$$\fint_{B_{i+1}} |Du - Dv_i|\, dx \leq c_{200c_2} \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{i-1})}{R_{i-1}^{n-1}}\right], \tag{6.34}$$

where c_{200c_2} is a constant depending on n, ν, L, δ, g_0. Estimate (4.3) applied with $B_r, B_R \equiv B_{i+2}, B_{i+1}$ gives using (6.1)

$$E(Dv_i, B_{i+2}) \leq 2^{-4} E(Dv_i, B_{i+1})$$

so we get the thesis performing a computation similar to (6.32):

$$E(Du, B_{i+2}) \leq 2E(Dv_i, B_{i+2}) + 2 \fint_{B_{i+2}} |Du - Dv_i|\, dx$$

$$
\begin{aligned}
&\leq 2^{-3}E(Dv_i, B_{i+1}) + 2\eta^{-n} \fint_{B_{i+i}} |Du - Dv_i|\, dx \\
&\leq 2^{-2}E(Du, B_{i+1}) + 2(\eta^{-n}+1) \fint_{B_{i+i}} |Du - Dv_i|\, dx \\
&\leq 2^{-2}E(Du, B_{i+1}) + c_3 \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right], \qquad (6.35)
\end{aligned}
$$

where we used (6.34). The proof is concluded. □

Now we proceed with the proof of Theorem 1.1 and we define

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

Recalling the definition in (6.20) and (6.25), we have

$$
\sum_{j=i_e-2}^{i_e} m_j + \eta^{-n} A_{i_e} \leq C_{i_e} \leq \frac{\lambda}{8}. \qquad (6.36)
$$

Our goal is to prove by induction that

$$
m_j + A_j \leq \lambda \qquad \text{for all } j \geq i_e. \qquad (6.37)
$$

The case $j = i_e$ holds true from the definition of C_{i_e} and the exit time (6.25):

$$
m_{i_e} + A_{i_e} \leq 3 \fint_{B_{i_e}} |Du|\, dx \leq 3\frac{\lambda}{8}.
$$

Assume now that (6.37) holds true for $j = i_e, \dots, i$. By (6.36) for $j = i_e, \dots, i$ and directly from the definition of C_{i_e} and from (6.25) for $j = i_e - 2, i_e - 1$ – indeed for these two exponents $A_j \leq 2 \fint_{B_j} |Du|\, dx$ – we have in particular that

$$
\fint_{B_j} |Du|\, dx \leq \lambda \qquad \text{for } j = i_e - 2, \dots, i.
$$

Since assumption (6.33) is satisfied, we can apply the excess decay Lemma 6.3 that gives

$$
A_{j+2} \leq \frac{1}{4} A_{j+1} + c_3 \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{j-1})}{R_{j-1}^{n-1}}\right] \qquad \text{for } j = i_e - 1, \dots, i. \qquad (6.38)
$$

When $j = i - 1$ the previous inequality in particular gives

$$
A_{i+1} \leq \frac{1}{4} A_i + c_3 \frac{\lambda}{g(\lambda)} \left[\frac{|\mu|(B_{i-2})}{R_{i-2}^{n-1}}\right] \leq \frac{\lambda}{4} + \frac{1}{4}\frac{\lambda}{g(\lambda)} g(\lambda) \leq \frac{\lambda}{2}. \qquad (6.39)
$$

since by inductive hypothesis $A_i \leq \lambda$ and since the following inequality holds true for all i (recall $i \geq i_e \geq 3$):

$$
c_3 \frac{|\mu|(B_{i-2})}{R_{i-2}^{n-1}} \leq c_3 \eta^{-n} \mathbf{I}_1^{|\mu|}(x, 2R) \leq c_3 \frac{\eta^{-n}}{H_2} g(\lambda) \leq \frac{1}{4} g(\lambda),
$$

see (6.22) and recall that $c_3 = 4\eta^{-n} c_{200c_2}$. Moreover, summing (6.38) for $i \in \{i_e - 1, i - 2\}$ and performing some algebraic manipulations leads to

$$
\sum_{j=i_e}^{i} A_j \leq A_{i_e} + \frac{1}{4} \sum_{j=i_e}^{i-1} A_j + c_3 \frac{\lambda}{g(\lambda)} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{R_j^{n-1}}
$$

which gives, after reabsorption,

$$
\sum_{j=i_e}^{i} A_j \leq 2A_{i_e} + 2c_3 \frac{\lambda}{g(\lambda)} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{R_j^{n-1}}. \qquad (6.40)
$$

On the other hand,

$$m_{i+1} - m_{i_e} = \sum_{j=i_e}^{i} (m_{j+1} - m_j) \leq \sum_{j=i_e}^{i} \fint_{B_{j+1}} \left|Du - (Du)_{B_j}\right| dx \leq \sum_{j=i_e}^{i} \frac{|B_j|}{|B_{j+1}|} E(Du, B_j)$$

and therefore, using (6.40), (6.36) and (6.22),

$$\begin{aligned} m_{i+1} &\leq m_{i_e} + \eta^{-n} \sum_{j=i_e}^{i} A_j \\ &\leq m_{i_e} + 2\eta^{-n} A_{i_e} + 2\eta^{-n} c_3 \frac{\lambda}{g(\lambda)} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-1}} \leq 2\frac{\lambda}{8} + \frac{\lambda}{4} \leq \frac{\lambda}{2}. \end{aligned} \tag{6.41}$$

Merging together (6.39) and (6.41) gives $m_{i+1} + A_{i+1} \leq \lambda$. Being $x \in \Omega$ a Lebesgue point of Du then we have

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

7. VERY WEAK SOLUTIONS & SKETCHES OF OTHER PROOFS

We recall here that Theorem 1.1 is stated as *a priori estimate* for solutions belonging to the energy space, while we consider poorly regular data; this choice can be made more clear by the following lines. Once considering measure data as we do, one of the customary approaches to get existence (and usually, at the same time, regularity) results is to smoothen the data by convolution, then to find regular solutions (together with properties stable when passing to the limit) and finally prove convergence, see for instance the classics [2, 3, 34].

After this introduction, we hence quickly show how to extend the potential estimate Theorem 1.1 to a certain class of the so-called *very weak solutions*. Being a the vector field considered in (1.8), for Dirichlet problems of the type

$$\begin{cases} -\operatorname{div} a(Du) = \mu & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{7.1}$$

we give the following definition, recalling the meaning of the Orlicz-Sobolev space $W_0^{1,g}(\Omega)$ described in Paragraph 3.2:

Definition 7.1. *A very weak solution to* (7.1) *is a function* $u \in W_0^{1,g}(\Omega)$ *such that*

$$\int_\Omega \langle a(Du), D\varphi \rangle \, dx = \int_\Omega \varphi \, d\mu \tag{7.2}$$

for every $\varphi \in C_c^\infty(\Omega)$.

Note that the requirement $u \in W^{1,g}(\Omega)$ is the minimal assumption which gives meaning to the distributional formulation (7.2), when the vector field a satisfies assumptions (2.1). Existence of very weak solutions cannot clearly be obtained by using usual monotonicity methods (which would require the right-hand side to belong to the dual of the energy space), but can be deduced by an adaption of the method developed in [2, 3]; this would lead to a particular class of very weak solutions, usually called SOLAs (Solutions Obtained as Limit of Approximations). We briefly recall it; notice however that uniqueness in general remains an open issue in this class of solutions.

Consider a sequence of approximating regular functions $f_k \in L^\infty(\Omega)$ weakly-$*$ converging to μ and such that

$$|f_k|(B_{R+1/k}) = \int_{B_{R+1/k}} |f_k|\,dx \leq |\mu|(B_R).$$

Solve the weak formulation (7.2) with datum f_k, i.e with right-hand side $\int_\Omega f_k\varphi\,dx$, and by monotonicity methods get regular solutions $u_k \in W_0^{1,G}(\Omega)$. By the compactness and truncation arguments in [2, 3] we can get pointwise convergence of both u_k and Du_k to a limit function belonging at least to $W^{1,g}(\Omega)$ (an approach following Lemma 5.3 should give a priori estimates, and therefore existence, in $W^{1,g^\xi}(\Omega)$ with ξ as in (5.33)). Therefore, also using $f_k \rightharpoonup \mu$, we get that the limit function solves the distributional formulation (7.2). Note that due to (1.6) $L^g(\Omega)$ turns out to be reflexive.

After this introduction we can state a variant of Theorem 1.1:

Theorem 7.2. *Let $u \in W^{1,g}(\Omega)$ be a SOLA to the Dirichlet problem (7.1), with μ and a as in Theorem 1.1. Then there exists a constant c, depending on n, ν, L, δ, g_0, such that the pointwise estimate*

$$g\big(|Du(x_0)|\big) \leq c\,\mathbf{I}_1^{|\mu|}(x_0, 2R) + c\,g\bigg(\fint_{B_R(x_0)} |Du|\,dx\bigg)$$

holds for almost every $x_0 \in \Omega$ and for every ball $B_{2R}(x_0) \subset \Omega$.

The proof of the previous Theorem is quite simple: it consists in passing to the limit not in the final estimate, but in the proofs. One, indeed, has in particular to consider the approximating solutions u_k instead of the energy solution u and pass to the limit in the comparison Lemmata (Lemma 5.1 and subsequent ones), using quite standard arguments (as extraction of diagonal subsequences), in order to get that they hold also in the case u just belongs to $W^{1,g}(\Omega)$. Note that one can also see that, if $\mu \in L^\infty$, then actually (1.9) holds at *every point* in Ω, since solutions with regular data have continuous gradient and therefore every $x_0 \in \Omega$ is a Lebesgue's point of Du; for more details, see [27, Section 5] and [29, Section 14].

Now we come to the proof of Theorem 1.3. More precisely, we will just sketch the proof of the following Proposition, since once having the potential estimate (1.9), the Lemmata proved in the previous sections and the Proposition 7.3 at hand, using simple tricks extensively used in the preceding pages, as, for example, subadditivity of Remark 3.1, this proof is just an adaption of the one in [26]. Note that then Corollary 1.4 would follow in a straightforward way, since the locally uniform convergence in (1.10) is obviously implied by (2) and also by (1), see [13, Lemmata 1 & 3].

Proposition 7.3. *Let $u \in W^{1,G}(\Omega)$ be as in Corollary 1.2. If $x \mapsto |\mu|(x, R)$ is locally bounded in Ω for some $R > 0$ and if*

$$\lim_{R\to 0} \frac{|\mu|\big(B_R(x)\big)}{R^{n-1}} = 0 \quad \textit{locally uniformly in } \Omega \textit{ w.r.t. } x, \tag{7.3}$$

then Du is locally VMO-regular in Ω.

Again we recall that the stated regularity of Du means that for every $\Omega' \Subset \Omega$

$$\lim_{R\to 0} \omega(R) = 0 \quad \text{where} \quad \omega(R) \equiv \omega_{\Omega'}(R) := \sup_{\substack{B_r \Subset \Omega' \\ r \leq R}} \fint_{B_r} \big|Du - (Du)_{B_r}\big|\,dx.$$

Proof. Consider an intermediate open set Ω'' such that $\Omega' \Subset \Omega'' \Subset \Omega$. Since by Corollary (1.2) and by our assumptions Du is locally bounded, it makes sense to prove that for every $\varepsilon \in (0,1)$, there exists a positive radius $r_\varepsilon < \operatorname{dist}(\Omega', \partial\Omega'')$, depending on $n, \delta, p_g, \nu, L, \mu(\cdot), \|Du\|_{L^\infty(\Omega'')}, \operatorname{dist}(\Omega', \partial\Omega), \varepsilon$, such that

$$\fint_{B_\rho(y)} \left|Du - (Du)_{B_\rho(y)}\right| dx \le \varepsilon\lambda, \qquad \lambda := \|Du\|_{L^\infty(\Omega'')} \tag{7.4}$$

holds whenever $\rho \in (0, r_\varepsilon)$ and $y \in \Omega'$. This would give the local VMO regularity of Du. For ε given as in the statement and fixed, consider the quantity

$$\eta := \left(\frac{\varepsilon^2}{2^{3\alpha+10} c_o c_2^2 c_h}\right)^{\frac{1}{\alpha}} \le \frac{1}{(2^3 c_h)^{1/\alpha}}$$

where the involved quantities are the ones appearing also in (6.1). This gives that η is a constant depending only on $n, \nu, L, \delta, g_0, \varepsilon$. Then take the constant $c_{16c_2/\varepsilon}$ as the constant c_H appearing in Lemma 6.1 with the choice $H = 16c_2/\varepsilon$. Then it also depends upon $n, \nu, L, \delta, g_0, \varepsilon$. Finally chose a radius $R_0 < \operatorname{dist}(\Omega'', \partial\Omega')$ depending on $n, \nu, L, \delta, g_0, \mu(\cdot), \|Du\|_{L^\infty(\Omega'')}, \operatorname{dist}(\Omega', \partial\Omega), \varepsilon$ such that

$$\sup_{0<\rho\le R_0} \sup_{x\in\Omega'} \frac{|\mu|\big(B_\rho(x)\big)}{\rho^{n-1}} \le g\left(\varepsilon\frac{\eta^{2n}}{4c_1}\left(\frac{\eta^n}{16c_1 c_{16c_2/\varepsilon}}\right)^{1/\delta}\lambda\right); \tag{7.5}$$

this is allowed by (7.3). Finally for $x \in \Omega'$ fixed define for $i \in \mathbb{N}_0$ the chain of ball $B_i \equiv B_{r_i}$ as in (6.2), with radius $r_i := \eta^i r$ and where $r \in (\eta R_0, R_0]$ is fixed. Define subsequently the comparison functions v_i over B_i as in (6.3). Note that by the definition of λ and the fact that $B_i \subset \Omega''$ then

$$\fint_{B_i} |Du|\, dx \le \lambda \tag{7.6}$$

for all $i \in \mathbb{N}_0$. What we want to prove is

$$E(Du, B_{i+2}) \le \varepsilon\lambda \qquad i \in \mathbb{N}. \tag{7.7}$$

Fix hence an index $i \in \mathbb{N}$ and suppose without loss of generality

$$\fint_{B_{i+2}} |Du|\, dx \ge \varepsilon\frac{\lambda}{2}. \tag{7.8}$$

On the contrary (7.7) would plainly follow by triangle's inequality. Now using an approach similar to Lemmata 6.2 and 6.3 we will prove that

$$\begin{aligned} E(Du, B_{i+2}) &\le \frac{\varepsilon}{4} E(Du, B_{i+1}) + 4c_{16c_2/\varepsilon}\eta^{-n}\frac{\lambda}{g(\lambda)}\left[\frac{|\mu|(B_{i-1})}{R_{i-1}^{n-1}}\right] \\ &\le \frac{\varepsilon}{4} E(Du, B_{i+1}) + \frac{\varepsilon}{4}\lambda \end{aligned} \tag{7.9}$$

by (7.5). This would be enough to get (7.7) by induction. First apply Lemma 5.6 with $B_R \equiv B_i$, $\tilde\eta \equiv \eta^2$, $\vartheta \equiv \varepsilon/4$ which, together with (7.8) gives

$$\varepsilon\frac{\lambda}{4} \le \fint_{B_{i+2}} |Du|\, dx - \varepsilon\frac{\lambda}{4} \le \fint_{B_{i+2}} |Dv_i|\, dx.$$

Note now that Lemma 5.5 with $B_R \equiv B_i$ yields in particular $\sup_{B_i/4} |Dv_i| \le c_2\lambda$ by (7.6); therefore Lemma 4.3 with $B \equiv B_i$, $\sigma \equiv \eta$, $m = 2$ and $C = 4c_2/\varepsilon$ gives

$$\varepsilon\frac{\lambda}{16c_2} \le |Dv_i| \qquad \text{in } B_{i+1}.$$

A similar reasoning yields $\varepsilon\lambda/16c_2 \le |Dv_{i-1}| \le c_2\lambda$ in B_i. Then we apply Lemma 6.1 for $H = 16c_2/\varepsilon$ and we have

$$\fint_{B_{i+1}} |Du - Dv_i|\, dx \le c_{16c_2/\varepsilon}\frac{\lambda}{g(\lambda)}\left[\frac{|\mu|(B_{i-1})}{r_{i-1}^{n-1}}\right]$$

At this point since the choice of η gives $E(Dv_i, B_{i+2}) \leq 2^{-3}E(Dv_i, B_{i+1})$ by (4.3), performing a calculation completely similar to (6.35) brings us to (7.9).

A brief argument similar to the one in [26] concludes the proof. Since all these estimates are uniform with respect the choice of $x \in \Omega'$ and the radius $r \in (\eta R_0, R_0]$, then we obtain (7.4) with $r_\varepsilon := \eta^3 R_0$. Indeed let $\rho \leq \eta^3 R_0$: then exists an integer $m \geq 3$ such that $\eta^{m+1}R_0 < \rho \leq \eta^m R_0$. Therefore $\rho = \eta^m r$ for some $r \in (\eta R_0, R_0]$ and (7.4) follows from (7.7). □

Acknowledgements. This research has been supported by the ERC grant 207573 "Vectorial Problems".

REFERENCES

[1] R.A. ADAMS: *Sobolev Spaces*, Academic Press, New York, 1975.

[2] L. BOCCARDO, T. GALLOUËT: Nonlinear elliptic equations with right hand side measures, *Comm. Partial Differential Equations* **17**: 641–655, 1992.

[3] L. BOCCARDO, T. GALLOUËT: Non-linear elliptic and parabolic equations involving measure data, *J. Funct. Anal.* **87**: 149–169, 1989.

[4] M. CHIPOT, L.C. EVANS: Linearization at innity and Lipschitz estimates for certain problems in the calculus of variations, *Proc. R. Soc. Edinb., Sect.* A **102**: 291–303, 1986.

[5] A. CIANCHI: Maximizing the L^∞ norm of the gradient of solutions to the Poisson equation, *J. Geom. Anal.* **2 (6)**: 499–515, 1992.

[6] A. CIANCHI: Continuity properties of functions from Orlicz-Sobolev, spaces and embedding theorems, *Ann. Scuola Norm. Sup. Pisa Cl. Sci (4)* **23 (3)**: 575–608, 1996.

[7] A. CIANCHI: Boundedness of solutions to variational problems under general growth conditions, *Comm. Partial Differential Equations* **22**: 1629–1646. 1997.

[8] A. CIANCHI, N. FUSCO: Gradient regularity for minimizers under general growth conditions, *J. reine angew. Math.* **507**: 15–36, 1999.

[9] A. CIANCHI, V.G. MAZ'YA: Global Lipschitz regularity for a class of quasilinear elliptic equations, *Comm. Partial Differential Equations* **36 (1)**, 100–133, 2011.

[10] A. CIANCHI, V.G. MAZ'YA: Gradient regularity via rearrangements for p-Laplacian type elliptic boundary value problems, *J. Eur. Math. Soc.* (*JEMS*) **16 (3)**: 571–595, 2014.

[11] E. DIBENEDETTO: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations, *Nonlinear Anal.* **7**: 827–850, 1983.

[12] L. DIENING, F. ETTWEIN: Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20**: 523–556, 2008.

[13] F. DUZAAR, G. MINGIONE, Gradient continuity estimates, *Calc. Var. Partial Differential Equations* **39**: 379–418, 2010.

[14] F. DUZAAR, G. MINGIONE, Gradient estimates via linear and nonlinear potentials, *J. Funct. Anal.* **259**: 2961–2998, 2010.

[15] F. DUZAAR, G. MINGIONE: Gradient estimates via non-linear potentials, *Amer. J. Math.* **133 (4)**: 1093–1149, 2011.

[16] N. FUSCO, C. SBORDONE: Higher integrability of the gradient of minimizers of functionals with nonstandard growth conditions, *Comm. Pure Appl. Math.* **43**: 673–683, 1990.

[17] E. GIUSTI: *Direct methods in the Calculus of Variations*, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[18] M. HAVIN, V.G. MAZYA: A nonlinear potential theory, *Russ. Math. Surveys* **27**: 71–148, 1972.

[19] J. HEINONEN, T. KILPELÄINEN, O. MARTIO: *Nonlinear potential theory of degenerate elliptic equations*, Oxford Mathematical Monographs, New York, 1993.

[20] T. KILPELÄINEN: Hölder continuity of solutions to quasilinear elliptic equations involving measures, *Potential Anal.* **3 (3)**: 265–272, 1994.

[21] T. KILPELÄINEN, J. MALÝ: Degenerate elliptic equations with measure data and nonlinear potentials, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (IV) **19 (4)**: 591–613, 1992.

[22] T. KILPELÄINEN, J. MALÝ: The Wiener test and potential estimates for quasilinear elliptic equations, *Acta Math.* **172 (1)**: 137–161, 1994.

[23] T. KUUSI, G. MINGIONE: Universal potential estimates, *J. Funct. Anal.* **262 (10)**: 4205–4638, 2012.

[24] T. KUUSI, G. MINGIONE: Gradient regularity for nonlinear parabolic equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5)* **12 (4)**: 755–822, 2013.

[25] T. KUUSI, G. MINGIONE: New perturbation methods for nonlinear parabolic problems, *J. Math. Pures Appl.* **98**: 390–427, 2012.

[26] T. KUUSI, G. MINGIONE: Linear potentials in nonlinear potential theory, *Arch. Rat. Mech. Anal.* **207 (1)**: 215–246, 2013.

[27] T. KUUSI, G. MINGIONE: Linear potential estimates for nonlinear parabolic equations, *Arch. Rat. Mech. Anal.* **212 (3)**: 727–780, 2014.
[28] T. KUUSI, G. MINGIONE: A nonlinear Stein theorem, *Calc. Var. Partial Differential Equations*, doi: 10.1007/s00526-013-0666-9
[29] T. KUUSI, G. MINGIONE: Guide to nonlinear potential estimates, *Bulletin of Mathematical Sciences* **4 (1)**: 1–82, 2014.
[30] G. M. LIEBERMAN: Boundary regularity for solutions of degenerate elliptic equations, *Nonlinear Anal.* **12**: 1203–1219, 1988.
[31] G. M. LIEBERMAN: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations, *Comm. Partial Differential Equations* **16**: 311–361, 1991.
[32] G. M. LIEBERMAN: Sharp forms of estimates for subsolutions and supersolutions of quasilinear elliptic equations involving measures, *Comm. Partial Differential Equations* **18**: 1191–1212, 1993.
[33] J. MALÝ: Wolff potential estimates of superminimizers of Orlicz type Dirichlet integrals, *Manuscripta Math.* **110 (4)**: 513– 525, 2003.
[34] G. MINGIONE: The Calderón-Zygmund theory for elliptic problems with measure data, *Ann. Scuola Norm. Sup. Pisa Cl. Sci (5)* **6 (2)**: 195–261, 2007.
[35] G. MINGIONE: Gradient estimates below the duality exponent, *Math. Ann.* **346 (3)**: 571–627, 2010.
[36] G. MINGIONE: Gradient potential estimates, *J. Eur. Math. Soc.* (*JEMS*) **13**: 459–486, 2011.
[37] M.M. RAO, Z.D. REN: *Theory of Orlicz spaces*, Marcel Dekker Inc., 1991.
[38] C. SBORDONE: On some integral inequalities and their applications to the calculus of variations, *Boll. Un. Mat. Ital., Analisi Funzionale e Applicazioni, Ser. VI* **5**: 73–94, 1986.
[39] C. SCHEVEN, T. SCHMIDT: Asymptotically regular problems II: Partial Lipschitz continuity and a singular set of positive measure *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)* **8 (3)**: 469–507, 2009.
[40] E. M. STEIN: Editor's note: the differentiability of functions in $\mathbb{R}^n$, *Ann. of Math.* **113 (2)**: 383–385, 1981.
[41] G. TALENTI: An embedding theorem, *Essays of Math. Analysis in honor of E. De Giorgi,* Birkhäuser Verlag, Boston, 1989.
[42] N. S. TRUDINGER: On imbeddings into Orlicz spaces and some applications, *J. Math. Mech.* **17**: 473–483, 1967.

PAOLO BARONI, DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, LÄGERHYDDSVÄGEN 1, SE-751 06, UPPSALA, SWEDEN.
E-mail address: paolo.baroni@math.uu.se