A FRACTIONAL GEHRING LEMMA, WITH APPLICATIONS TO NONLOCAL EQUATIONS

TUOMO KUUSI, GIUSEPPE MINGIONE, AND YANNICK SIRE

To Carlo Sbordone on his 65th birthday

Abstract. We describe a fractional version of the classical Gehring lemma. As a consequence, new self-improving regularity properties of solutions to integrodifferential equations emerge.

1. The classical Gehring lemma

The Gehring lemma [7, 9] is a fundamental tool in modern nonlinear analysis, with crucial implications in several different fields, ranging from nonlinear elliptic and parabolic equations to the calculus of variations, from quasiconformal geometry to stability issues [2, 6, 11]. Its ultimate essence relies on a basic, self-improving property of certain kind of inequalities, called reverse Hölder type inequalities. This can be described as follows: if one can control the L^p-means of a given function $f \in L^p$, at all scales, with similar L^q-means, and $p > q$, then the function f is necessarily better than just being in L^p. Starting from the original work of Gehring, there have been several different versions of this result; see [9] for a panorama. The following one, involving reverse inequalities with increasing support, can be for instance found in [8].

Theorem 1.1. *Let $f \in L^p_{\mathrm{loc}}(\Omega)$, $p > 1$ be a non-negative function such that the following reverse Hölder type inequality holds whenever B is a ball in the open subset $\Omega \subset \mathbb{R}^n$:*

$$\left(\fint_{B/2} f^p \, dx \right)^{1/p} \leq c \left(\fint_B f^q \, dx \right)^{1/q} ,$$

where $0 < q < p$. Then there exists a number $\delta > 0$, depending only on n, q and the constant c appearing in the previous inequality, such that $f \in L^{p+\delta}_{\mathrm{loc}}(\Omega)$. Moreover the following inequality holds whenever $B \subset \Omega$ is a ball:

$$\left(\fint_{B/2} f^{p+\delta} \, dx \right)^{1/(p+\delta)} \leq \tilde{c} \left(\fint_B f^q \, dx \right)^{1/q} ,$$

for a new constant $\tilde{c} \equiv \tilde{c}(n, q, c)$.

In the previous statement, as in the rest of this paper, we are using the standard notation

$$(h)_{\mathcal{O}} \equiv \fint_{\mathcal{O}} h \, d\mu := \frac{1}{\mu(\mathcal{O})} \int_{\mathcal{O}} h \, d\mu$$

to denote the average of an integrable function h with respect to a measure μ, over a measurable set $\mathcal{O}$ with positive measure $\mu(\mathcal{O}) > 0$.

The applications of Theorem 1.1 to solutions to linear and nonlinear PDE are particularly relevant. A model result is about the higher gradient integrability of weak energy solutions to divergence form equations of the type

$$-\operatorname{div}(A(x)Du) = 0 \qquad \text{in } \Omega \subset \mathbb{R}^n , \qquad n \geq 2 . \tag{1.1}$$

The matrix of coefficients $A(\cdot)$ is supposed to have measurable entries, and to be bounded and elliptic, i.e., both

$$\Lambda^{-1}|\xi| \leq \langle A(x)\xi, \xi\rangle \qquad \text{and} \qquad |A(x)| \leq \Lambda$$

hold whenever $x \in \Omega$ and $\xi \in \mathbb{R}^n$, where $\Lambda \geq 1$. The ultimate outcome in this case is the higher gradient integrability of energy distributional solutions, that is

$$u \in W^{1,2} \Longrightarrow u \in W^{1,2+\delta} \tag{1.2}$$

holds for some δ depending only on n and Λ. This result was first proved by Meyers [14] for linear equations; modern proofs extending to nonlinear ones are indeed based on Theorem 1.1 [5, 8]. The key fact is that weak solutions satisfy energy inequalities, often called Caccioppoli type inequalities - i.e. inequalities of the type (1.3) below; in turn these imply higher integrability. This is summarised in the next

Theorem 1.2. *Let $u \in W^{1,2}(\mathbb{R}^n)$ be a function such that the following Caccioppoli type inequality holds for every ball $B \equiv B(x_0, r) \subset \mathbb{R}^n$ with centre x_0 and radius $r > 0$:*

$$\int_B |D(u\psi)|^2 dx \leq \frac{c}{r^2}\int_B |u(x) - (u)_B|^2\, dx\,, \tag{1.3}$$

whenever $\psi \in C_0^\infty(B(x_0, 3r/4))$ is a cut-off function such that $|D\psi| \leq c/r$. Then there exists a positive number $\delta \in (0,1)$, depending only on c and n, such that $u \in W^{1,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$.

The route from inequality (1.3) to higher gradient integrability is straightforward. Indeed, applying Sobolev-Poincaré inequality we get that the following reverse type inequality with increasing support holds for a constant that depends only on n, c and for a new constant $c_0 \equiv c_0(n, c)$:

$$\left(\fint_{B/2} |Du|^2 dx\right)^{1/2} \leq c_0 \left(\fint_B |Du|^{2n/(n+2)} dx\right)^{(n+2)/2n}.$$

At this point Theorem 1.1 finally implies that $Du \in L^{2+\delta}_{\mathrm{loc}}$ for some $\delta > 0$, depending only on n, c, but not on the specific function u.

2. The fractional Gehring Lemma

Here we are going to report the main facts from [13], to which we refer for a more complete presentation and for the detailed proofs. With applications to regularity of solutions to nonlocal problems in mind, we here present a fractional analog of the classical Gehring's lemma. For the sake of simplicity, we shall confine ourselves to a simpler situation. More general cases can be indeed found in [13].

Let us recall that a function v belongs to the fractional Sobolev space $W^{s,\gamma}(\mathbb{R}^n)$, with $s \in (0,1)$ and $\gamma \geq 1$ iff $v \in L^s(\mathbb{R}^n)$ and

$$[v]^\gamma_{s,\gamma} := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^\gamma}{|x-y|^{n+\gamma s}}\, dx\, dy < \infty\,. \tag{2.1}$$

Local variants of the space $W^{s,\gamma}(\mathbb{R}^n)$ are defined in the usual way, while in this paper we shall always consider the case $n \geq 2$. The main novelty in our result is the fact that, on the contrary of what happens in the local case, the self-improvement happens in the differentiability scale, which is the leading one. As we shall see later, when applied to solutions to nonlocal equations, this will lead us to discover a new regularity property of solutions to nonlocal equations that has no parallel in the theory of classical local elliptic equations; see Remark 3.1 below. Our fractional version of Gehring's lemma will show that, starting from a $W^{\alpha,2}$-function satisfying the natural Caccioppoli's inequality, we will observe the improvement

$$W^{\alpha,2} \Longrightarrow W^{\alpha+\delta,2+\delta}\,.$$

This a surprising new feature of nonlocal problems, since the natural analog of (1.2) is in this case

$$W^{\alpha,2} \Rightarrow W^{\alpha,2+\delta}\,. \tag{2.2}$$

See Remark 3.2 below. We indeed have the following:

Theorem 2.1 (Fractional Caccioppoli inequality)**.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a function such that the following nonlocal Caccioppoli type inequality holds for every ball $B \equiv B(x_0, r) \subset \mathbb{R}^n$:*

$$\begin{aligned}\int_B\int_B &\frac{|[u(x)-(u)_B]\psi(x)-[u(y)-(u)_B]\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\\ &\leq \frac{c}{r^{2\alpha}}\int_B |u(x)-(u)_B|^2\,dx\\ &\quad +c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy\int_B |u(x)-(u)_B|\,dx\end{aligned} \tag{2.3}$$

whenever $\psi \in C^\infty_c(B(x_0, 3r/4))$ is a cut-off function such that $|D\psi| \leq c/r$. Then there exists a positive number $\delta \in (0, 1-\alpha)$, depending only on c and n, such that $u \in W^{\alpha+\delta,2+\delta}_{\rm loc}(\mathbb{R}^n)$.

The type of Caccioppoli inequality involved in the previous lemma is the natural analogue of the local one in display (1.3). We note the presence of an additional "tail" term on the right hand side of (2.3). The presence of this term encodes the fact that the problems inequality (2.3) is typically stemming from are nonlocal and defined on the whole $\mathbb{R}^n$. Ultimately, the last term in (2.3) serves to take into account the long distance interactions which are typical of nonlocal problems. In the previous theorem the crucial information is really given by the differentiability gain. Indeed, assuming that $u \in W^{\alpha+\delta,2}$ for some $\delta > 0$ allows to deduce, via the fractional version of Sobolev embedding theorem, that $u \in W^{\alpha+\delta',2+\delta'}$ for some positive $\delta' < \delta$.

As mentioned above, the key to the proof of the previous theorem is a new type of fractional Gehring lemma. Rather than holding for functions this new version holds for what we are going to call "dual pairs". These are introduced in the following:

Definition 1. Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ and let $\varepsilon \in (0, \alpha/2)$. Define the function

$$U(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{\alpha+\varepsilon}}\,. \tag{2.4}$$

whenever $x \neq y$ and the measure

$$\mu(A) := \int_A \frac{dx\,dy}{|x-y|^{n-2\varepsilon}}\,, \tag{2.5}$$

whenever $A \subset \mathbb{R}^{2n}$ is a measurable subset. The couple (μ, U) is called a dual pair generated by the function u.

The use of the terminology "dual pair" is then motivated by the following equivalence, which holds whenever $u \in L^2(\mathbb{R}^n)$:

$$u \in W^{\alpha,2}(\mathbb{R}^n) \iff U \in L^2(\mathbb{R}^{2n};\mu)\,.$$

The idea is now the following: the problem of proving self-improving properties for a function $u \in W^{\alpha,2}$ in $\mathbb{R}^n$ is lifted in $\mathbb{R}^{2n}$; we then prove a higher integrability result for U with respect to the measure μ. Essentially, this is a higher integrability result for the dual pair (μ, U). This eventually implies the higher differentiability of u. We indeed have

Theorem 2.2 (Fractional Gehring lemma)**.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ for $\alpha \in (0,1)$, and let (μ, U) be the dual pair generated by u in the sense of (2.4)-(2.5) and Definition 1. Assume that the following reverse Hölder type inequality with tail holds for every $\sigma \in (0,1)$ and for every ball $B \subset \mathbb{R}^n$:*

$$\begin{aligned} \left(\fint_{\mathcal{B}} U^2 d\mu \right)^{1/2} &\leq \frac{c(\sigma)}{\sigma \varepsilon^{1/q-1/2}} \left(\fint_{2\mathcal{B}} U^q d\mu \right)^{1/q} \\ &\quad + \frac{\sigma}{\varepsilon^{1/q-1/2}} \sum_{k=2}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}} U^q \, d\mu \right)^{1/q} , \end{aligned} \tag{2.6}$$

where $q \in (1,2)$ is a fixed exponent and $\mathcal{B} = B \times B$. Then there exists a positive number $\delta \in (0, 1-\alpha)$, depending only on α, ε, q and $c(\sigma)$, such that $U \in L^{2+\delta}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu)$ and $u \in W^{\alpha+\delta, 2+\delta}_{\mathrm{loc}}(\Omega)$. Moreover, the following inequality holds whenever $B \subset \mathbb{R}^n$, for a constant $c \equiv c(n, \alpha, \varepsilon, c(\sigma), q)$:

$$\left(\fint_{\mathcal{B}} U^{2+\delta} \, d\mu \right)^{1/(2+\delta)} \leq c \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}} U^2 \, d\mu \right)^{1/2} . \tag{2.7}$$

The main point of the previous theorem is that we are not asserting the higher integrability of any function U satisfying (2.6). In other words, we are not proving an extension of Gehring's lemma with respect to general measures, something which is on the other hand already available in the literature. Indeed, in the case of Theorem 2.2, the reverse inequality (2.6) is assumed to hold only on diagonal balls

$$\mathcal{B} \equiv \mathcal{B}(x_0, r) = B(x_0, r) \times B(x_0, r) \subset \mathbb{R}^{2n} , \tag{2.8}$$

thereby, we do not assume any information on those zones of $\mathbb{R}^{2n}$ which are far from the diagonal, which is here defined by

$$\mathrm{Diag} := \{(x,x) \in \mathbb{R}^{2n} \, : \, x \in \mathbb{R}^n\} . \tag{2.9}$$

In other words, no reverse inequality holds on non-diagonal balls, or on sets of the type $B(x_0, r) \times B(y_0, r)$. What we are really doing with Theorem 2.2 is asserting the higher integrability of U in $L^{2+\delta}(\mu)$ provided (μ, U) is a dual pair, and this is the crucial point allowing to recover the missing information on non-diagonal balls. Once Theorem 2.2 is proved, we can then get the higher differentiability of functions satisfying a Caccioppoli type inequality with tail, that is Theorem 2.1.

Sketch of the proof of Theorem 2.1. Let us consider, as in Theorem 2.1, a cut-off function $\psi \in C^\infty_{\mathrm{c}}(B(x_0, 3r/4))$ such that $|D\psi| \leq c/r$ and $\psi \equiv 1$ of $B(x_0, r/2)$; from now on we shall denote $B \equiv B(x_0, r)$ and $\mathcal{B} \equiv B \times B$. A direct computation using the definition in (2.5) gives that

$$\mu(\mathcal{B}) = \frac{c(n) r^{n+2\varepsilon}}{\varepsilon} .$$

where $c(n)$ is a constant depending only on n, and this holds whenever $\mathcal{B}$ is a diagonal ball as in (2.8). We then have, using the formula in the last display, that $\psi \equiv 1$ on $B(x_0, r/2)$ and then inequality (2.3), the estimations below

$$\begin{aligned} \frac{r^{2\varepsilon}}{\varepsilon} \fint_{\mathcal{B}/2} U^2 d\mu &\leq \frac{c(n)}{|B|} \int_{\mathcal{B}/2} U^2 \, d\mu \\ &\leq \frac{c}{|B|} \int_{\mathcal{B}/2} \frac{|[u(x)-(u)_B]\psi(x) - [u(y)-(u)_B]\psi(y)|^2}{|x-y|^{2\alpha+2\varepsilon}} \, d\mu \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{c}{r^{2\alpha}} \int_B |u(x)-(u)_B|^2\,dx \\
&\qquad + c\int_{\mathbb{R}^n\setminus B} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy \fint_B |u(x)-(u)_B|\,dx\,. \qquad (2.10)
\end{aligned}
$$

We find an upper bound for the two terms appearing on the right hand side of (2.10). The fractional version of Sobolev embedding theorem provides us with the inequality

$$
r^{-2\alpha}\fint_B |u(x)-(u)_B|^2\,dx \leq \frac{cr^{2\varepsilon}}{\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q} \tag{2.11}
$$

for a constant c depending only on n and α, where

$$
q := \frac{2n+4\varepsilon}{n+2\alpha+2\varepsilon} < 2\,.
$$

For this see [13, Lemma 4.2]. The estimation of the last term on the right hand side of (2.10) is similar, and follows taking into account the geometric decay of the kernel; for this we refer to [13, Proposition 4.2]. The final outcome is the following inequality, which holds whenever $\sigma \in (0,1)$:

$$
\begin{aligned}
&\int_{\mathbb{R}^n\setminus B} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy \fint_B |u(x)-(u)_B|\,dx \\
&\quad \leq \frac{cr^{2\varepsilon}}{\sigma^2\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q} + \frac{\sigma^2 r^{2\varepsilon}}{\varepsilon^{2/q}}\left[\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}\right]^2.
\end{aligned}
$$

Combining the last estimate with (2.11) and (2.10) yields (2.6). We can therefore apply Theorem 2.2 that implies the existence of $\delta > 0$ such that $U \in L^{2+\delta}(\mathcal{B};\mu)$ whenever $\mathcal{B} = B \times B$ and $B \subset \mathbb{R}^n$ is a ball centred at the origin. We conclude that $U \in L^{2+\delta}_{\rm loc}(\mathbb{R}^{2n};\mu)$. We now translate this information in terms of fractional norms of the original function u. In fact this means that, whenever $B \subset \mathbb{R}^n$ is a ball centred at the origin, then we have

$$
\int_{B\times B} U^{2+\delta}\,d\mu = \int_B\int_B \frac{|u(x)-u(y)|^{2+\delta}}{|x-y|^{n+(2+\delta)\alpha+\varepsilon\delta}}\,dx\,dy < \infty\,.
$$

The last integral can be now written as

$$
\int_B\int_B \frac{|u(x)-u(y)|^{2+\delta}}{|x-y|^{n+(2+\delta)[\alpha+\varepsilon\delta/(2+\delta)]}}\,dx\,dy
$$

and this means that $u \in W^{\alpha+\varepsilon\delta/(2+\delta),2+\delta}_{\rm loc}(\mathbb{R}^n)$. We have therefore improved the regularity of u both in the fractional and in the differentiability scale, and Theorem 2.1 follows by renaming the number δ considered in its statement and using the fractional Sobolev embedding theorem. □

3. Nonlocal equations

We now come to nonlocal equations, and report the main facts from [13]. Our results actually rely on techniques which are nonlinear in nature, and therefore they hold for nonlinear equations as well. We shall therefore consider forms of the type

$$
\mathcal{E}_K^{\varphi}(u,\eta) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(u(x)-u(y))[\eta(x)-\eta(y)]K(x,y)\,dx\,dy\,, \tag{3.1}
$$

where the function $\varphi\colon \mathbb{R}\to\mathbb{R}$ satisfies

$$
|\varphi(t)| \leq \Lambda|t|\,, \qquad \varphi(t)t \geq t^2, \qquad \forall\, t\in\mathbb{R}\,, \qquad \Lambda \geq 1\,. \tag{3.2}
$$

The measurable kernel is instead assumed to satisfy

$$\frac{1}{\Lambda|x-y|^{n+2\alpha}} \leq K(x,y) \leq \frac{\Lambda}{|x-y|^{n+2\alpha}} \qquad 0<\alpha<1\,. \tag{3.3}$$

Assumptions (3.1)-(3.3) make the form $\mathcal{E}_K^{\varphi}(u,\eta)$ coercive in $W^{\alpha,2}$, and the related nonlocal equations elliptic. We shall also denote

$$\mathcal{E}_K^{\varphi}(u,\eta) \equiv \mathcal{E}_K(u,\eta) \qquad \text{for} \ \ \varphi(t)=t\,.$$

We are considering solutions $u \in W^{\alpha,2}(\mathbb{R}^n)$ to equations of the type

$$\mathcal{E}_K^{\varphi}(u,\eta) = \mathcal{E}_H(g,\eta) + \int_{\mathbb{R}^n} f\eta\, dx \qquad \forall\ \eta \in C_c^{\infty}(\mathbb{R}^n)\,, \tag{3.4}$$

where, a main point, is that on the right hand side of the previous equation there appears an operator of order $\beta \in (0,1)$ in the sense that the kernel $H(\cdot)$ is assumed to satisfy

$$|H(x,y)| \leq \frac{\Lambda}{|x-y|^{n+2\beta}} \qquad 0<\beta<1\,. \tag{3.5}$$

The family of equations considered in (3.4) allows us to reach the largest nonlocal generalisation of the classical higher integrability results for solutions to elliptic equations. These hold for quasilinear non-homogeneous equations of the type

$$-\operatorname{div} a(x,Du) = -\operatorname{div}(B(x)g) + f\,. \tag{3.6}$$

Note that a main feature of the previous equation is that we have zero and first order operators on the right. In the nonlocal case this is naturally replaced by considering a right hand side that involves the form $\mathcal{E}_H(\cdot)$ with β that can be assumed to be even larger than α, as we shall see in a few moments. The assumptions on the data f and g in the right hand side of (3.4) are now as follows; their optimality will be discussed in Remark 3.3 below. First, we assume that

$$f \in L^{2_*+\delta_0}_{\mathrm{loc}}(\mathbb{R}^n) \tag{3.7}$$

for some $\delta_0>0$. The exponent 2_* is the conjugate of the relevant fractional Sobolev embedding exponent, that is

$$2_* := \frac{2n}{n+2\alpha}\,, \qquad 2^* := \frac{2n}{n-2\alpha}\,, \qquad \frac{1}{2^*}+\frac{1}{2_*}=1\,.$$

The terminology is motivated by the fractional version of the classical Sobolev embedding theorem $W^{\alpha,2} \hookrightarrow L^{2^*}$. Second, we describe the assumptions on g, which are necessarily more involved. We state them first considering the case $2\beta \geq \alpha$. In this case we assume the existence of a positive number $\delta_0>0$ such that

$$g \in W^{2\beta-\alpha+\delta_0,2}(\mathbb{R}^n)\,. \tag{3.8}$$

Needless to say, we also assume that $2\beta-\alpha+\delta_0 \in (0,1)$ to give $[g]_{2\beta-\alpha+\delta_0,2}$ sense according to the definition in (2.1); this in particular implies that $\beta<(1+\alpha)/2$. In the case $0<2\beta<\alpha$ we instead do not consider any differentiability on g, but only integrability:

$$g \in L^{p_0+\delta_0}(\mathbb{R}^n)\,, \qquad p_0 := \frac{2n}{n+2(\alpha-2\beta)}\,. \tag{3.9}$$

The *main result* of [13] is then

Theorem 3.1. *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (3.4) under the assumptions (3.2)-(3.3) and (3.5)-(3.9). Then there exists a positive number $\delta \in (0,1-\alpha)$, depending only on $n,\alpha,\Lambda,\beta,\delta_0$, but otherwise independent of the solution u and of the kernels $K(\cdot),H(\cdot)$, such that*

$$u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)\,. \tag{3.10}$$

An immediate corollary follows when considering the case $H(\cdot) = K(\cdot)$ and $\varphi(t) = t$, thereby considering the linear equation

$$\mathcal{E}_K(u,\eta) = \mathcal{E}_K(g,\eta) + \int_{\mathbb{R}^n} f\eta\, dx \qquad \forall\, \eta \in C_c^\infty(\mathbb{R}^n)\,.$$

In this case Theorem 3.1 allows to get the following particularly neat result:

$$f \in L^{2_*+\delta_0}_{\mathrm{loc}}(\mathbb{R}^n), g \in W^{\alpha+\delta_0,2}(\mathbb{R}^n) \implies u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)\,, \tag{3.11}$$

for some positive $\delta \in (0,\delta_0)$.

The proof of Theorem 3.1 is based on an extended version of Theorem 2.2, that takes into account additional terms stemming from the right hand side of the equation in (3.4). More precisely, the starting point of the proof of Theorem 3.1 is the following Caccioppoli type inequality:

$$\begin{aligned}
&\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\\
&\quad\leq \frac{c}{r^{2\alpha}}\int_B |u(x)|^2\,dx + c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy\int_B |u(x)|\,dx\\
&\quad\quad + cr^{n+2\alpha}\left(\fint_B |f(x)|^{2_*}\,dx\right)^{2/2_*}\\
&\quad\quad + cr^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2\,.
\end{aligned} \tag{3.12}$$

This holds whenever $B(x_0,r) \subset \mathbb{R}^n$ is a ball, with $\psi \in C_c^\infty(B(x_0,3r/4))$ being a cut-off function such that $|D\psi| \leq c/r$; the constant c depends only on n,Λ,α,β and the exponent p depends on n,α,β,δ_0, while $\alpha < \gamma < 2\beta-\alpha$. Note that (3.12) reduces to (2.3), when $f = g = 0$. As in the sketch of the proof of Theorem 2.1, from (2.3) it follows a reverse inequality of the type in (2.6), but with additional terms. From this (3.10) follows again by an extension of Theorem 2.2, that takes into account reverse inequalities with additional terms (in this case, those coming from f and g). In this note we prefer giving a description of the simpler case of Theorem 2.2 for the sake of brevity and clarity of exposition.

Remark 3.1 (Peculiarity of the nonlocal case)**.** At first sight, the natural nonlocal analog of the results valid for local equations would be to prove that $u \in W^{\alpha,2+\delta}_{\mathrm{loc}}$, for some $\delta > 0$. Therefore, Theorem 3.1 reveals a new, unexpected property of solutions to nonlocal equations that has in fact no analog in the local case. Indeed, in order to get some higher gradient differentiability of the $W^{1,2}$-solutions to (1.1), it is then necessary to assume that the entries of the matrix $A(\cdot)$ belong themselves to a fractional Sobolev space, as shown for instance in [12, 15]. To see this already in the one dimensional case $n = 1$, it is sufficient to consider the following equation:

$$\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) = 0\,, \qquad \frac{1}{\Lambda} \leq a(x) \leq \Lambda\,, \tag{3.13}$$

and to note that

$$x \to \int_0^x \frac{dt}{a(t)}$$

is a solution with $a(\cdot)$ being any measurable function satisfying nothing but the inequalities in (3.13). It is then easy to build similar multidimensional examples.

Remark 3.2 (Previous results)**.** We mention a very recent and interesting paper of Bass & Ren [1] who considered the function (called Marcinkiewicz integral)

$$\Gamma(x) := \left(\int_{\mathbb{R}^n}\frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\,dy\right)^{1/2}\,,$$

and proved that $\Gamma \in L^{2(1+\delta)}(\mathbb{R}^n)$ for some positive δ depending only on n, α, Λ and δ_0. The equations considered in [1] are of the type

$$\mathcal{E}_K(u,\eta) = \int_{\mathbb{R}^n} f\eta\, dx \qquad \forall\, \eta \in C_c^\infty(\mathbb{R}^n)$$

with $f \in L^{2+\delta_0}$, for some $\delta_0 > 0$. The function $\Gamma(\cdot)$ can be interpreted, dimensionally speaking, as a fractional gradient of u of order α. Once this result is achieved, the higher integrability as stated in (2.2) then follows via a deep characterisation characterisation of fractional Sobolev spaces via Bessel potential spaces that rests on Littlewood-Paley theory ([4, 16]).

Remark 3.3 (Optimality of the assumptions on f and g)**.** The assumptions on f and g considered in (3.7)-(3.9) are the natural counterparts of those usually considered for the classical case (3.6). Their optimality can be checked by using a few formal arguments applied on the linear model equation

$$(-\triangle)^\alpha u = (-\triangle)^\beta g + f \tag{3.14}$$

that can be indeed treated by different means via Fourier analysis or Riesz potentials. Needless to say, the case of (3.14) belongs to the family described by (3.4), as can be seen by taking $K(x,y) = |x-y|^{-n-2\alpha}$ and $H(x,y) = |x-y|^{-n-2\beta}$. The following arguments will be purely formal; they are only aimed at checking that the exponents considered for f and g in (3.7)-(3.9) are the right ones. First of all, since here we are dealing with self-improving properties, and since all the numbers δ_0 and δ are bound to be small, then with no loss of generality we will check the optimality of the exponents for f and g in the "limit case" $\delta_0 = \delta = 0$. We start by f, therefore considering the equation $(-\triangle)^\alpha u = f$, for simplicity when $2\alpha < 1$. In this case we have that $f \in L^q$ implies $u \in W^{2\alpha,q}$. Since on the other hand we are dealing with equations with measurable coefficients, $W^{2\alpha,q}$-regularity is not achievable, and we look for the corresponding $W^{\alpha,2}$-regularity. Therefore we recall the imbedding

$$W^{2\alpha,q} \hookrightarrow W^{\alpha,2} \quad \text{if} \quad 2\alpha - \frac{n}{q} = \alpha - \frac{n}{2}\,.$$

This in fact gives $q = 2_*$, that is (3.7) for $\delta_0 = 0$. As for g, we now consider the equation $(-\triangle)^\alpha u = (-\triangle)^\beta g$. Let us first observe that in the case $\alpha = \beta$ it is obvious to take $g \in W^{\alpha,2}$, as in (3.11) with $\delta_0 = 0$. In the case $2\beta > \alpha$ let us formally write $\partial^\alpha u \approx \triangle^{\beta-\alpha/2} g \approx \partial^{2\beta-\alpha} g \in L^2$. Therefore, in order to obtain that $u \in W^{\alpha,2}$ it remains to require that $g \in W^{2\beta-\alpha,2}$. Finally, in the case $2\beta < \alpha$, we use the same formal argument, interpreting $W^{2\beta-\alpha,2}$ as the dual of $W^{\alpha-2\beta,2}$. The fractional Sobolev embedding theorem then gives

$$W^{\alpha-2\beta,2} \hookrightarrow L^{\frac{2n}{n-2(\alpha-2\beta)}}\,.$$

But now

$$\left(L^{\frac{2n}{n-2(\alpha-2\beta)}}\right)' = L^{\frac{2n}{n+2(\alpha-2\beta)}}\,,$$

and therefore we conclude that a sufficient condition for g to belong to the dual of $W^{\alpha-2\beta,2}$ is

$$g \in L^{\frac{2n}{n+2(\alpha-2\beta)}}\,,$$

that is (3.9) with $\delta_0 = 0$.

4. Ideas from the proof of Theorem 2.2

The proof of Theorem 2.2 is rather complex, and we can only try to give a brief sketch of the arguments, referring the reader to [13] for the rest. The whole issue can be reduced to prove the following level set inequality of the function U:

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \le c\lambda^{2-q}\int_{\mathcal{B}(x_0,s)\cap\{U>\lambda\}} U^q d\mu + \frac{c\lambda_0}{(s-t)^{n+2\varepsilon}}\,, \tag{4.1}$$

which is bound to hold whenever $\mathcal{B}(x_0,\varrho_0)\subset\mathcal{B}(x_0,t)\subset\mathcal{B}(x_0,s)\subset\mathcal{B}(x_0,3\varrho_0/2)$ with $r<t<s<3r/2$, and for those λ satisfying

$$\lambda_0 := \frac{1}{(s-t)^{n+2\varepsilon}}\sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^2\,d\mu\right)^{1/2} \lesssim \lambda\,.$$

The assertion, that is (2.7), then follows using truncation arguments, Cavalieri's principle, and an iteration lemma. We therefore briefly discuss the proof of (4.1). Since the main information at our disposal, that is the reverse inequality (2.6), is available only on diagonal balls as in (2.8), and not on every ball in $\mathbb{R}^{2n}$, we start the estimation of the integral on the right hand side of (4.1) by splitting

$$\begin{aligned}\int_{\{U>\lambda\}} U^2\,d\mu \quad &= \quad \int_{\{U>\lambda\}\cap\text{“zone close to the diagonal”}} U^2\,d\mu\\ &\qquad + \int_{\{U>\lambda\}\cap\text{“zone far from the diagonal”}} U^2\,d\mu\,.\end{aligned} \tag{4.2}$$

This actually means that we are going to use two different exit time arguments to build two Calderón-Zygmund coverings of the level set $\mathcal{B}(x_0,t)\cap\{U>\lambda\}$; the first is aimed to cover the zone close to the diagonal, while the second to cover the zone far from the diagonal. The diagonal covering is obtained via a direct exit time argument based on Vitali's covering lemma, and is made of a countable family of diagonal balls $\{\mathcal{B}_j\}$ of the type in (2.8), on which it happens that

$$⨍_{\mathcal{B}_j} U^2\,d\mu \approx \lambda^2\,.$$

These balls are aimed at covering that part of the level set surrounding the diagonal. The second covering is instead obtained directly using the classical Calderón-Zygmund covering argument and is made of a countable family dyadic cubes $\{\mathcal{K}\}$, for it happens that

$$⨍_{\mathcal{K}} U^2 d\mu \approx M\lambda^2$$

and

$$U \le M\lambda \qquad \text{holds a.e. in} \quad \mathcal{B}(x_0,t)\setminus\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda}\mathcal{K}\,.$$

The constant $M\ge 1$ is chosen large enough to make, in a sense, the cubes $\mathcal{K}$ smaller than the balls from the family $\{\mathcal{B}_j\}$.

We then proceed in sorting the cubes from the non-diagonal covering in two classes: those that are close to the diagonal Diag (defined in (2.9)), and those cubes which are suitably far from the diagonal. The cubes that are close enough to the diagonal can be covered by the diagonal balls $\{\mathcal{B}_j\}$ coming from the diagonal covering. The other ones need a different treatment. How to decide if a cube $\mathcal{K}$ is far from the diagonal? For us this means that, with $l(\mathcal{K})$ denoting the side length of the cube $\mathcal{K}$, it happens that

$$\operatorname{dist}(\mathrm{Diag},\mathcal{K}) \ge l(\mathcal{K})\,. \tag{4.3}$$

This condition relates in some sense the size of the exit time cube $\mathcal{K}$ with the distance to the diagonal and, ultimately, to the size of the kernel $K(\cdot)$ on the cubes $\mathcal{K}$. The analysis then proceeds in two different stages. In a first one, we use inequality (2.6) on the diagonal balls $\{\mathcal{B}_j\}$; this, together with a proper use of the exit time condition to treat the tail terms, allows to deal with the first integral appearing on the right hand side of (4.2).

In a second stage, we deal with the cubes $\{\mathcal{K}\}$ which are far from the diagonal, and that therefore satisfy (4.3). The lack of reverse inequalities on the cubes $\mathcal{K}$ is compensated by the fact that, far from the diagonal, a different type of reverse inequality automatically hold. This inequality reads as

$$
\begin{aligned}
\left(\fint_{\mathcal{K}} U^2 \, d\mu\right)^{1/2} \leq \; & c\left(\fint_{\mathcal{K}} U^q \, d\mu\right)^{1/q} \\
& + \frac{c}{\varepsilon^{1/q}}\left(\frac{l(\mathcal{K})}{\operatorname{dist}(\mathrm{Diag},\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{K_1\times K_1} U^q \, d\mu\right)^{1/q} \\
& + \frac{c}{\varepsilon^{1/q}}\left(\frac{l(\mathcal{K})}{\operatorname{dist}(\mathrm{Diag},\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{K_2\times K_2} U^q \, d\mu\right)^{1/q} ,
\end{aligned} \tag{4.4}
$$

where $\mathcal{K}$, being a dyadic cube in $\mathbb{R}^{2n}$, admits a decomposition $\mathcal{K} = K_1 \times K_2$ and K_1, K_2 are themselves dyadic cubes from $\mathbb{R}^n$ with the same side lengths. The constant c depends only on n, α and the main point is that the exponent q is such that $q < 2$.

Inequality (4.4) is bound to replace (2.6) far from the diagonal, but it unfortunately involves two remainder terms that prevents it to be a real reverse Hölder inequality. These terms are those in the last two lines of (4.4) and involve integrals on additional dyadic cubes, that are $K_1 \times K_1$ and $K_2 \times K_2$. The main problem is now that these cubes have not been selected via an exit time argument and therefore there is no control on the average of U over them. In turn, this does not allow to employ the usual covering arguments. Instead, we make use of very delicate combinatorial arguments that at the end will work via a subtle combination of geometric information coming from the sorting of the cubes, the size of the measure when certain distance conditions from the diagonal are considered, and finally the size of the coefficients appearing on the right hand side of (4.4). Once this is achieved both integrals appearing on the right hand side of (4.4) can be estimated, and this opens the way to the proof of (4.1). The details are at this point extremely technical, and we once again refer to [13].

Acknowledgment. This paper reports the content of a talk given by the second-named author at the Accademia dei Lincei on November 26, 2013.

References

[1] Bass R. F. & Ren H.: Meyers inequality and strong stability for stable-like operators. *J. Funct. Anal.* 265 (2013), 28–48.

[2] Bojarski B. & Iwaniec T.: Analytical foundations of the theory of quasiconformal mappings in $\mathbb{R}^n$. *Ann. Acad. Sci. Fenn. Ser. A I Math.* 8 (1983), 257–324.

[3] Caffarelli L. & Chan C. H. & Vasseur A.: Regularity theory for parabolic nonlinear integral operators. *J. Amer. Math. Soc.* 24 (2011), 849–869.

[4] Dorronsoro J.R.: A characterization of potential spaces. *Proc. Amer. Math. Soc.* 95 (1985), 21–31.

[5] Elcrat A. & Meyers N.: Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions. *Duke Math. J.* 42 (1975), 121–136.

[6] Fusco N. & Sbordone C.: Higher integrability of the gradient of minimizers of functionals with nonstandard growth conditions. *Comm. Pure Appl. Math.* 43 (1990), 673–683.

[7] Gehring F. W.: The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* 130 (1973), 265–277.

[8] Giaquinta M. & Modica G.: Regularity results for some classes of higher order nonlinear elliptic systems. *J. reine angew. Math. (Crelles J.)* 311/312 (1979), 145–169.
[9] Iwaniec T.: The Gehring lemma. *Quasiconformal mappings and analysis (Ann Arbor, MI, 1995)*, 181–204, Springer, New York, 1998.
[10] Kassmann M.: A priori estimates for integro-differential operators with measurable kernels. *Calc. Var. PDE* 34 (2009), 1–21.
[11] Kinnunen J. & Lewis J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* 102 (2000), 253–271.
[12] Kuusi T. & Mingione G.: Universal potentials estimates. *J. Funct. Anal.* 262 (2012), 4205–4269.
[13] Kuusi T. & Mingione G. & Sire Y.: Nonlocal self-improving properties. *Analysis & PDE*, to appear.
[14] Meyers N. G.: An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa C. Sci. (3)* 17 (1963), 189–206.
[15] Mingione G.: The singular set of solutions to non-differentiable elliptic systems. *Arch. Ration. Mech. Anal.* 166 (2003), 287–301.
[16] Stein E. M.: The characterization of functions arising as potentials. *Bull. Amer. Math. Soc.* 67 (1961), 102–104.

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

Giuseppe Mingione, Dipartimento di Matematica, Università di Parma, Parco Area delle Scienze 53/a, Campus, 43100 Parma, Italy
E-mail address: giuseppe.mingione@unipr.it.

Yannick Sire, Université Aix-Marseille and LATP-CMI,, 9, rue F. Joliot Curie, 13453 Marseille Cedex 13, France , France
E-mail address: sire@cmi.univ-mrs.fr