ON CERTAIN ANISOTROPIC ELLIPTIC EQUATIONS ARISING IN CONGESTED OPTIMAL TRANSPORT: LOCAL GRADIENT BOUNDS

LORENZO BRASCO AND GUILLAUME CARLIER

Abstract. Motivated by applications to congested optimal transport problems, we prove higher integrability results for the gradient of solutions to some anisotropic elliptic equations, exhibiting a wide range of degeneracy. The model case we have in mind is the following:

$$\partial_x \left[(|u_x|-\delta_1)_+^{q-1}\,\frac{u_x}{|u_x|}\right]+\partial_y \left[(|u_y|-\delta_2)_+^{q-1}\,\frac{u_y}{|u_y|}\right]=f,$$

for $2\le q<\infty$ and some non negative parameters δ_1,δ_2. Here $(\,\cdot\,)_+$ stands for the positive part. We prove that if $f\in L^\infty_{loc}$, then $\nabla u\in L^r_{loc}$ for every $r\ge 1$.

Contents

1. Introduction

1.1. **Background and motivations.** Let $\Omega\subset\mathbb{R}^N$ be an open bounded (smooth) set and let us consider a variational integral of the type

$$\mathcal{F}(u)=\int_\Omega F(\nabla u(x))\,dx+\int_\Omega f(x)\,u(x)\,dx, \tag{1.1}$$

with $z \mapsto F(z)$ being a convex energy with $q-$growth at infinity (here $q > 1$), uniformly convex for $|z| \gg 1$, but not necessarily strictly convex in the whole. The prototypical case of such an energy is given by

$$F(z) = \frac{1}{q}\,(|z| - \delta)_+^q, \qquad z \in \mathbb{R}^N, \tag{1.2}$$

for some $\delta \geq 0$, which identically vanishes on the ball $\{z\,:\,|z| \leq \delta\}$. More general functionals of this type have been considered for example in [10, 13, 14].

As pointed out in [7], regularity results for minimizers of such a kind of functionals are tightly connected with *optimal transport problems with congestion effects*. In order to neatly motivate the purposes of this paper, we want to spend some words about this point. Suppose that the positive f^+ and negative parts f^- of f stand for the densities of centers of production and consumption of a given commodity in the region $\Omega \subset \mathbb{R}^N$ (the physical case clearly corresponds to $N = 2$). The transportation programs are represented by vector fields ϕ satisfying the balance laws

$$\operatorname{div}\phi = f^+ - f^- \qquad + \qquad \text{Neumann boundary conditions},$$

where these boundary conditions are zero if $\int_\Omega f^+ = \int_\Omega f^-$, i.e. if the region is economically balanced, so there is no need for import/export activities. Observe that the constraint on the divergence simply states that in each point the incoming/outcoming transportation flow is ruled by the difference between the demand and the supply. Then a function $G : \mathbb{R}^N \to \mathbb{R}^+$ is given, such that for every transportation program ϕ, the quantity

$$\mathcal{G}(\phi) = \int_\Omega G(\phi(x))\,dx,$$

gives the total transportation cost. In order to capture the effects of congestion, the function G is typically taken to be strictly convex and superlinear. In economical terms, this comes from the fact that in a congested situation, the marginal cost ∇G is strictly monotone and divergent at infinity. In other words, the transport problem we are facing is the so called *Beckmann's problem*, introduced in [2] and defined by

$$\min\left\{\mathcal{G}(\phi)\,:\,\operatorname{div}\phi = f,\ \text{in } \Omega,\ \langle \phi, \nu_\Omega\rangle = 0 \text{ on } \partial\Omega\right\}, \tag{1.3}$$

where for simplicity we are considering the balanced case, i.e. we assume $\int_\Omega f\,dx = 0$. The link between (1.3) and our original problem (1.1) is given by

$$\min_\phi \mathcal{G}(\phi) = \max_u -\mathcal{F}(u),$$

provided that F is the Legendre-Fenchel transform of G. Then optimizers of the two problems are linked by the primal-dual optimality conditions

$$\nabla u_0 \in \partial G(\phi_0) \qquad \text{or equivalently} \qquad \phi_0 \in \partial F(u_0)$$

i.e. (1.1) is the dual (in the sense of convex analysis) of Beckmann's problem. As always in Optimal Transport, the dual variables u of problem (1.1) have to be thought as *price systems* for a company, handling the transport in a congested situation. An optimizers u_0 then gives the price system which maximizes the profit of the company.

Observe that the function (1.2) considered in [6, 7] corresponds to Beckmann's problem with cost

$$G(\xi) = \frac{1}{p}\,|\xi|^p + \delta\,|\xi|, \qquad \xi \in \mathbb{R}^N,$$

where $p = q/(q-1)$. In this case, a Lipschitz estimate and the higher differentiability of an optimal price u_0 can be proved, by looking at the corresponding Euler-Lagrange equation

$$\operatorname{div}\left((|\nabla u_0| - \delta)_+^{q-1} \frac{\nabla u_0}{|\nabla u_0|} \right) = f. \tag{1.4}$$

Appealing to the primal-duality optimality conditions, these in turn give that the optimal ϕ_0 in (1.3) is a bounded Sobolev vector field, provided f is smooth enough (see [6, Theorem 2.1] and [7, Theorem 3.4]).

It is crucial to observe that for such a cost G, the linear part $|\xi|$ prevails on the strictly convex one as $|\xi| \ll 1$. This means that *congestion effects are negligible, in the small mass regime*, an hypothesis which is very natural. Of course, the lack of strict convexity for the Lagrangian (1.2) is precisely a consequence of this assumption. This implies that *elliptic equations that are relevant for these transport problems, typically exhibit a severe degeneracy*, like in (1.4).

1.2. **More degeneracy: technical issues.** However, the hypothesis of an isotropic cost function G is not always well-motivated. For example, as shown in [1], the analysis of discrete congested transport problems settled on a network grid of small size ε, naturally leads to (1.3) with a transportation cost of the form

$$G(\xi) = \sum_{i=1}^{N} h_i(|\xi_i|), \qquad \xi \in \mathbb{R}^N,$$

as the parameter ε goes to 0. Roughly speaking, this anisotropic cost keeps memory of the rigid geometry (i.e. the network grid) of the approximating discrete problems.

Here again, the functions $h_1, \dots, h_N : \mathbb{R}^+ \to \mathbb{R}^+$ are strictly convex and superlinear, such that $h_i(0) = 0$ and $h_i'(0) = \delta_i > 0$. As before, this last hypothesis is motivated by the reasonable assumption that G behaves linearly, for small masses. Back to our original problem (1.1), it is then natural to ask which kind of Lagrangians we are lead to study, with such a choice. It is easily seen that in this case, we have

$$F(z) = \sum_{i=1}^{N} h_i^*(|z_i|), \qquad z \in \mathbb{R}^N,$$

where h_i^* are C^1 functions of one variable, constantly equal to 0 on the interval $[0, \delta_i]$, due to the assumption $h_i'(0) = \delta_i > 0$. A significant instance of such a Lagrangian is given by

$$F(z) = \sum_{i=1}^{N} \frac{(|z_i| - \delta_i)_+^q}{q}, \qquad z \in \mathbb{R}^N. \tag{1.5}$$

This function considerably differs from (1.2), in that this time the Hessian matrix D^2F is given by a diagonal matrix, whose $i-$th entry $(h_i^*)''(|z_i|)$ is constantly zero as $0 \leq |z_i| \leq \delta_i$. In terms of the corresponding Euler-Lagrange equation

$$\operatorname{div} \nabla F(\nabla u) = f, \tag{1.6}$$

this implies that *ellipticity breaks down at every point where a single component of the gradient is small.* Also observe that due to the particular structure of D^2F, the function F not only lacks strict convexity, but it is not even uniformly convex "at infinity", i.e. outside a ball, contrary to the case of (1.2). This fact is the main source of difficulties. Typically, in order to derive higher integrability results for the gradient of minimizers of $\mathcal{F}$, one considers the differentiated equations

$$\mathrm{div}\left(D^2F(\nabla u)\nabla u_{x_j}\right) = f_{x_j}, \qquad j=1,\dots,N, \tag{1.7}$$

which are solved by the (components of the) gradient of a minimizer u_0. Then it is sufficient to know that this equation is uniformly elliptic at least "at infinity", for example[1]

- there exists $M_0 \geq 0$, such that $\quad (1+|z|^2)^{\frac{q-2}{2}} \lesssim \min\limits_{|z|\geq M_0}\left(\min\limits_{|\vartheta|=1}\langle D^2F(z)\,\vartheta,\vartheta\rangle\right), \quad$ for every $z\in\mathbb{R}^N$;
- $|D^2F(z)| \lesssim (1+|z|^2)^{\frac{q-2}{2}}, \quad$ for every $z\in\mathbb{R}^N$;

to conclude that ∇u_0 is in L^∞. In this case the natural idea, which is somehow common to [6, 10, 13, 14], would be that of cutting away the degeneracy region, by localizing equation (1.7) "in a neighborhood of infinity". In a nutshell, this is done by selecting suitable test functions in the weak formulation of (1.7), for examples quantities like $(|\nabla u_0|^\beta - M_0^\beta)_+$ would do the job. Thanks to the hypotheses on D^2F, one can derive Caccioppoli inequalities for these quantities, which combined with the Sobolev inequality give a recursive scheme of reverse Hölder inequalities, leading to the L^∞ estimate for ∇u_0.

Here on the contrary, with the choice (1.5) we have

$$\min_{|z|\geq M}\left(\min_{|\xi|=1}\langle D^2F(z)\,\xi,\xi\rangle\right) = 0, \qquad \text{for every } M>0,$$

so that we have an obstruction in deriving a true Caccioppoli inequality for positive subsolutions of the linearized equation (1.7). This is the main technical difficulty linked with this type of degeneracy. However, something can be done. First of all, a weak form of the Caccioppoli inequality can still be derived (see Lemma 2.7 below). This time, rather than having an integral control on ∇u_{x_j} in terms of u_{x_j} itself, we have a control on a "weighted" norm of ∇u_{x_j}, where the weights depend on *all the other components u_{x_i} of the gradient*, through the nonlinear functions $(h_i^*)''$. Once we have this, we avoid the use of Sobolev inequality for these weighted integrals. Rather, we apply a sort of *very weak weighted Gagliardo-Nirenberg inequality* (see Lemma 2.6 below), valid for solutions of equation (1.6). This can be derived by means of a weird choice for the test function to be inserted in the weak formulation of (1.6). This is given by a mixture of the solution and its gradient. We point out that a similar idea can be found in the papers [4, 5] (dealing with variational integrals similar to those considered here, in less degenerate situations) and[2] [16]. We learnt this trick from Di Benedetto's celebrated paper on $C^{1,\alpha}$ regularity for solutions of $p-$Laplacian type equations (see [12, Lemma 2.4 and Proposition 3.1]).

Once we have these two surrogates of the classical tools, we can join them to get the desired recursive scheme of reverse Hölder inequalities (Lemma 2.10). Here there is a drawback, since at each step the gain of integrability for the gradient is quite poor and does not permit to arrive at an L^∞ estimate. However, boundedness on each Lebesgue space with finite exponent can be obtained and this is the main achievement of this paper.

[1]As shown in [14], in general the upper bound on D^2F is not really necessary and can be replaced by a growth assumption on ∇F. Here for ease of exposition, we stick to a more classical hypothesis.

[2]In this paper, the expression "rather unorthodox choice of the test function" is used to describe a similar procedure (see [16, Lemma 6.2])

Last but not least, some final words concerning the hypothesis on the datum f, which is required to be in L^∞. One may wonder whether this hypothesis is optimal or not, since usually $f \in L^{N+\varepsilon}$ is the sharp assumption (on the scale of Lebesgue spaces) that could guarantee a Lipschitz estimate for the solution. The key point is that the usual proof of this result uses the Sobolev inequality, to treat the term f as a lower order perturbation. For the reason described before, such a strategy seems not to work in the present situation. We leave as an (interesting) open question to know if the L^∞ hypothesis on f can be weakened.

1.3. **The result of this paper.** Motivated by the previous discussion, the main aim of this paper is to investigate regularity properties of local minimizers of the following convex energy

$$\mathfrak{F}_q(u;\Omega) = \sum_{i=1}^N \int_\Omega g_i(u_{x_i})\,dx + \int_\Omega f(x)\,u(x)\,dx, \qquad u \in W^{1,q}(\Omega),$$

where $q \geq 2$, the functions g_i are defined by

$$g_i(t) = \frac{1}{q}\,(|t| - \delta_i)_+^q, \qquad t \in \mathbb{R},$$

for some $\delta_1, \ldots, \delta_N \geq 0$ and $f \in L^\infty(\Omega)$ is given. For the reader's convenience, we recall that $u \in W^{1,q}_{loc}(\Omega)$ is said to be a *local minimizer* of $\mathfrak{F}_q$ if for every $\Omega' \Subset \Omega$ we have

$$\mathfrak{F}_q(u;\Omega') \leq \mathfrak{F}_q(u+\varphi;\Omega'), \qquad \text{for every } \ \varphi \ \text{ such that } \ v \in W^{1,q}_0(\Omega').$$

In this paper, we are going to prove the following higher integrability result.

Main Theorem. *Let $f \in L^\infty_{loc}(\Omega)$ and $q \geq 2$. If $u \in W^{1,q}_{loc}(\Omega)$ is a local minimizer of $\mathfrak{F}_q$, then $u \in W^{1,r}_{loc}(\Omega)$ for every $r \geq 1$.*

In the particular case $\delta_1 = \cdots = \delta_N = 0$, the gradient term of $\mathfrak{F}_q$ coincides with the anisotropic Dirichlet energy, i.e.

$$\frac{1}{q}\int_\Omega \|\nabla u(x)\|_{\ell^q}^q\,dx + \int_\Omega f(x)\,u(x)\,dx, \qquad u \in W^{1,q}(\Omega),$$

where for every $z \in \mathbb{R}^N$, we set $\|z\|_{\ell^q} = \left(\sum_{i=1}^N |z_i|^q\right)^{1/q}$. Observe that a local minimizer of this anisotropic energy is a local weak solution of the equation

$$\sum_{i=1}^N \left(|u_{x_i}|^{q-2}\,u_{x_i}\right)_{x_i} = f, \tag{1.8}$$

where the differential operator on the left-hand side is sometimes called *pseudo $q-$Laplacian*. Equations like (1.8) have been around for a long time. One of the first paper to address regularity issues for them has been [17]. Recently, they started to attract an increasing interest, due to their applications in biology and physics (see for example [18] and the references therein). We also cite the papers [3, 8], where some spectral properties of this operator are investigated.

1.4. **Plan of the paper.** The rest of the paper is devoted to prove the Main Theorem. In Section 2 we will derive local uniform estimates for the gradients of minimizers of some regularized problems. Then in Section 3 we will show how to take these estimates to the limit, in order to prove the desired result. Since the functional $\mathfrak{F}_q$ is not strictly convex, a further penalization argument will be needed, in order to be sure to select the desired local minimizer in the limit. Finally, the concluding Section 4 gives an application of the Main Theorem to the relevant optimal transport problem.

2. Regularity estimates for approximating problems

Let us fix an open set $\mathcal{O} \subset \mathbb{R}^N$. For every $\varepsilon \ll 1$, we consider the following functional

$$\mathfrak{F}_q^\varepsilon(u) = \sum_{i=1}^N \int_{\mathcal{O}} g_i^\varepsilon(u_{x_i}(x))\,dx + \varepsilon \int_{\mathcal{O}} H(\nabla u(x))\,dx + \int_{\mathcal{O}} b(x, u(x))\,dx, \qquad u \in W^{1,q}(\mathcal{O}),$$

where:

- for every $i = 1, \dots, N$, we simply set $g_i^\varepsilon(t) = g_i(t)$ if $q > 2$, while if $q = 2$ this is given by
$$g_i^\varepsilon(t) = \begin{cases} 0, & \text{if } |t| \le \delta_i - \varepsilon, \\ \dfrac{1}{12\,\varepsilon}\,(|t| - \delta_i + \varepsilon)^3, & \text{if } \delta_i - \varepsilon \le |t| \le \delta_i + \varepsilon, \\ \dfrac{1}{6}\varepsilon^2 + \dfrac{1}{2}\,(|t| - \delta_i)^2, & \text{if } |t| \ge \delta_i + \varepsilon, \end{cases}$$
which converges in C^1 to $1/2\,(|t| - \delta_i)_+^2$ as ε goes to 0;
- $H : \mathbb{R}^N \to \mathbb{R}$ is the C^∞ strictly convex function given by
$$H(z) = \frac{1}{q}\,(1 + |z|^2)^{\frac{q}{2}}, \qquad z \in \mathbb{R}^N;$$
- $b : \mathcal{O} \times \mathbb{R} \to \mathbb{R}$ is of class C^∞ and such that

$$|b(x,u)| \le C_1\,(|u| + 1) \qquad \text{and} \qquad \left| b'(x,u) := \frac{\partial}{\partial u} b(x,u) \right| \le C_2, \qquad (x,u) \in \mathcal{O} \times \mathbb{R}. \tag{2.1}$$

In this section, we will prove the following result.

Proposition 2.1. *Let $q \ge 2$ and $\zeta \in W^{1,q}(\mathcal{O})$. If $u^\varepsilon \in W^{1,q}(\mathcal{O})$ is a solution of*

$$\min \left\{ \mathfrak{F}_q^\varepsilon(v) \,:\, v - \zeta \in W_0^{1,q}(\mathcal{O}) \right\}, \tag{2.2}$$

then $u^\varepsilon \in W^{1,r}_{loc}(\mathcal{O})$, for every $r \ge 1$. Moreover, for every $\Sigma \Subset \mathcal{O}$ and every $r \ge 1$, we have the following estimate

$$\|u^\varepsilon\|_{W^{1,r}(\Sigma)} \le C \tag{2.3}$$

for some constant C depending on q, r, N, $\|u^\varepsilon\|_{W^{1,q}}$, Σ and the constants C_1, C_2 in (2.1), but not on ε.

The rest of this section is devoted to prove Proposition 2.1. For the sake of readability, we divide the proof in five main steps, each corresponding to a subsection.

2.1. **Step 1: machinery and preliminary results.** In what follows, we will drop the superscript ε and we will simply use u to denote the solution of (2.2). The same convention will be used for the functions g_i. Observe that the Euler-Lagrange equation of problem (2.2) is given by

$$\sum_{i=1}^{N} \int_{\mathcal{O}} g_i'(u_{x_i})\, \varphi_{x_i}\, dx + \varepsilon \int_{\mathcal{O}} \langle \nabla H(\nabla u), \nabla \varphi \rangle\, dx + \int_{\mathcal{O}} b'(x,u)\, \varphi\, dx = 0, \quad \text{for every } \varphi \in W_0^{1,q}(\mathcal{O}). \tag{2.4}$$

Let us first collect some basic properties of the convex functions g_i. The proof being elementary, it is left to the reader.

Lemma 2.2. *For every $i = 1, \dots, N$ and every $q \geq 2$, the function g_i is $C^{2,\alpha}$, with $\alpha = \min\{q-2, 1\}$ for $q > 2$ and $\alpha = 1$ for $q = 2$ (regularized case). Moreover, we have the following estimates*

$$g_i''(t) \leq (q-1)\, t^{q-2} \qquad \text{and} \qquad \frac{g_i''(t)\, t^2}{q-1} \geq 2^{-q}\, t^q - C,\ (C = C(\delta, q)), \qquad \text{for every } t \in \mathbb{R}, \tag{2.5}$$

and also

$$g_i'(t)\, t \geq \frac{1}{2(q-1)}\, g_i''(t)\, t^2 - \frac{\delta_i^2}{2\,(q-1)}\, g_i''(t) \qquad \text{and} \qquad \frac{|g_i'(t)|}{|t|} \leq \frac{g_i''(t)}{q-1}, \qquad \text{for every } t \in \mathbb{R}. \tag{2.6}$$

Remark 2.3. We observe that the integrand of $\mathfrak{F}_q^\varepsilon$ is a $C^{2,\alpha}$ function, whose Hessian with respect to the gradient variable is bounded from below and such that the ratio between its minimal and maximal eigenvalue is bounded. Then we can infer the $C^{2,\alpha}$ local regularity for the solutions u^ε (see [15, Theorem 10.18]). This implies that quantities of the type $h(u_{x_i})$ are admissible test functions, for every $h : \mathbb{R} \to \mathbb{R}$ locally Lipschitz.

First of all, we need the following classical L^∞ result, for local minimizers of integral having $q-$growth conditions in the gradient variable. The important point is the dependence of the constant of the L^∞ estimate. For a proof of this standard result, the reader can consult [15, Theorem 7.5]. The statement has been adapted to suit our simplified hypotheses.

Lemma 2.4. *Let $F : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N$ be a Caratheodory function satisfying the growth conditions*

$$|z|^q - M\,(|u| + 1) \leq F(x,u,z) \leq L\,|z|^q + M\,(|u|+1), \qquad (x,u,z) \in \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N, \tag{2.7}$$

for some positive constants L and M. Then every local minimizer $u \in W^{1,q}(\mathcal{O})$ of the functional

$$\int F(x, u(x), \nabla u(x))\, dx,$$

belongs to $L^\infty_{loc}(\mathcal{O})$. Moreover, there exists a constant C depending on $q, N, \|u\|_{W^{1,q}}$ and M, such that for every pair of concentric balls $B_\varrho(x_0) \subset B_R(x_0) \Subset \mathcal{O}$, we have

$$\|u\|_{L^\infty(B_\varrho(x_0))} \leq C \left[\frac{1}{(R-\varrho)^{\frac{N}{q}}}\, \|u\|_{L^q(B_R(x_0))} + M^{\frac{1}{q}}\, R \right].$$

Particularizing the previous result to our problem (2.2), we have the following.

Corollary 2.5. *Let $u^\varepsilon \in W^{1,q}(\mathcal{O})$ be a solution of problem (2.2). Then for every $\Sigma \Subset \mathcal{O}$ we have*

$$\|u^\varepsilon\|_{L^\infty(\Sigma)} \leq C,$$

for some constant C depending on $q, N, \|u^\varepsilon\|_{W^{1,q}}, \Sigma, \max\{\delta_1,\dots,\delta_N\}$ and the constant C_1 in (2.1), but not ε.

Proof. It is sufficient to check that $\mathfrak{F}^\varepsilon_q$ verifies hypothesis (2.7), then we can apply the estimate of Lemma 2.4. To this aim, we simply use the first hypothesis (2.1) on b, the definition of H and the estimates of Lemma 2.2 for the functions g_i. $\square$

We now fix some notation that we will use throughout the rest of the paper:

$$w(x) = 1 + |\nabla u(x)|^2, \qquad k_j = \delta_j + 1, \qquad \text{and} \qquad v_j = (u_{x_j} - k_j)_+^2 + 1, \quad j = 1,\dots,N. \tag{2.8}$$

2.2. **Step 2: a Sobolev-type inequality.** As already remarked in the Introduction, we need a sort of Sobolev inequality for solutions of (2.4). In this sense, the most important term in the right-hand side of (2.9) below is the gradient term. It is not difficult to see that the sum of the powers of the right-hand side is smaller than that on the left-hand one. Heuristically, this means that we are facing a a Gagliardo-Nirenberg inequality. However, things are more complicated, since the partial derivatives u_{x_j} and u_{x_i} are mixed.

Lemma 2.6. *Let α, β be two positive exponents such that*

$$0 \le \alpha < \beta,$$

then using the notation introduced in (2.8), for every $\xi \in C^1_0(\mathcal{O})$ we have

$$\begin{aligned}\sum_{i=1}^N \int g_i''(u_{x_i})\, |u_{x_i}|^2\, v_j^\beta\, \xi^2 + \varepsilon \int w^{\frac{q}{2}}\, v_j^\beta\, \xi^2 &\le C \sum_{i=1}^N \int_{\{u_{x_j} > k_j\}} |g_i''(u_{x_i})| \left|\partial_{x_i}\left(v_j^{\beta - \frac{\alpha}{2}}\right)\right|^2 \xi^2 \\ &\quad + C \int w^{\frac{q}{2}}\, v_j^\alpha\, \xi^2 + C \int w^{\frac{q-2}{2}}\, v_j^\beta\, \left(|\nabla \xi|^2 + \xi^2\right) \\ &\quad + \varepsilon\, C \int_{\{u_{x_j} > k_j\}} w^{\frac{q-2}{2}} \left|\nabla\left(v_j^{\beta - \frac{\alpha}{2}}\right)\right|^2 \xi^2, \quad j = 1,\dots,N,\end{aligned} \tag{2.9}$$

for some constant C depending on q, N and $\|u\|_{W^{1,q}}$, but neither on ε, nor on α and β.

Proof. We take the following test function

$$\varphi^+_{j,\beta} = u\, v_j^\beta\, \xi^2, \qquad j = 1,\dots,N, \quad \beta > 0,$$

which is admissible thanks to Remark 2.3. Inserting it into (2.4), we get

$$\begin{aligned}\sum_{i=1}^N \int_\Omega g_i'(u_{x_i})\, u_{x_i}\, v_j^\beta\, \xi^2 + \sum_{i=1}^N \int g_i'(u_{x_i})\ \partial_{x_i}\left(v_j^\beta\right) u\, \xi^2 &+ 2 \sum_{i=1}^N \int_\Omega g_i'(u_{x_i})\, \xi_{x_i}\, \xi\, u\, v_j^\beta \\ &+ \varepsilon \int_\Omega \langle \nabla H(\nabla u), \nabla u\rangle\, v_j^\beta\, \xi^2 \\ &+ \varepsilon \int \langle \nabla H(\nabla u), \nabla(v_j^\beta)\rangle\, \xi^2\, u \\ &+ 2\,\varepsilon \int_\Omega \langle \nabla H(\nabla u), \nabla \xi\rangle\, \xi\, u\, v_j^\beta \\ &= -\int b'\, u\, v_j^\beta\, \xi^2, \qquad j = 1,\dots,N.\end{aligned} \tag{2.10}$$

We start estimating the second term in (2.10): observe that using Young inequality we have

$$\begin{aligned}\left|\partial_{x_i}\left(v_j^\beta\right)\right| = \beta\, v_j^{\beta-1}\, |\partial_{x_i} v_j| &\le \frac{1}{2}\beta^2\, v_j^{2\beta-\alpha-2}\, |\partial_{x_i} v_j|^2 + \frac{1}{2}\, v_j^\alpha\, 1_{\{u_{x_j}>k_j\}}\\ &= \frac{2\,\beta^2}{(2\,\beta-\alpha)^2}\left|\partial_{x_i}\left(v_j^{\beta-\frac{\alpha}{2}}\right)\right|^2 + \frac{1}{2}\, v_j^\alpha\, 1_{\{u_{x_j}>k_j\}},\end{aligned}$$

where we used that $\partial_{x_i} v_j = 0$ on the set $\{u_{x_j} \le k_j\}$. Also observe that thanks to the fact that we are assuming $\alpha < \beta$, in the previous we can further estimate

$$\frac{2\,\beta^2}{(2\,\beta-\alpha)^2} \le 2,$$

so that the constant C that will appear in (2.9) will not depend on α and β. Using the previous estimate and the fact that $u \in L^\infty$, the second term can be estimated by

$$\begin{aligned}\left|\sum_{i=1}^N \int g_i'(u_{x_i})\ \partial_{x_i}\,(v_j^\beta)\, u\,\xi^2\right| &\le C\sum_{i=1}^N \int \frac{|g_i'(u_{x_i})|}{|u_{x_i}|}\left|\partial_{x_i}\left(v_j^{\beta-\frac{\alpha}{2}}\right)\right|^2 \xi^2\\ &\quad + C\sum_{i=1}^N \int_{\{u_{x_j}>k_j\}} |g_i'(u_{x_i})|\,|u_{x_i}|\, v_j^\alpha\,\xi^2,\end{aligned} \tag{2.11}$$

for some constant $C > 0$, clearly depending on $\|u\|_{L^\infty}$. Observe that the second integral can be easily estimated by

$$\sum_{i=1}^N \int_{\{u_{x_j}>k_j\}} |g_i'(u_{x_i})|\,|u_{x_i}|\, v_j^\alpha\,\xi^2 \le C\int_{\{u_{x_j}>k_j\}} w^{\frac{q}{2}}\, v_j^\alpha\,\xi^2,$$

using the growth of g_i and the definition of w. The third term in (2.10) is estimated by

$$\begin{aligned}\left|\sum_{i=1}^N \int g_i'(u_{x_i})\,\xi_{x_i}\,\xi u\, v_j^\beta\right| &\le C\,\tau\sum_{i=1}^N \int |g_i'(u_{x_i})|\,|u_{x_i}|\,\xi^2\, v_j^\beta\\ &\quad + \frac{C}{\tau}\sum_{i=1}^N \int \frac{|g_i'(u_{x_i})|}{|u_{x_i}|}\,|\nabla\xi|^2\, v_j^\beta\\ &\le \frac{C}{\tau}\int w^{\frac{q-2}{2}+\beta}\,|\nabla\xi|^2 + C\,\tau\sum_{i=1}^N \int |g_i'(u_{x_i})|\,|u_{x_i}|\,\xi^2\, v_j^\beta,\end{aligned}$$

and the second term can be absorbed in the left-hand side, by taking $\tau > 0$ small enough and observing that $g_i'(t)\,t \ge 0$.

We now come to the estimates of the $\varepsilon-$terms: for the fourth term in (2.10), we first observe that

$$\int \langle \nabla H(\nabla u), \nabla u\rangle\, v_j^\beta\,\xi^2 = \int w^{\frac{q}{2}}\, v_j^\beta\,\xi^2 - \int w^{\frac{q-2}{2}}\, v_j^\beta\,\xi^2,$$

so that we will collect the first integral in left-hand side and put the second one in the right-hand side, since this is a lower-order term with respect to the first (just check the sum of the powers). As for the second

$\varepsilon-$term, we use:

$$\begin{aligned}|\langle \nabla H(\nabla u), \nabla(v_j^\beta)\rangle| &\le C\,|\nabla v_j|\,v_j^{\beta-1}\,|\nabla u|\,(1+|\nabla u|^2)^{\frac{q-2}{2}}\\ &\le C\,|\nabla v_j|\,v_j^{\beta-\frac{\alpha}{2}-1}\,v_j^{\frac{\alpha}{2}}\,w^{\frac{q-1}{2}}\\ &\le C\,\left|\nabla\left(v_j^{\beta-\frac{\alpha}{2}}\right)\right|^2\,w^{\frac{q-2}{2}}+C\,w^{\frac{q}{2}}\,v_j^\alpha,\end{aligned}$$

and thus obtain

$$\left|\int \langle \nabla H(\nabla u), \nabla(v_j^\beta)\rangle\,\xi^2\,u\right| \le C\int_{\{u_{x_j}>k_j\}} w^{\frac{q-2}{2}}\,\left|\nabla\left(v_j^{\beta-\frac{\alpha}{2}}\right)\right|^2\,\xi^2+C\int_{\{u_{x_j}>k_j\}} w^{\frac{q}{2}}\,v_j^\alpha\,\xi^2,$$

and finally, still in the same way as before we get

$$\left|\int_\Omega \langle \nabla H(\nabla u), \nabla \xi\rangle\,\xi\,u\,v_j^\beta\right| \le C\,\tau\int_\Omega w^{\frac{q}{2}}\,\xi^2\,v_j^\beta+\frac{C}{\tau}\int_\Omega w^{\frac{q-2}{2}}\,|\nabla\xi|^2\,v_j^\beta,$$

so that the first integral can be once again absorbed in the left-hand side.

Finally, we estimate the term containing b': since this is bounded, it is readily seen that we have

$$\left|\int b'\,u\,v_j^\beta\,\xi^2\right| \le C\int v_j^\beta\,\xi^2 \le C\int w^{\frac{q-2}{2}}\,v_j^\beta\,\xi^2.$$

Collecting all the estimates that we derived and using once again (2.6), we arrive at (2.9). □

2.3. **Step 3: a Caccioppoli-type inequality.** In order to derive a Caccioppoli-type inequality for the gradient, we have to differentiate the equation (2.4) with respect to x_j. More precisely, let us take the test function $\varphi=\eta_{x_j}$ in (2.4), for some $\eta\in C_0^\infty(\mathcal{O})$. By recalling that $u\in C^{2,\alpha}$ and integrating by parts, the resulting equation takes the form

$$-\sum_{i=1}^N\int g_i''(u_{x_i})\,u_{x_i\,x_j}\,\eta_{x_i}\,dx-\varepsilon\int\langle D^2H(\nabla u)\,D_j^2u,\nabla\eta\rangle\,dx+\int b'(x,u)\,\eta_{x_j}\,dx=0, \tag{2.12}$$

for every $\eta\in C_0^\infty(\mathcal{O})$. Here D_j^2u stands for the $j-$th column of the Hessian matrix, i.e.

$$D_j^2u=\begin{bmatrix} u_{x_1\,x_j}\\ \vdots\\ u_{x_N\,x_j}.\end{bmatrix},\qquad j=1,\dots,N.$$

By a density argument, we then get that (2.12) holds for every $\eta\in W_0^{1,q}(\mathcal{O})$.

Lemma 2.7. *Using the notation* (2.8), *for every* $s>0$ *and every* $\xi\in C^1_0(\mathcal{O})$, *we have*

$$
\begin{aligned}
s\sum_{i=1}^N \int g_i''(u_{x_i})\left|\partial_{x_i}\left(v_j^{\frac{s+1}{2}}\right)\right|^2\xi^2 &+ \sum_{i=1}^N\int_{\{u_{x_j}>k_j\}} g_i''(u_{x_i})\,u_{x_i\,x_j}^2\,v_j^s\,\xi^2\\
&+\varepsilon\, s\int w^{\frac{q-2}{2}}\left|\nabla\left(v_j^{\frac{s+1}{2}}\right)\right|^2\xi^2\\
&+\varepsilon\int_{\{u_{x_j}>k_j\}}\langle D^2H(\nabla u)\,D_j^2u, D_j^2u\rangle\, v_j^s\,\xi^2\\
&\le C\,(s+1)\int_{\{u_{x_j}>k_j\}} w^{\frac{q-2}{2}}\,v_j^{s+1}\,(|\nabla\xi|^2+\xi^2),\qquad j=1,\dots,N,
\end{aligned}
\tag{2.13}
$$

for a constant C *depending on* q, N *and* $\|b'\|_\infty$, *but not on* ε.

Proof. We insert the test function

$$
\psi^+_{j,s}=(u_{x_j}-k_j)_+\,v_j^s\,\xi^2,
$$

in equation (2.12), with $s>0$. Then we obtain the following 3 groups of terms that have to be estimated: the terms containing the functions g_i

$$
\begin{aligned}
\sum_{i=1}^N\int_{\{u_{x_j}>k_j\}} g_i''(u_{x_i})\,u_{x_i\,x_j}^2\,v_j^s\,\xi^2 &+ 2\,s\sum_{i=1}^N\int g_i''(u_{x_i})\,u_{x_i\,x_j}^2\,v_j^{s-1}\,(u_{x_j}-k_j)_+^2\,\xi^2\\
&+2\sum_{i=1}^N\int g_i''(u_{x_i})\,u_{x_i\,x_j}\,(u_{x_j}-k_j)_+\,v_j^s\,\xi_{x_i}\,\xi\\
&=: G_1+2\,s\,G_2+2\,G_3,
\end{aligned}
$$

the terms containing H

$$
\begin{aligned}
\int_{\{u_{x_j}>k_j\}}\langle D^2H(\nabla u)\,D_j^2u, D_j^2u\rangle\,v_j^s\,\xi^2 &+ 2\,s\int\langle D^2H(\nabla u)\,D_j^2u, D_j^2u\rangle\,v_j^{s-1}\,(u_{x_j}-k_j)_+^2\,\xi^2\,dx\\
&+2\int\langle D^2H(\nabla u)\,D_j^2u,\nabla\xi\rangle\,\xi\,(u_{x_j}-k_j)_+\,v_j^s\,dx\\
&=: H_1+2\,s\,H_2+2\,H_3,
\end{aligned}
$$

and the terms containing b', i.e.

$$
\begin{aligned}
-\int_{\{u_{x_j}>k_j\}} b'\,u_{x_j\,x_j}\,v_j^s\,\xi^2dx &+ 2\,s\int b'\,(u_{x_j}-k_j)_+^2\,v_j^{s-1}\,u_{x_j\,x_j}\,\xi^2\\
&-2\int b'\,(u_{x_j}-k_j)_+\,v_j^s\,\xi_{x_j}\,\xi =: B_1+2\,s\,B_2+2\,B_3.
\end{aligned}
$$

Terms G_i. Let us start with the term G_2: by noticing that

$$
u_{x_i\,x_j}^2\,v_j^{s-1}\,(u_{x_j}-k_j)_+^2=\left|v_j^{\frac{s-1}{2}}\,u_{x_i\,x_j}\,(u_{x_j}-k_j)_+\right|^2=\frac{1}{(s+1)^2}\left|\partial_{x_i}\left(v_j^{\frac{s+1}{2}}\right)\right|^2,
$$

we get

$$G_2 = \frac{1}{(s+1)^2} \sum_{i=1}^N \int g_i''(u_{x_i}) \left| \partial_{x_i} \left(v_j^{\frac{s+1}{2}} \right) \right|^2 \xi^2.$$

For the term G_3, we estimate it from above: we use Young inequality, so to get

$$\begin{aligned} |G_3| &\le \sum_{i=1}^N \int g_i''(u_{x_i})\, |u_{x_i\,x_j}|\, (u_{x_j}-k_j)_+\, v_j^s\, |\xi_{x_i}|\, \xi \\ &\le \frac{1}{\tau} \sum_{i=1}^N \int_{\{u_{x_j}>k_j\}} g_i''(u_{x_i})\, v_j^{s+1}\, |\nabla \xi|^2 \\ &+ \tau \sum_{i=1}^N \int g_i''(u_{x_i})\, |u_{x_i\,x_j}|^2\, (u_{x_j}-k_j)_+^2\; v_j^{s-1}\, \xi^2, \end{aligned}$$

and the last integral is exactly the same as in G_2.

Terms H_i. We keep the term H_1, which is positive, and we estimate H_2 from below by

$$H_2 \ge \frac{1}{(s+1)^2} \int w^{\frac{q-2}{2}} \left| \nabla \left(v_j^{\frac{s+1}{2}} \right) \right|^2 \xi^2.$$

For H_3 we proceed similarly to G_3, then getting

$$|H_3| \le \frac{1}{\tau} \int_{\{u_{x_j}>k_j\}} w^{\frac{q-2}{2}}\, v_j^{s+1}\, |\nabla \xi|^2 + \tau \int \langle D^2 H(\nabla u)\, D_j^2 u, D_j^2 u\rangle\; v_j^s\, \xi^2,$$

having used Cauchy-Schwarz inequality in the following form

$$\left| \langle D^2 H(\nabla u)\, D_j^2 u, \nabla \xi\rangle \right| \le \sqrt{\langle D^2 H(\nabla u)\, D_j^2 u, D_j^2 u\rangle}\, \sqrt{\langle D^2 H(\nabla u)\, \nabla \xi, \nabla \xi\rangle},$$

the growth of $|D^2 H| \simeq w^{\frac{q-2}{2}}$ and the simple fact $(u_{x_j}-k_j)_+^2 \le v_j$.

Terms B_i. We estimate from above each of these terms, replacing $|b'|$ by a constant, thanks to our assumption (2.1). Then we have

$$|B_1| \le C \int_{\{u_{x_j}>k_j\}} |u_{x_j\,x_j}|\, v_j^s\, \xi^2 \le C\, \tau \int_{\{u_{x_j}>k_j\}} |u_{x_j\,x_j}|^2\, v_j^s\, \xi^2 + \frac{C}{\tau} \int_{\{u_{x_j}>k_j\}} v_j^s\, \xi^2$$

and then we observe that we have

$$1_{\{u_{x_j}>k_j\}} \le \frac{1}{q-1}\, g_j''(u_{x_j})\, 1_{\{u_{x_j}>k_j\}}, \tag{2.14}$$

thanks to the fact that $k_j = \delta_j + 1$. Inserting this information in the previous estimate, we finally get

$$|B_1| \le C\, \tau \int_{\{u_{x_j}>k_j\}} g_j''(u_{x_j})\, |u_{x_j\,x_j}|^2\, v_j^s\, \xi^2 + \frac{C}{\tau} \int_{\{u_{x_j}>k_j\}} g_j''(u_{x_j})\, v_j^{s+1}\, \xi^2,$$

possibly with a different constant C. Notice that we further estimated $v_j^s \le v_j^{s+1}$, thanks to the fact that $v_j \ge 1$.

The term B_2 is readily estimated in a similar manner: we have

$$|B_2| \le C \int (u_{x_j} - k_j)_+^2 \, v_j^{s-1} \, |u_{x_j\,x_j}| \, \xi^2 \le C\,\tau \int g_j''(u_{x_j})\,(u_{x_j} - k_j)_+^2 \, v_j^{s-1} \, |u_{x_j\,x_j}|^2 \, \xi^2 + \frac{C}{\tau} \int g_j''(u_{x_j})\, v_j^{s+1} \, \xi^2,$$

where we used again (2.14) and $(u_{x_j} - k_j)_+^2 \le v_j \le v_j^2$. Finally, we come to the term B_3: we obtain

$$\begin{aligned} |B_3| &\le C \int (u_{x_j} - k_j)_+ \, v_j^s \, |\nabla \xi| \, \xi \le \frac{C}{2} \int (u_{x_j} - k_j)_+^2 \, v_j^s \, \xi^2 + \frac{C}{2} \int_{\{u_{x_j} > k_j\}} v_j^s \, |\nabla \xi|^2 \\ &\le C \int g_j''(u_{x_j}) \, v_j^{s+1} \, \left(\xi^2 + |\nabla \xi|^2\right), \end{aligned}$$

still using (2.14).

We are now ready to put all these estimates together. We keep the lower estimates on G_2 and H_2 on the left, while we put all the other terms on the right. By taking $\tau > 0$ small enough, in order to absorb all the terms appearing on the right and containing the Hessian of u, we finally get

$$\begin{aligned} s \sum_{i=1}^N \int g_i''(u_{x_i}) \left| \partial_{x_i} \left(v_j^{\frac{s+1}{2}} \right) \right|^2 \xi^2 &+ \sum_{i=1}^N \int_{\{u_{x_j} > k_j\}} g_i''(u_{x_i}) \, u_{x_i\,x_j}^2 \, v_j^s \, \xi^2 \\ &+ \varepsilon s \int w^{\frac{q-2}{2}} \left| \nabla \left(v_j^{\frac{s+1}{2}} \right) \right|^2 \xi^2 \\ &+ \varepsilon \int_{\{u_{x_j} > k_j\}} \langle D^2 H(\nabla u) \, D_j^2 u, D_j^2 u \rangle \, v_j^s \, \xi^2 \\ &\le C \sum_{i=1}^N \int_{\{u_{x_j} > k_j\}} g_i''(u_{x_i}) \, v_j^{s+1} \, |\nabla \xi|^2 \\ &+ \varepsilon \, C \int_{\{u_{x_j} > k_j\}} w^{\frac{q-2}{2}} \, v_j^{s+1} \, \left(\xi^2 + |\nabla \xi|^2\right), \qquad j = 1, \dots, N, \end{aligned}$$

i.e. we showed the validity of (2.13). We only have to observe that

$$g_i''(u_{x_i}) \le w^{\frac{q-2}{2}}, \qquad i = 1, \dots, N.$$

Observe that the first integral on the left-hand side is equally performed on the set $\{u_{x_j} > k_j\}$, since otherwise v_j is constant. $\square$

Let us now pay special attention to the case $s = 0$. Computations are very much the same.

Lemma 2.8. *For every $\xi \in C^1_0(\mathcal{O})$, we have*

$$\sum_{i=1}^N \int g_i''(u_{x_i}) \left|\partial_{x_i}\left(v_j^{\frac{1}{2}}\right)\right|^2 \xi^2 + \varepsilon \int w^{\frac{q-2}{2}} \left|\nabla\left(v_j^{\frac{1}{2}}\right)\right|^2 \xi^2 \le C \int_{\{u_{x_j}>k_j\}} w^{\frac{q-2}{2}}\, v_j\, (|\nabla\xi|^2 + \xi^2), \tag{2.15}$$

for $j = 1, \dots, N$, for some constant C independent of ε.

Proof. We repeat the previous computations, using the test function $\psi^+_{j,0} = (u_{x_j} - k_j)_+\, \xi^2$. This gives

$$\begin{aligned}\sum_{i=1}^N \int_{\{u_{x_j}>k_j\}} g_i''(u_{x_i})\, u^2_{x_i\,x_j}\, \xi^2 + \varepsilon \int_{\{u_{x_j}>k_j\}} \langle D^2H(\nabla u)\, D^2_j u\, D^2_j u\rangle\, \xi^2 &\\ \le C \sum_{i=1}^N \int_{\{u_{x_j}>k_j\}} g_i''(u_{x_i})\, v_j\, |\nabla\xi|^2 &\\ + \varepsilon\, C \int_{\{u_{x_j}>k_j\}} w^{\frac{q-2}{2}}\, v_j\, \left(\xi^2 + |\nabla\xi|^2\right), \qquad j = 1, \dots, N.&\end{aligned}$$

By combining this together with

$$u^2_{x_ix_j} \ge \left|\partial_{x_i}\left(v_j^{\frac{1}{2}}\right)\right|^2 \qquad \text{and} \qquad \int_{\{u_{x_j}>k_j\}} \langle D^2H(\nabla u)\, D^2_j u, D^2_j u\rangle\, \xi^2 \ge \int w^{\frac{q-2}{2}} \left|\nabla\left(v_j^{\frac{1}{2}}\right)\right|^2 \xi^2,$$

and using again $g_i''(u_{x_i}) \le w^{\frac{q-2}{2}}$, we readily get the thesis. □

Remark 2.9. The previous estimates are valid for the functions v_j, which are different from 0 if u_{x_j} is large and positive. If on the contrary u_{x_j} is large in absolute value but negative, we can repeat the same estimates of Lemma 2.6 and Lemma 2.7, this time using as test functions

$$\varphi^-_{j,\beta} = u\, z_j^\beta\, \xi^2 \qquad \text{and} \qquad \psi^-_{j,s} = (-u_{x_j} - k_j)_+\, z_j^s\, \xi^2,$$

where z_j is given by

$$z_j = (-u_{x_j} - k_j)^2_+ + 1, \qquad j = 1, \dots, N.$$

Then we derive inequalities (2.9), (2.13) and (2.15), with z_j in place of v_j.

2.4. **Step 4: a slow Moser's iteration.** Gluing together the estimates of Lemmas 2.6, 2.7 and 2.8 and tuning the exponents β and s, we obtain the following intermediate estimate, that we enunciate as a separate result for the sake of readability.

Lemma 2.10. *Let $\alpha \ge 0$, for every $\xi \in C^1_0(\mathcal{O})$, we have*

$$\int w^{\frac{q}{2}}\, v_j^{\frac{1}{2}+\alpha}\, \xi^2 \le C\,(\alpha+1) \int w^{\frac{q}{2}}\, v_j^\alpha\, \left(\xi^2 + |\nabla\xi|^2\right), \qquad j = 1, \dots, N, \tag{2.16}$$

for some constant C independent of ε.

Proof. First of all, we consider the case $\alpha > 0$ and make the choices

$$\beta = \frac{1}{2} + \alpha \qquad \text{and} \qquad s = 2\,\beta - \alpha - 1 = \alpha,$$

in (2.9) and (2.13). Then we drop the $\varepsilon-$term in the left-hand side of (2.9). In this way, using $v_j \le w$, we obtain

$$\begin{aligned}\sum_{i=1}^N \int g_i''(u_{x_i})\,|u_{x_i}|^2\, v_j^{\frac{1}{2}+\alpha}\,\xi^2 &\le C\,(\alpha+1)\,\int w^{\frac{q-2}{2}}\,v_j^{\alpha+1}\,\left(\xi^2+|\nabla\xi|^2\right)+C\int w^{\frac{q}{2}} v_j^{\alpha}\xi^2 \\ &\le C\,(\alpha+1)\int w^{\frac{q}{2}}\,v_j^{\alpha}\,\left(\xi^2+|\nabla\xi|^2\right),\end{aligned} \tag{2.17}$$

for some constant C independent of ε. We now estimate from below the left-hand side of (2.17): thanks to (2.5), we have

$$\sum_{i=1}^N g_i''(u_{x_i})\,|u_{x_i}|^2 \ge C_q\,\sum_{i=1}^N |u_{x_i}|^q - C_q' \ge C_{q,N}\,w^{\frac{q}{2}} - C_{q,N}'', \tag{2.18}$$

thanks to the fact that in $\mathbb{R}^N$ all norms are equivalent and to the simple estimate

$$(t-1)^{\frac{q}{2}} \ge 2^{\frac{2-q}{2}}\,t^{\frac{q}{2}}-1, \qquad t\ge 0.$$

Using this into (2.17) and using as always C as a generic constant depending on the data of the problem, but not on ε, we can thus infer

$$\int w^{\frac{q}{2}}\,v_j^{\frac{1}{2}+\alpha}\,\xi^2 \le C\,(\alpha+1)\int w^{\frac{q}{2}}\,v_j^{\alpha}\,\left(\xi^2+|\nabla\xi|^2\right)+C\,(\alpha+1)\int v_j^{\frac{1}{2}+\alpha}\,\xi^2,$$

which finally yields the thesis, by exploiting again that $1\le v_j\le w$ and $q\ge 2$.

To treat the case $\alpha=0$, which gives the first gain of integrability in Moser's iterations, we proceed similarly. We combine (2.15) and (2.9) with $\beta=1/2$ and $\alpha=0$. Then we use $v_j^{1/2}\le v_j$, thus arriving at

$$\sum_{i=1}^N \int g_i''(u_{x_i})\,|u_{x_i}|^2\, v_j^{\frac{1}{2}}\,\xi^2 \le C\,\int w^{\frac{q-2}{2}}\,v_j\,\left(\xi^2+|\nabla\xi|^2\right)+C\int w^{\frac{q}{2}}\,\xi^2 \le 2\,C\,\int w^{\frac{q}{2}}\,\left(\xi^2+|\nabla\xi|^2\right).$$

Again using (2.18), we immediately deduce the thesis. □

2.5. **Step 5: proof of Proposition 2.1.** Keeping in mind Remark 2.9, the same estimate (2.16) holds with z_j in place of v_j, so that summing up we get

$$\int w^{\frac{q}{2}}\,\left(v_j^{\frac{1}{2}+\alpha}+z_j^{\frac{1}{2}+\alpha}\right)\xi^2\,dx \le C\,(\alpha+1)\int w^{\frac{q}{2}}\,\left(v_j^{\alpha}+z_j^{\alpha}\right)\,\left(\xi^2+|\nabla\xi|^2\right)\,dx, \qquad j=1,\dots,N.$$

If we set

$$T_j=\max\{v_j,z_j\}, \qquad j=1,\dots,N,$$

from the previous we can easily infer

$$\int w^{\frac{q}{2}}\,T_j^{\frac{1}{2}+\alpha}\,\xi^2\,dx \le 2\,C\,(\alpha+1)\,\int w^{\frac{q}{2}}\,T_j^{\alpha}\,\left(\xi^2+|\nabla\xi|^2\right)\,dx, \qquad j=1,\dots,N.$$

Summing up all these inequalities and setting $T=\max\{T_1,\dots,T_N\}$, we get

$$\int w^{\frac{q}{2}}\,T^{\frac{1}{2}+\alpha}\,\xi^2\,dx \le 2\,N\,C\,(\alpha+1)\,\int w^{\frac{q}{2}}\,T^{\alpha}\,\left(\xi^2+|\nabla\xi|^2\right)\,dx.$$

Finally, we observe that there exist two constants $\gamma_1, \gamma_2 > 0$ depending only on the dimension N and $\max\{\delta_1, \dots, \delta_N\}$, such that

$$\gamma_1\, T \le w \le \gamma_2\, T.$$

Inserting this information in the previous inequality, we get

$$\int T^{\frac{q}{2}+\alpha+\frac{1}{2}}\, \xi^2\, dx \le C\,(\alpha+1) \int T^{\frac{q}{2}+\alpha}\, \left(\xi^2 + |\nabla \xi|^2\right)\, dx.$$

Fixing two balls $B_\varrho(x_0) \subset B_R(x_0)$ and suitably choosing a sequence of cut-off functions $\{\xi_k\}_{k\in\mathbb{N}} \subset C^1_0(\mathcal{O})$ supported on an infinite family of shrinking balls $B_\varrho(x_0) \subset B_{r_{k+1}}(x_0) \subset B_{r_k}(x_0) \subset B_R(x_0)$, we can iterate the previous estimate, taking $\alpha_k = k/2$, with $k \in \mathbb{N}$. Then a standard covering argument concludes the proof of Proposition 2.1.

3. Proof of the Main Theorem

We now come to the proof of the Main Theorem. Let $f \in L^\infty_{loc}(\Omega)$ and call u a local minimizer. We have to show that $u \in W^{1,r}_{loc}(\Omega)$ for every $r \ge 1$. At this aim, it is sufficient to show that for every ball $B \Subset \Omega$, then $u \in W^{1,r}(B)$ for every $r \ge 1$.

We then fix a ball $B \Subset \Omega$ and consider a slightly larger ball $B' \Subset \Omega$. For $\delta > 0$ sufficiently small, we take a standard mollification kernel η_δ with compact support and set

$$u^\delta = (u * \eta_\delta) \cdot 1_{B'} \in C^\infty(\overline{B'}).$$

Now we fix $\varepsilon \ll 1$ and take $f^\varepsilon \in C^\infty(\Omega)$, such that f^ε $*-$weak converges in L^∞_{loc} to f. In particular, for every $\Omega' \Subset \Omega$ the norms $\|f^\varepsilon\|_{L^\infty(\Omega')}$ can be taken to be uniformly bounded, by a constant depending on $\|f\|_{L^\infty(\Omega')}$, but not on ε.

Then, let $u^{\varepsilon,\delta}$ be a solution of

$$\min\left\{\mathfrak{F}^\varepsilon_q(v; B') + P(v - u^\delta; B')\, :\, v - u^\delta \in W^{1,q}_0(B')\right\}, \tag{3.1}$$

where

$$\mathfrak{F}^\varepsilon_q(v; B') = \sum_{i=1}^N \int_{B'} g^\varepsilon_i(v_{x_i})\, dx + \int_{B'} f^\varepsilon(x)\, v(x)\, dx + \varepsilon \int_{B'} H(\nabla v(x))\, dx, \qquad v \in W^{1,q}(B'),$$

and P is a penalization term, given by

$$P(v; B') = \int_{B'} \exp(-|v(x)|^2)\, dx, \qquad v \in W^{1,q}(B').$$

Lemma 3.1 (Uniform estimates). *The following estimate holds*

$$\|u^{\varepsilon,\delta}\|_{W^{1,q}(B')} \le C\, \left(\|u\|_{W^{1,q}(B')} + 1\right), \tag{3.2}$$

for some constant C not depending on δ and ε. Moreover, $u^{\varepsilon,\delta} \in W^{1,r}(B)$ for every $r \ge 1$ and we have the estimate

$$\|u^{\varepsilon,\delta}\|_{W^{1,r}(B)} \le C,$$

for a constant C depending on $q, r, \operatorname{dist}(B, \partial\Omega), \|u\|_{W^{1,q}(B)}$ and $\|f\|_{L^\infty(B)}$, but not on ε and δ.

Proof. We consider the Euler-Lagrange equation of problem (3.1), tested with $\varphi = u^{\varepsilon,\delta} - u^\delta \in W^{1,q}_0(B')$. This yields

$$\sum_{i=1}^N \int_{B'} g_i'(u^{\varepsilon,\delta}_{x_i})\, u^{\varepsilon,\delta}_{x_i} + \varepsilon \int_{B'} \langle H(\nabla u^{\varepsilon,\delta}), \nabla u^{\varepsilon,\delta}\rangle \le \sum_{i=1}^N \int_{B'} |g_i'(u^{\varepsilon,\delta})|\, |u^\delta_{x_i}| + \varepsilon \int_{B'} |\nabla H(\nabla u^{\varepsilon,\delta})|\, |\nabla u^\delta| + C\, \|u^{\varepsilon,\delta} - u^\delta\|_{L^q(B')},$$

for a constant depending on q, $|B'|$ and $\|f\|_{L^\infty(B')}$, but not on ε and δ. Using the growth conditions on ∇H and g_i' and Young inequality, from the previous we can infer

$$\int_{B'} |\nabla u^{\varepsilon,\delta}(x)|^q\, dx \le C \int_{B'} |\nabla u^\delta(x)|^q\, dx + C\, \|u^{\varepsilon,\delta} - u^\delta\|_{L^q(B')}\, dx + C.$$

Finally using Poincaré inequality for the function $u^{\varepsilon,\delta} - u^\delta \in W^{1,q}_0(B')$, we obtain

$$\|u^{\varepsilon,\delta}\|_{W^{1,q}(B')} \le C \left(\|u^\delta\|_{W^{1,q}(B')} + 1 \right),$$

for a constant independent of ε and δ. Recalling the definition of u^δ, this finally gives (3.2).

The second part of the statement is a consequence of Proposition 2.1, applied with $\mathcal{O} = B'$, $\zeta = u^\delta$ and

$$b(x,t) = f^\varepsilon(x)\, t + \exp(-|t - u^\delta(x)|^2), \qquad x \in B',\ t \in \mathbb{R},$$

in conjunction with (3.2). □

We also need a $\Gamma-$convergence result. The reader is referred to [11] for an introduction to the general theory of $\Gamma-$convergence. The main point here is that of inferring the convergence of minimizers.

Lemma 3.2. *Let $\delta > 0$ be given and $\{\varepsilon_k\}_{k\in\mathbb{N}}$ be a sequence of positive reals converging to 0, then the functionals*

$$u \mapsto \mathfrak{F}^{\varepsilon_k}_q(u) + P(u - u^\delta), \qquad u \in W^{1,q}(B'),$$

are $\Gamma-$converging to $\mathfrak{F}_q + P(\cdot - u^\delta)$ with respect to the $W^{1,q}(B')$ weak topology. Moreover, a sequence of minimizers $\{u^{\varepsilon_k,\delta}\}_{k\in\mathbb{N}}$ weakly converges (up to a subsequence) in $W^{1,q}(B')$ to a minimizer $u^{0,\delta}$ of

$$\min\left\{ \mathfrak{F}_q(v; B') + P(v - u^\delta; B') \,:\, v - u^\delta \in W^{1,q}_0(B') \right\}. \tag{3.3}$$

Proof. First of all, we observe that the additive term $u \mapsto P(u - u^\delta)$ is not dependent on ε and it is continuous with respect to the $W^{1,q}(B')$ weak convergence, then it is sufficient to prove the $\Gamma-$convergence of the functionals $\mathfrak{F}^{\varepsilon_k}_q$, thanks to [11, Proposition 6.21].

$\Gamma-$*liminf inequality.* Let $u \in W^{1,q}(B')$ and $\{u^{\varepsilon_k}\}_{k\in\mathbb{N}} \subset W^{1,q}(B')$ a sequence weakly converging to u. Thanks to the strong convergence of $\{u^\varepsilon_k\}_{k\in\mathbb{N}}$ to u in $L^p(B')$, one immediately gets

$$\lim_{k\to\infty} \int_{B'} f^{\varepsilon_k}(x)\, u^{\varepsilon_k}(x)\, dx = \lim_{k\to\infty} \int_{B'} f^{\varepsilon_k}(x)\, u(x)\, dx = \int_{B'} f(x)\, u(x)\, dx,$$

where we furter used that $\{f^{\varepsilon_k}\}_{k\in\mathbb{N}}$ is $*-$weak convergent in $L^\infty(B')$. Using that $|g_i^\varepsilon(t)-g_i(t)|\le C\,\varepsilon^2$, we obtain:

$$\begin{aligned}\liminf_{k\to\infty}\int_{B'} g_i^{\varepsilon_k}(u^{\varepsilon_k}_{x_i}(x))\,dx &\ge \liminf_{k\to\infty}\left[\int_{B'} g_i(u^{\varepsilon_k}_{x_i}(x))\,dx-\int_{B'}|g_i^{\varepsilon_k}(u^{\varepsilon_k}_{x_i}(x))-g_i(u^{\varepsilon_k}_{x_i}(x))|\,dx\right]\\ &\ge \liminf_{k\to\infty}\left[\int_{B'} g_i(u^{\varepsilon_k}_{x_i}(x))\,dx - C\,\varepsilon_k^2\,\left|B'\right|\right]\\ &=\int_{B'} g_i(u_{x_i}(x))\,dx,\qquad i=1,\dots,N,\end{aligned}$$

where we used the convexity of g_i, which guarantees the lower semicontinuity of the first integral. Using the previous estimates and dropping the term containing $H(\nabla u^{\varepsilon_k})$ in the definition of $\mathfrak{F}_q^{\varepsilon_k}$, we finally end up with

$$\begin{aligned}\liminf_{k\to\infty}\mathfrak{F}_q^{\varepsilon_k}(u^{\varepsilon_k}) &\ge \liminf_{k\to\infty}\left[\sum_{i=1}^N\int_{B'} g_i^{\varepsilon_k}(u^{\varepsilon_k}_{x_i})\,dx+\int_{B'} f^{\varepsilon_k}(x)\,u^{\varepsilon_k}(x)\,dx\right]\\ &\ge \sum_{i=1}^N\int_{B'} g_i(u_{x_i})\,dx+\int_{B'} f(x)\,u(x)\,dx=\mathfrak{F}_q(u).\end{aligned}$$

$\Gamma-$*limsup inequality.* Let $u\in W^{1,q}(B')$ and choose the constant sequence $u^{\varepsilon_k}=u$, for every $k\in\mathbb{N}$. Again thanks to the uniform convergence of the functions $g_i^{\varepsilon_k}$, an application of the Lebesgue Dominated Convergence Theorem leads to

$$\lim_{k\to\infty}\left[\sum_{i=1}^N\int_{B'} g_i^{\varepsilon_k}(u_{x_i})\,dx+\int_{B'} f^{\varepsilon_k}(x)\,u(x)\,dx+\varepsilon\int_\Omega H(\nabla u(x))\,dx\right]=\mathfrak{F}_q(u),$$

which finally gives the desired $\Gamma-$convergence result.

For the last part of the statement, we first observe that introducing

$$M_\delta(u)=\left\{\begin{array}{ll}0, & \text{if } u-u^\delta\in W^{1,q}_0(B'),\\ +\infty, & \text{otherwise}\end{array}\right.$$

then the functional

$$u\mapsto \mathfrak{F}_q^{\varepsilon_k}(u)+P(u-u^\delta)+M_\delta(u),$$

is $\Gamma-$converging to $\mathfrak{F}_q(u)+P(u-u^\delta)+M_\delta(u)$. Indeed, it is sufficient to repeat the previous proof. The $\Gamma-\liminf$ inequality remains true, thanks to the lower semicontinuity of M_δ. On the other hand, for every $u\in W^{1,q}(B')$, if we take the constant sequence $u^{\varepsilon_k}=u$, then we have

$$\lim_{k\to\infty}\mathfrak{F}^{\varepsilon_k}(u)+M_\delta(u)=\mathfrak{F}_q(u)+M_\delta(u).$$

We can conclude by using [11, Corollary 7.20], once it is observed that the minimizers $\{u^{\varepsilon_k,\delta}\}_{k\in\mathbb{N}}$ satisfy the equi-coercivity condition (3.2). □

Proof of the Main Theorem. Now, we first pass to the limit as ε goes to 0. Thanks to (3.2), there exists a subsequence $\{u^{\varepsilon_k,\delta}\}_{k\geq 0}$ weakly converging in $W^{1,q}(B')$ to a limit function $u^{0,\delta}$. Moreover, thanks to Lemma 3.2, this $u^{0,\delta}$ is a minimizer of (3.3). We now take the limit as δ goes to 0. Still from (3.2) we have

$$\|u^{0,\delta}\|_{W^{1,q}(B')} \leq C\,\left(\|u\|_{W^{1,q}(B')}+1\right),$$

then there exists a subsequence $\{u^{0,\delta_k}\}_{k\geq 0}\subset W^{1,q}(B')$ which weakly converges in $W^{1,q}(B')$ to a function u^0. Using the minimality of u^{0,δ_k}, the semicontinuity of the penalized functional and the continuity of the functional $\mathfrak{F}_q(\cdot\,;B')$ with respect to the strong convergence, we get

$$\begin{aligned}\mathfrak{F}_q(u^0;B')+P(u^0-u;B') &\leq \liminf_{k\to\infty}\left[\mathfrak{F}_q(u^{0,\delta_k};B')+P(u^{0,\delta_k}-u^{\delta_k};B')\right]\\ &\leq \liminf_{k\to\infty}\mathfrak{F}_q(u^{\delta_k};B')=\mathfrak{F}_q(u;B').\end{aligned}$$

Finally, we use the fact that u is a local minimum and that $u^0-u\in W^{1,q}_0(B')$, then

$$\mathfrak{F}_q(u^0;B')+P(u^0-u;B')\leq \mathfrak{F}_q(u;B')\leq \mathfrak{F}_q(u^0;B'),$$

which shows that $u=u^0$ almost everywhere in B'.

Let us now observe that thanks to Lemma 3.1, we have that $u^{\varepsilon_k,\delta}\in W^{1,r}_{loc}(B')$, for every $r\geq 1$. In particular, $u^{\varepsilon_k,\delta}\in W^{1,r}(B)$ for every $r\geq 1$ and we have a uniform estimate of the type

$$\|\nabla u^{\varepsilon_k,\delta}\|_{W^{1,r}(B)}\leq C,$$

with C independent of ε_k and δ. Using the fact that $u^{\varepsilon_k,\delta}$ converges to $u^{0,\delta}$, we get that $u^{0,\delta}\in W^{1,r}(B)$ for every $r\geq 1$ as well, with an estimate uniform in δ. Finally, taking the limit as δ goes to 0, from the previous discussion we get that $u\in W^{1,r}(B)$, for every $r\geq 1$. This finally concludes the proof of the Main Theorem. □

4. Applications to Beckmann's problem

Going back to our original purpose, it is mandatory to conclude the paper with some applications to Beckmann's problem (1.3).

Corollary 4.1. *Let $f\in L^\infty(\Omega)$ be such that $\int_\Omega f(x)\,dx=0$ and set*

$$W^{1,q}_\diamond(\Omega)=\left\{v\in W^{1,q}(\Omega)\,:\,\int_\Omega v(x)\,dx=0\right\}.$$

Every solution u of the following variational problem

$$\min\left\{\mathfrak{F}_q(v;\Omega)\,:\,v\in W^{1,q}_\diamond(\Omega)\right\},\tag{4.1}$$

satisfies $u\in W^{1,r}_{loc}(\Omega)$, for every $r\geq 1$.

Proof. Let u be a minimizer and let us take $\Omega'\Subset\Omega$. We take $\varphi\in W^{1,q}_0(\Omega')$ and extend it by 0 to the whole Ω, then we set

$$c=\frac{1}{|\Omega|}\int_\Omega \varphi(x)\,dx=\frac{1}{|\Omega|}\int_{\Omega'}\varphi(x)\,dx.$$

The function $u+\varphi-c$ is admissible for problem (4.1), so that

$$\mathfrak{F}_q(u;\Omega) \le \mathfrak{F}_q(u+\varphi;\Omega),$$

where we used that $\mathfrak{F}_q(v+c)=\mathfrak{F}_q(v)$ for every constant c. We then observe that φ is supported in Ω', so that we can write

$$\mathfrak{F}_q(u+\varphi;\Omega)=\mathfrak{F}_q(u+\varphi;\Omega')+\mathfrak{F}_q(u;\Omega\setminus\Omega'),$$

that is, using this information in the previous inequality, we get

$$\mathfrak{F}_q(u;\Omega')+\mathfrak{F}_q(u;\Omega\setminus\Omega') \le \mathfrak{F}_q(u+\varphi;\Omega')+\mathfrak{F}_q(u;\Omega\setminus\Omega').$$

This finally shows that u is a local minimizer of the functional $\mathfrak{F}_q$. The thesis now follows by simply applying the Main Theorem to u. □

An "almost" L^∞ estimate for the optimal transportation program is now an easy consequence of the previous result and the primal-dual optimality condition.

Corollary 4.2. *Let $f\in L^\infty(\Omega)$ be such that $\int_\Omega f(x)\,dx=0$ and $1<p\le 2$. Then the (unique) vector field $\widetilde{\phi}\in L^p(\Omega;\mathbb{R}^N)$ which solves*

$$\min_{\phi\in L^p(\Omega;\mathbb{R}^N)}\left\{\sum_{i=1}^N\int_\Omega\left[\frac{|\phi_i(x)|^p}{p}+\delta_i\,|\phi_i(x)|\right]dx\,:\,\operatorname{div}\phi=f\ \text{in}\ \Omega,\ \langle\phi,\nu_\Omega\rangle=0\ \text{on}\ \partial\Omega\right\},$$

is in $L^r_{loc}(\Omega;\mathbb{R}^N)$, for every $r\ge 1$.

Proof. Existence and uniqueness of $\widetilde{\phi}\in L^p(\Omega;\mathbb{R}^N)$ is straightforward, since we are minimizing a strictly convex energy with $p-$growth, under a linear and closed constraint. By standard convex duality, we get the primal-dual optimality condition

$$\widetilde{\phi}_i=(|u_{x_i}|-\delta_i)_+^{q-1}\,\frac{u_{x_i}}{|u_{x_i}|},\qquad i=1,\dots,N,$$

with $q=p/(p-1)\ge 2$ and $u\in W^{1,q}(\Omega)$ solution of (4.1). Then the result follows from Corollary 4.1. □

Acknowledgements. This work has been supported by the ANR through the projects ANR-09-JCJC-0096-01 EVAMEF and ANR-07-BLAN- 0235 OTARIE, as well as by the ERC Advanced Grant n. 226234. Nina Uralt'seva is gratefully acknowledged for having pointed out to us the reference [17].

References

[1] J.-B. Baillon, G. Carlier, From discrete to continuous Wardrop equilibria, Netw. Heterogenous Media, **7** (2012), 219–241.
[2] M. J. Beckmann, A continuous model of transportation, Econometrica **20** (1952), 643–660.
[3] M. Belloni, B. Kawohl, The pseudo $p-$Laplace eigenvalue problem and viscosity solution as $p\to\infty$, ESAIM Control Optim. Calc. Var., **10** (2004), 28–52.
[4] M. Bildhauer, M. Fuchs, X. Zhong, A regularity theory for scalar local minimizers of splitting-type variational integrals, Ann. Sc. Norm. Super. Pisa Cl. Sci., **6** (2007), 385–404.
[5] M. Bildhauer, M. Fuchs, X. Zhong, Variational integrals with a wide range of anisotropy, St. Petersburg Math. J., **18** (2007), 717–736.
[6] L. Brasco, Global L^∞ gradient estimates for solutions to a certain degenerate elliptic equation, Nonlinear Anal., **72** (2011), 516–531.
[7] L. Brasco, G. Carlier, F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations, J. Math. Pures Appl., **93** (2010), 652–671.

[8] L. Brasco, G. Franzina, A nonlinear eigenvalue problem of Stekloff type, in preparation.

[9] G. Carlier, C. Jimenez, F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria, SIAM J. Control Optim. **47** (2008), 1330-1350.

[10] P. Celada, G. Cupini, M. Guidorzi, Existence and regularity of minimizers of nonconvex integrals with $p-q$ growth, ESAIM Control Optim. Calc. Var., **13** (2007), 343–358.

[11] G. Dal Maso, An Introduction to $\Gamma-$Convergence. Progress in Nonlinear Differential Equations and their Applications, **8**. Birkhäuser Boston, Inc., Boston, MA, 1993.

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

[13] L. Esposito, G. Mingione, C. Trombetti, On the Lipschitz regularity for certain elliptic problems, Forum Math. **18** (2006), 263–292.

[14] I. Fonseca, N. Fusco, P. Marcellini, An existence result for a nonconvex variational problem via regularity, ESAIM Control Optim. Calc. Var. **7** (2002), 69–95.

[15] E. Giusti, Metodi diretti nel calcolo delle variazioni. (Italian) [Direct methods in the calculus of variations], Unione Matematica Italiana, Bologna, 1994.

[16] G. Mingione, A. Zatorksa-Goldstein, X. Zhong, Gradient regularity for elliptic equations in the Heisenberg group, Adv. Math., **222** (2009), 62–129.

[17] N. Uralt'seva, N. Urdaletova, The boundedness of the gradients of generalized solutions of degenerate quasilinear nonuniformly elliptic equations, Vest. Leningr. Univ. Math., **16** (1984), 263–270.

[18] J. Vétois, The blow-up of critical anistropic equations with critical directions, NoDEA Nonlinear Differential Equations and Applications, **18** (2011), 173–197.

L. B., LATP, Aix–Marseille Université, 13453, Marseille, France
E-mail address: `brasco@cmi.univ-mrs.fr`

G. C., Ceremade, Université Paris Dauphine, 75775, Paris Cedex 16, France
E-mail address: `carlier@ceremade.dauphine.fr`