

PARTIAL REGULARITY FOR VECTORIAL LOCAL-NONLOCAL PROBLEMS

CARLO ALBERTO ANTONINI  AND PAOLO BARONI 

ABSTRACT. We establish sharp interior regularity for solutions to vectorial, nonlinear local-nonlocal systems.

1. INTRODUCTION

Operators that couple a local and a nonlocal diffusion term naturally arise in the modeling of processes combining continuous and jump-type behavior [33], and in recent years they have been studied with increasing attention, both for their analytical structure and for this modeling relevance. The prototype of this class is the case where a local Laplacian diffusion is coupled with a fractional Laplacian diffusion

$$-\Delta u + (-\Delta)^s u.$$

For this class of operators, several aspects of the elliptic and parabolic theory have by now been thoroughly investigated, including Green function estimates, boundary Harnack principles, maximum principles, and regularity properties, both in the elliptic and in the parabolic setting; since the literature on the linear theory is by now vast, we only refer to [16, 17, 9, 10, 22, 45] and the references therein for a detailed account. More recently, attention has turned to nonlinear counterparts: first to models preserving the linear local structure while allowing more general stochastic nonlocal terms [11], and then, more substantially, to mixed quasilinear operators modeled on

$$-\Delta_p u + (-\Delta_q)^s u, \quad p, q \in (1, \infty), \quad s \in (0, 1);$$

a number of existence and qualitative results have been obtained, covering both variational and nonvariational formulations; see, for instance, [4, 5, 12, 18, 21, 29, 31, 32]. The regularity theory for these operators is shaped by the balance between the local and nonlocal components, and this balance in turn depends on the relation between p , q and s ; the present paper adds to this general picture by showing how this interplay manifests itself for vector-valued solutions with quadratic growth. In the scalar case, this balance has already been made precise: De Filippis & Mingione indeed established in [23] the interior Hölder continuity of Du via perturbation methods under the structural condition $p > sq$, as this assumption ensures that at small scales the local operator $-\Delta_p u$ dominates the nonlocal one $(-\Delta_q)^s u$, so that the regularizing effect of the equation is essentially governed by its local part. Consequently, solutions inherit the same type of regularity of solutions to purely local operators, namely the $C^{1,\alpha}$ -regularity for some non-explicit $\alpha \in (0, 1)$, which does not exceed the optimal Hölder exponent of the gradient of p -harmonic functions [3, 6, 28]. The gradient Hölder regularity in the complementary regime $p \leq sq$ was recently treated by Biswas and Topp [13] through viscosity methods.

We initiate, in this paper, the study of quasilinear vectorial mixed operators by treating the quadratic growth case, namely $p = q = 2$, allowing for a class of local and nonlocal terms more general than the model operators, see (1.1)–(1.2) and the related assumptions (1.4) and (1.5). Although this is only the first step toward the more general theory we aim to develop, it already exhibits the essential difficulties of the vectorial local-nonlocal interaction, whose treatment requires a substantially new combination of classic local partial regularity methods, developed for the vectorial setting, and nonlocal tail estimates, adapted here to interact with them. We show that regularity of u and of its gradient Du hold on the very same regular set, whose singular complement has Hausdorff dimension explicitly bounded, and that on this set a full scale of results is available, improving as the assumptions on the local term are strengthened: from BMO and VMO regularity of u , up to its continuity and Hölder continuity, and finally to Lipschitz continuity, continuity and Hölder continuity of the gradient itself; none of these improvements requires any additional regularity on the nonlocal term, which is allowed to remain as irregular and non-coercive as in our basic assumptions throughout.

Our main novelties can be summarized as follows:

- we work in the vectorial setting and study the partial regularity of solutions (Theorems 2.1 to 2.8), obtaining everywhere regularity under additional structural assumptions on the local term (Theorem 2.9);
- we allow for very general nonlocal operators, possibly irregular and non-coercive (see (1.4)–(1.5));
- we obtain sharp regularity results in terms of the right-hand side of the equation, with explicit Hölder exponents together with quantitative local estimates that also keep track of the nonlocal tail of the solution (for instance, see (2.15) and (2.24) and the related regularity results).

The present paper provides, to the best of our knowledge, the first systematic development of the vectorial regularity theory for mixed local-nonlocal systems. Its closest predecessor is the recent vectorial theory for purely nonlocal systems developed by De Filippis, Mingione & Nowak in [25], where classical tools from the local partial regularity theory – excess decay, tail estimates, blow-up procedures and harmonic approximation – were implemented for the first time in a genuinely nonlocal vectorial setting. Here we show that, in the mixed regime, the local component can be exploited as the leading operator, while the nonlocal part can be treated as a lower-order perturbation even under minimal assumptions, namely without regularity or coercivity on the integrand, so that under appropriate assumptions on the local part solutions retain the highest regularity dictated by the local theory. This dominance, however, has a natural limit at the level of the gradient: under the sole structural assumptions on the nonlocal term, the Hölder exponent of Du cannot in general exceed a threshold dictated by s , reflecting the fact that differentiating the equation makes the lack of regularity of the nonlocal term visible; remarkably, this threshold turns out to be sharp, matching the optimal boundary regularity known for nonlocal equations. When the nonlocal term is assumed to have the natural structure of the fractional Laplacian, this limitation disappears entirely, and the full, sharp Hölder regularity of the gradient is recovered.

The precise setting we work in is the following: let $n, N \in \mathbb{N}$, with $n \geq 2$ and $N \geq 1$, let $\Omega \subset \mathbb{R}^n$ be open, let $u : \mathbb{R}^n \rightarrow \mathbb{R}^N$, and let $f : \Omega \rightarrow \mathbb{R}^N$ be a measurable function. We deal with systems of the form

$$(1.1) \quad Q_A^l u + Q_\Phi^{nl} u = f \quad \text{in } \Omega,$$

where Q_A^l is the local operator given by

$$Q_A^l u(x) = -\operatorname{div}(A(x, u)Du),$$

and Q_Φ^{nl} is the nonlocal operator defined by

$$(1.2) \quad Q_\Phi^{nl} u(x) := 2 \operatorname{P.V.} \int_{\mathbb{R}^n} \Phi(x, y, u(x), u(y), u(x) - u(y)) \frac{dy}{|x - y|^{n+2s}}, \quad s \in (0, 1);$$

here P.V. denotes the Cauchy principal value.

The coefficients $A(x, z) = \{A_{ij}^{\alpha\beta}(x, z)\}_{i,j=1,\dots,n}^{\alpha,\beta=1,\dots,N} : \Omega \times \mathbb{R}^N \rightarrow \mathbb{R}^{nN \times nN}$ are Carathéodory functions satisfying the Legendre and the boundedness conditions

$$(1.3) \quad \langle A(x, z)\xi, \xi \rangle \geq \Lambda^{-1} |\xi|^2, \quad |A(x, z)| \leq \Lambda,$$

for almost all $x \in \Omega$, every $z \in \mathbb{R}^N$ and $\xi = \{\xi_i^\alpha\}_{i=1,\dots,n}^{\alpha=1,\dots,N} \in \mathbb{R}^{nN}$, for some constant $\Lambda \geq 1$. We refer to Paragraph 3.1 for the relevant notational conventions.

As for the nonlocal term $\Phi : \mathbb{R}^{2n+3N} \rightarrow \mathbb{R}^N$, $\Phi(x, y, u, v, t) = \{\Phi^\alpha(x, y, u, v, t)\}_{\alpha=1,\dots,N}$, we assume that it has linear growth in the t -variable

$$(1.4) \quad |\Phi(x, y, u, v, t)| \leq \nu |t|$$

for some $\nu > 0$, and it satisfies the symmetry properties

$$(1.5) \quad \begin{cases} \Phi(x, y, u, v, t) = \Phi(y, x, v, u, t) \\ \Phi(x, y, u, v, -t) = -\Phi(x, y, u, v, t) \end{cases},$$

for almost every $x, y \in \mathbb{R}^n$, all $u, v \in \mathbb{R}^N$ and $t \in \mathbb{R}^N$; no additional regularity is required on the nonlocal term.

We consider weak solutions to equation (1.1), i.e., functions $u \in W_{\operatorname{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ that satisfy the weak formulation of (1.1):

$$(1.6) \quad \int_{\Omega} \langle A(x, u)Du, D\varphi \rangle dx + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \Phi(x, y, u(x), u(y), u(x) - u(y)) \cdot (\varphi(x) - \varphi(y)) \frac{dx dy}{|x - y|^{n+2s}}$$

$$= \int_{\Omega} f \cdot \varphi \, dx$$

for all smooth test functions with compact support in Ω , $\varphi \in C_c^\infty(\Omega; \mathbb{R}^N)$. Here, we employ the notation L_{2s}^1 for the tail space

$$L_{2s}^1 := \left\{ w \in L_{\text{loc}}^1(\mathbb{R}^n; \mathbb{R}^N) : \int_{\mathbb{R}^n} \frac{|w(x)|}{1 + |x|^{n+2s}} \, dx < \infty \right\}.$$

To better illustrate the main results of the paper, we state the following two model theorems, the first concerning the Hölder continuity of solutions.

Theorem 1.1 (Partial Hölder continuity). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (1.1), with (1.3)–(1.5) in force. Suppose that $A(\cdot, \cdot)$ satisfies the continuity assumption*

$$(1.7) \quad |A(x, u) - A(y, v)| \leq \Lambda \omega(|x - y| + |u - v|) \quad \text{for all } x, y \in \Omega \text{ and } u, v \in \mathbb{R}^N,$$

for some modulus of continuity $\omega : [0, \infty) \rightarrow [0, 1]$. Then there exists an open set $\Omega_u \subseteq \Omega$ with Hausdorff dimension $\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) \leq n - 2 - \zeta$ for some $\zeta > 0$ if $n \geq 3$, and $\Omega_u = \Omega$ if $n = 2$ such that, if

$$f \in L_{\text{loc}}^{n/(2-\beta_0)}(\Omega; \mathbb{R}^N) \quad \text{for some } \beta_0 \in (0, 1),$$

then

$$u \in C_{\text{loc}}^{0, \beta_0}(\Omega_u; \mathbb{R}^N).$$

If in addition $A(x, u) \equiv A(x)$ is independent of u , then full regularity holds, i.e., $\Omega_u = \Omega$.

The next model theorem addresses the quantitative Hölder continuity of the gradient.

Theorem 1.2 (Partial gradient Hölder continuity). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (1.1), with (1.3)–(1.5) in force. Let $\beta \in (0, 1)$, and suppose that $A(\cdot, \cdot)$ fulfils (1.7) with $\omega(r) = r^\beta$. If*

$$f \in L_{\text{loc}}^{n/(1-\beta)}(\Omega; \mathbb{R}^N),$$

then

$$u \in C_{\text{loc}}^{1, \min\{\beta, 2(1-s)\}}(\Omega_u; \mathbb{R}^N),$$

with Ω_u as in Theorem 1.1. In particular, if $A(x, u) \equiv A(x)$, then full gradient regularity holds. If in addition the nonlocal term is of the form

$$(1.8) \quad \Phi(x, y, u, v, t) = a(x - y)t$$

for some matrix-valued function $a : \mathbb{R}^n \rightarrow \mathbb{R}^{N \times N}$ such that

$$(1.9) \quad a(z) = a(-z) \quad \text{and} \quad |a(z)| \leq \nu, \quad \text{for all } z \in \mathbb{R}^n,$$

then

$$(1.10) \quad u \in C_{\text{loc}}^{1, \beta}(\Omega_u; \mathbb{R}^N).$$

We refer to Section 2 for the precise statements of the results, where we also provide finer quantitative estimates. We now briefly comment on these results. Theorem 1.1 provides a quantitative Hölder continuity result for solutions, in analogy with the purely local equation [34]:

$$-\operatorname{div}(A(x, u) Du) = f \in L_{\text{loc}}^{n/(2-\beta_0)}(\Omega; \mathbb{R}^N) \implies u \in C_{\text{loc}}^{0, \beta_0}(\Omega_u; \mathbb{R}^N).$$

This result is essentially sharp, as can be seen by a simple scaling argument¹. This behavior is not surprising since, as already mentioned, the local operator Q_A^1 represents the leading term in (1.1), and therefore it is this component that ultimately dictates the regularity of solutions.

Another key feature of our approach, which relies precisely on the dominance of the local operator, is that we do not impose regularity or coercivity assumptions on the nonlocal term Φ . In particular, we do not require the usual coercivity condition

$$\Phi(x, y, u, v, t) \cdot t \geq c|t|^2, \quad c > 0;$$

this is particularly relevant in applications, where the nonlocal term may exhibit irregular or poorly structured behavior, see for instance [33].

¹See Theorem 2.4, where a slightly weaker integrability condition on f is assumed in terms of a Marcinkiewicz space.

We also recall that everywhere regularity may fail in the vectorial setting, even when the coefficients $u \mapsto A(x, u)$ are analytic [34, Chapter II.3]. On the other hand, when the u -dependence on the coefficients is dropped, everywhere regularity is then recovered.

Turning to gradient regularity, Theorem 1.2 establishes the Hölder continuity of Du with an explicit Hölder exponent dictated by the integrability of f . However, under such low assumptions on the nonlocal term, the exponent cannot in general exceed the threshold $C^{1,2(1-s)}$ in the regime $s > 1/2$ for such general nonlocal terms Q_Φ^{nl} . This reflects the fact that, in the mildly local regime $s > 1/2$, the lack of regularity of the nonlocal term influences the regularity of u at the level of the gradient, in view of the fact that the proof implicitly involves differentiating equation (1.1). We also emphasize that the threshold $2(1-s)$ is consistent with the optimal regularity expected in this setting. In particular, it coincides with the sharp boundary regularity exponent established for solutions to nonlocal equations, see [1], and this highlights the role of the nonlocal operator in limiting the attainable gradient regularity, even when the equation exhibits a predominantly local character.

On the other hand, when we impose structural assumptions on Φ , namely when it has the form described in (1.8)–(1.9), this limitation on the exponent disappears in (1.10); thus, we recover the full, sharp Hölder continuity of the gradient

$$f \in L_{\text{loc}}^{n/(1-\beta)}(\Omega; \mathbb{R}^N) \implies u \in C_{\text{loc}}^{1,\beta}(\Omega_u; \mathbb{R}^N) \quad \text{for all } \beta \in (0, 1).$$

The prototypical example of Φ fulfilling (1.8)–(1.9) is the fractional Laplacian $(-\Delta)^s u$, for which $\Phi(x, y, u, v, t) = \Phi(t) = t$.

Outline of the proof. The proof proceeds in several steps and follows, in principle, a perturbative approach as in [4, 23], but carried out in a finer and more quantitative form, making the interplay between the local and nonlocal operators more explicit.

As in the classical regularity theory for purely local systems, the first step consists in establishing higher integrability of the gradient, namely $Du \in L_{\text{loc}}^{2+\zeta}$ in Proposition 4.4. The delicate point here is to keep explicit track of the dependence on the growth factor ν associated with the nonlocal term in (1.4), see Remark 5.1.

The next step is to prove an approximate harmonicity property for solutions—see Proposition 5.3. The key observation is that if u solves (1.1) in the ball $B_\varrho(x_0) \Subset \Omega$, then the blow-up function $u_\varrho(x) := u(x_0 + \varrho x)$ solves a similar equation in which the nonlocal term is multiplied by the factor $\varrho^{2(1-s)}$, see equation (5.7). More explicitly, in the case $f = 0$, the function u_ϱ is solution of

$$Q_{A_\varrho}^1 u_\varrho + \varrho^{2(1-s)} Q_{\Phi_\varrho}^{\text{nl}} u_\varrho = 0 \quad \text{in } B_1$$

with coefficient matrix A_ϱ and Φ_ϱ still satisfying the structural assumptions (1.3) to (1.5). Consequently, by choosing the radius ϱ sufficiently small, the nonlocal contribution in this equation can be made arbitrarily small, see Lemma 5.2.

At the same time, the higher integrability result obtained in the first step guarantees that the $L^{2+\zeta}$ -norm of Du_ϱ remains uniformly controlled for any small radii ϱ —this is the reason of the care in tracking the dependence on the size of the nonlocal term ν . This allows us to apply the approximate harmonicity Lemma 3.5, see also (5.25); hence u is quantitatively close to a harmonic function, which in turn leads to a quantitative decay estimate for the excess functional—see Paragraph 5.2 and Proposition 5.6.

From this point on, the proof of the regularity of u follows a more standard route. In particular, we exploit the recently developed machinery of potential theory to control the source term f , together with suitable decay estimates for the nonlocal tail over shrinking balls, see Lemma 3.6 and Lemma 7.6.

Plan of the paper. In Section 2 we state the main results of the paper in full detail, while Section 3 provides the notation and some preliminary results. In Section 4 we prove the Caccioppoli inequality and the higher integrability of solutions. Section 5 is devoted to the proof of approximate harmonicity of solutions and the excess decay estimate, which will be the starting point of the proofs for the main results. In Section 6 we characterize the regular set of solutions and we provide the proof of the results involving the regularity of u , while in Section 7 we study the regularity of the gradient Du . Finally, in Section 8 we discuss in detail the everywhere regularity of solutions in the case of u -independent coefficients.

2. MAIN RESULTS

In this section we state in full detail the results of the paper. Unless otherwise stated, we shall assume that

$$(2.1) \quad f \in L_{\text{loc}}^\chi(\Omega; \mathbb{R}^N) \quad \text{for some } \chi > 2_*,$$

where we set

$$(2.2) \quad 2_* = \begin{cases} \frac{2n}{n+2} & n > 2 \\ \text{any number in } (1, 2) & n = 2 \end{cases}, \quad 2^* = \begin{cases} \frac{2n}{n-2} & n > 2 \\ \text{any number in } (2, +\infty) & n = 2 \end{cases}.$$

The exponent χ can be auxiliary: in the results where f is assumed to belong to a Marcinkiewicz space, χ is chosen below the corresponding Marcinkiewicz exponent, as explained after (3.5). We denote by **data** the set of parameters

$$(2.3) \quad \mathbf{data} = \{n, N, s, \Lambda, \chi\}.$$

Also, we shall always assume that $A(x, \cdot)$ is uniformly continuous; namely, there exists a modulus of continuity $\omega_u(\cdot)$ such that

$$(2.4) \quad |A(x, u) - A(x, v)| \leq \Lambda \omega_u(|u - v|) \quad \text{for almost all } x \in \Omega \text{ and } u, v \in \mathbb{R}^N.$$

Our first two main results are concerned with BMO and VMO regularity of solutions to (1.6). These are stated in terms of a smallness assumption on the x -mean oscillation of $A(\cdot, \cdot)$. Specifically, for any ball $B_r(x_0) \subseteq \Omega$, we define the averaged vector field

$$(2.5) \quad (A)_{B_r(x_0)}(z) := \oint_{B_r(x_0)} A(x, z) dx;$$

then, for every $x_0 \in \Omega$ and for some $r_0 > 0$, we define

$$(2.6) \quad E_{r_0}(A; x_0) = \sup_{\substack{\varrho \in (0, r_0]: B_\varrho(x_0) \subseteq \Omega, \\ z \in \mathbb{R}^N, z \neq 0}} \oint_{B_\varrho(x_0)} |A(x, z) - (A)_{B_\varrho(x_0)}(z)| dx,$$

the x -mean oscillation of $A(\cdot, \cdot)$.

Theorem 2.1 (Partial BMO Regularity). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1). Suppose that $A(\cdot, \cdot)$ satisfies (1.3), (2.4), and that Φ fulfills (1.4), (1.5). Then there exist constants*

$$\varepsilon_b = \varepsilon_b(\mathbf{data}, \omega_u(\cdot)) \in (0, 1), \quad \delta_0 = \delta_0(\mathbf{data}) \in (0, 1),$$

such that the following implication holds: if for every compact set $K \Subset \Omega$ there exist radii $r_A, r_f \in (0, \text{dist}(K, \partial\Omega))$ such that

$$(2.7) \quad \sup_{x_0 \in K} E_{r_A}(A; x_0) < \delta_0,$$

and

$$(2.8) \quad \sup_{\substack{x_0 \in K, \\ 0 < t \leq r_f}} t^2 \left(\oint_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} < \varepsilon_b,$$

where χ is as in (2.1), then there exists an open set $\Omega_u \subseteq \Omega$ such that $\Omega_u = \Omega$ if $n = 2$, while for $n \geq 3$,

$$\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) \leq n - 2 - \zeta$$

with $\zeta = \zeta(\mathbf{data}, \max\{\nu, 1\}) \in (0, 1)$, and

$$(2.9) \quad u \in BMO_{\text{loc}}(\Omega_u; \mathbb{R}^N).$$

Under assumptions stronger than (2.8), we are also able to obtain VMO-regularity of solutions.

Theorem 2.2 (Partial VMO Regularity). *Under the assumptions and notations of Theorem 2.1, if f satisfies*

$$(2.10) \quad \lim_{\varrho \rightarrow 0^+} \varrho^2 \left(\oint_{B_\varrho(x)} |f|^\chi dx \right)^{1/\chi} = 0 \quad \text{locally uniformly in } x \in \Omega,$$

χ as in (2.1), then

$$u \in VMO_{\text{loc}}(\Omega_u; \mathbb{R}^N).$$

Our next results deal with continuity of solutions and their gradient. In order to state them, we introduce the (truncated) Riesz-type potential of f :

$$(2.11) \quad \mathbf{I}_{\alpha,\chi}^f(x_0, \sigma) := \int_0^\sigma \varrho^\alpha \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} \frac{d\varrho}{\varrho}$$

for $x_0 \in \Omega$ and $\sigma > 0$. For $\chi = 1$, the definition reduces to the averaged form of the classical truncated Riesz potential, which can be evaluated both on functions and, by replacing $|f| dx$ with a measure, on locally finite signed measures. For $\chi > 1$, $\mathbf{I}_{\alpha,\chi}^f$ is a χ -averaged counterpart of the truncated Riesz potential and retains the same natural scaling properties, mutatis mutandis; see [24, Chapter 4].

Theorem 2.3 (Partial continuity). *Under the assumptions and notations of Theorem 2.1, suppose that f fulfills*

$$(2.12) \quad \lim_{\sigma \rightarrow 0^+} \mathbf{I}_{2,\chi}^f(\cdot, \sigma) = 0 \quad \text{locally uniformly in } \Omega;$$

then

$$u \in C^0(\Omega_u; \mathbb{R}^N).$$

For what concerns Hölder continuity of solutions, this can be quantified whenever the right-hand side of (1.1) belongs to a certain Marcinkiewicz, or weak-Lebesgue, space $\mathcal{M}^m(\Omega; \mathbb{R}^N)$ (often denoted by $L^{m,\infty}(\Omega; \mathbb{R}^N)$). Specifically, we say that $f \in \mathcal{M}^m(\Omega; \mathbb{R}^N)$, for $m \geq 1$, if

$$(2.13) \quad \|f\|_{\mathcal{M}^m(\Omega; \mathbb{R}^N)}^m := \sup_{\lambda > 0} \lambda^m |\{x \in \Omega : |f(x)| \geq \lambda\}| < \infty.$$

We recall that standard interpolation theory [43, Theorem 3.18.8] yields

$$(2.14) \quad f \in L^m(\Omega; \mathbb{R}^N) \implies f \in \mathcal{M}^m(\Omega; \mathbb{R}^N), \quad m \in [1, \infty)$$

and $\mathcal{M}^m(\Omega; \mathbb{R}^N) \subseteq L^{m-\varepsilon}(\Omega; \mathbb{R}^N)$ for $\varepsilon \in (0, m-1]$, see also (3.5).

Theorem 2.4 (Quantified partial Hölder regularity). *Assume that (1.3)–(1.5) and (2.4) are in force, and let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1), with $s \in (0, 1)$. Let $\beta_0 \in (0, 1)$ and assume that*

$$(2.15) \quad f \in \mathcal{M}_{\text{loc}}^{n/(2-\beta_0)}(\Omega; \mathbb{R}^N).$$

Then there exists a constant $\delta_0 = \delta_0(\text{data}, \beta_0) \in (0, 1)$ with the following property. If, for every compact set $K \Subset \Omega$, there exists a radius $r_A = r_A(K) > 0$ such that

$$(2.16) \quad \sup_{x_0 \in K} E_{r_A}(A; x_0) < \delta_0,$$

then there exists an open set $\Omega_u \subseteq \Omega$ such that $\Omega_u = \Omega$ if $n = 2$, while, if $n \geq 3$,

$$\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) \leq n - 2 - \zeta, \quad \zeta = \zeta(\text{data}, \max\{\nu, 1\}) \in (0, 1),$$

and

$$u \in C_{\text{loc}}^{0,\beta_0}(\Omega_u; \mathbb{R}^N).$$

Finally, for every $x_0 \in \Omega_u$, there exist radii $r_{x_0} \leq \varrho_{x_0}/2$ such that $B_{2\varrho_{x_0}}(x_0) \Subset \Omega$, and

$$(2.17) \quad [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} \leq c \varrho_{x_0}^{-\beta_0} \left(\int_{B_{\varrho_{x_0}}(x_0)} |u - (u)_{B_{\varrho_{x_0}}(x_0)}|^2 dx \right)^{1/2} \\ + c \varrho_{x_0}^{1-\beta_0-2s} \text{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)) + c \|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_{2\varrho_{x_0}}(x_0))},$$

with constant $c = c(\text{data}, \beta_0)$.

Observe that if $A(\cdot, \cdot)$ satisfies the continuity condition (1.7), then it certainly fulfills (2.16), since by (2.5)–(2.6) and (1.7) one easily deduces for every $K \Subset \Omega$

$$\sup_{x_0 \in K} E_r(A; x_0) \leq \Lambda \omega(r) \xrightarrow{r \rightarrow 0} 0$$

uniformly. Therefore, taking also into account (2.14), Theorem 2.4 is a stronger version of Theorem 1.1.

We also stress that, for general coefficients $A(x, u)$, the radii r_{x_0}, ϱ_{x_0} may depend also on the solution u . More precisely, they must fulfill the excess decay conditions (6.5) and (6.6). However, in the case $A(x, u) = A(x)$ these conditions are no longer necessary, hence r_{x_0}, ϱ_{x_0} will become independent of u , so that the quantitative

estimate (2.17) becomes particularly useful for compactness arguments. We also stress that estimate (2.17) is a direct consequence of a more general bound involving a maximal operator applied to f . We refer to Theorem 6.1 and to equation (6.14) below for the precise statement.

For what concerns gradient regularity of solutions, we need to make the following stronger assumptions on $A(\cdot, \cdot)$. Specifically, we shall assume that there exists another modulus of continuity $\omega_x(\cdot)$ such that

$$(2.18) \quad |A(x, u) - A(y, u)| \leq \Lambda \omega_x(|x - y|) \quad \text{for all } x, y \in \Omega \text{ and } u \in \mathbb{R}^N.$$

In order to obtain Lipschitz continuity of solutions, a standard condition to impose on the moduli of continuity ω_x, ω_u is that they satisfies the so-called the *Dini property*

$$(2.19) \quad \int_0^1 \omega(\varrho) \frac{d\varrho}{\varrho} < \infty.$$

Theorem 2.5 (Partial Lipschitz continuity). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1), for $s \in (0, 1)$. Assume that (1.3)–(1.4) and (2.18) are in force, that $\omega_u(\cdot), \omega_x(\cdot)$ fulfill the Dini continuity property (2.19), and that $f \in L_{\text{loc}}^X(\Omega; \mathbb{R}^N)$ is such that for every $K \Subset \Omega$ there exists $r_f = r_f(K) > 0$ such that*

$$(2.20) \quad \sup_{x_0 \in K} \mathbf{I}_{1,\chi}^f(x_0, r_f) < \infty.$$

Then there exists an open set $\Omega_u \subseteq \Omega$, with $\Omega_u = \Omega$ in dimension $n = 2$, satisfying $\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) \leq n - 2 - \zeta$, $\zeta = \zeta(\text{data}, \max\{\nu, 1\}) \in (0, 1)$ in higher dimension $n \geq 3$, such that

$$Du \in L_{\text{loc}}^\infty(\Omega_u; \mathbb{R}^N).$$

Moreover, for every $x_0 \in \Omega_u$, there exist three radii $\bar{r}_{x_0}, r_{x_0}$ and ϱ_{x_0} such that $\bar{r}_{x_0} \leq r_{x_0}/4$, $r_{x_0} \leq \varrho_{x_0}/8$, $B_{r_{x_0}}(x_0) \Subset \Omega_u$, $B_{\varrho_{x_0}}(x_0) \Subset \Omega$, and

$$(2.21) \quad \|Du\|_{L^\infty(B_r(x_0))} \leq c \left(\int_{B_{2r}(x_0)} |Du|^2 dx \right)^{1/2} + c \left(\int_{B_{\varrho_{x_0}}(x_0)} |Du|^2 dx \right)^{1/2} + c \|\mathbf{I}_{1,\chi}^f(\cdot, r_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))} \\ + c r \left[r_{x_0}^{-2s} \text{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}, B_{r_{x_0}}(x_0)) + \varrho_{x_0}^{-2s} \text{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}, B_{\varrho_{x_0}}(x_0)) \right]$$

for all $r \in (0, \bar{r}_{x_0}]$, with constant $c = c(\text{data}, \max\{\nu, 1\})$.

Observe that, by the absolute continuity of the Lebesgue integral, condition (2.20) implies

$$\lim_{\sigma \rightarrow 0} \mathbf{I}_{1,\chi}^f(x, \sigma) = 0 \quad \text{for every } x \in \Omega.$$

By upgrading this hypothesis to uniform convergence, we attain gradient continuity of solutions. This is the content of the following

Theorem 2.6 (Partial gradient continuity). *Under the assumptions of Theorem 2.5, suppose additionally that*

$$(2.22) \quad \lim_{\sigma \rightarrow 0} \mathbf{I}_{1,\chi}^f(\cdot, \sigma) = 0 \quad \text{locally uniformly in } \Omega;$$

then

$$Du \in C^0(\Omega_u; \mathbb{R}^N).$$

We point out that (2.22) is satisfied, for instance, under the Lorentz-space assumption

$$f \in L_{\text{loc}}^{n,1}(\Omega; \mathbb{R}^N),$$

see [30, Lemma 3]. $L_{\text{loc}}^{n,1}(\Omega; \mathbb{R}^N)$ consists of the measurable functions $f : \Omega \rightarrow \mathbb{R}^N$ such that

$$\|f\|_{L^{n,1}(\Omega; \mathbb{R}^N)} := \int_0^\infty \lambda |\{x \in \Omega : |f(x)| > \lambda\}|^{1/n} < \infty$$

and the local variant is defined in the usual way. We recall that, by standard interpolation theory [43, Proposition 8.1.8], we have the continuous embeddings

$$L^q(\Omega) \hookrightarrow L^{n,1}(\Omega) \hookrightarrow L^n(\Omega) \quad \text{for any } q > n.$$

The next theorem concerns Hölder continuity of the gradient up to the threshold $2(1 - s)$ in the case $s > 1/2$ for general nonlocal terms merely satisfying (1.4) and (1.5).

Theorem 2.7 (Partial $C^{1,\beta}$ -regularity I). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1), for $s \in (0, 1)$. Let $\beta \in (0, 1)$, and assume that $A(\cdot, \cdot)$ satisfies (1.3), (2.4) and (2.18) with*

$$(2.23) \quad \omega_x(r) = \omega_u(r) = r^\beta,$$

that Φ satisfies (1.4) and (1.5), and

$$(2.24) \quad f \in \mathcal{M}_{\text{loc}}^{n/(1-\beta)}(\Omega; \mathbb{R}^N).$$

Then there exists an open set $\Omega_u \subseteq \Omega$, with $\Omega_u = \Omega$ when $n = 2$, $\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) \leq n - 2 - \zeta$, $\zeta = \zeta(\text{data}, \max\{\nu, 1\}) \in (0, 1)$ if $n \geq 3$, such that

$$u \in C_{\text{loc}}^{1, \min\{\beta, 2(1-s)\}}(\Omega_u; \mathbb{R}^N).$$

Moreover, for every $x_0 \in \Omega_u$, there exist three radii $\bar{r}_{x_0}, r_{x_0}, \varrho_{x_0}$ such that $\bar{r}_{x_0} \leq r_{x_0}/4$, $r_{x_0} \leq \varrho_{x_0}/8$, $B_{r_{x_0}}(x_0) \Subset \Omega_u$, $B_{2\varrho_{x_0}}(x_0) \Subset \Omega$, and the quantitative estimates

$$(2.25) \quad \|Du\|_{L^\infty(B_r(x_0))} \leq c \left(\int_{B_{2r}(x_0)} |Du|^2 dx \right)^{1/2} + c \left(\int_{B_{\varrho_{x_0}}(x_0)} |Du|^2 dx \right)^{1/2} + c \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))} \\ + c \bar{r}_{x_0} r_{x_0}^{-2s} \text{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)) + c \bar{r}_{x_0} \varrho_{x_0}^{-2s} \text{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)),$$

for all $r \in (0, \bar{r}_{x_0}]$, and

$$(2.26) \quad [Du]_{C^{0, \min\{\beta, 2(1-s)\}}(B_{\bar{r}_{x_0}/8}(x_0))} \leq c \left[\bar{r}_{x_0}^{-\min\{\beta, 2(1-s)\}} \left(\int_{B_{\bar{r}_{x_0}}(x_0)} |Du - (Du)_{B_{\bar{r}_{x_0}}(x_0)}|^2 dx \right)^{1/2} \right. \\ \left. + \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}^{1+\beta} + (1 + \bar{r}_{x_0}^{1-2s}) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right. \\ \left. + \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\bar{r}_{x_0}}(x_0))} + \bar{r}_{x_0}^{-2s} \text{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + 1 \right]$$

hold true, with $c = c(\text{data}, \max\{\nu, 1\}, \beta)$.

Under additional assumptions on Φ , we are instead able to obtain $C^{1,\beta}$ regularity, regardless of the value of $s \in (1/2, 1)$. For the reader's convenience, we write them once again. We assume that there exists a matrix valued function $a : \mathbb{R}^n \rightarrow \mathbb{R}^{N^2}$, $a = \{a_{\alpha, \beta}\}_{\alpha, \beta=1, \dots, N}$ such that

$$(2.27) \quad \begin{cases} a(z) = a(-z) \\ |a(z)| \leq \nu \end{cases} \quad z \in \mathbb{R}^n,$$

and

$$(2.28) \quad \Phi(x, y, v, w, t) = a(x - y) t.$$

Theorem 2.8 (Partial $C^{1,\beta}$ regularity II). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1), for $s \in (0, 1)$. Let $\beta \in (0, 1)$, and assume that $A(\cdot, \cdot)$ satisfies (1.3), (2.4), (2.18) and (2.23). Suppose also that Φ is of the form (2.27)–(2.28), and that*

$$f \in \mathcal{M}_{\text{loc}}^{n/(1-\beta)}(\Omega; \mathbb{R}^N).$$

Then there exists an open set $\Omega_u \subseteq \Omega$, with $\Omega_u = \Omega$ if $n = 2$ and $\dim_{\mathcal{H}}(\Omega \setminus \Omega_u) < n - 2 - \zeta$ if $n \geq 3$, $\zeta = \zeta(\text{data}, \max\{\nu, 1\})$, such that

$$(2.29) \quad u \in C_{\text{loc}}^{1,\beta}(\Omega_u; \mathbb{R}^N)$$

and for every $x_0 \in \Omega_u$, there exist radii $R_{x_0} = \bar{r}_{x_0}/32 \leq \bar{r}_{x_0} \leq r_{x_0} \leq \varrho_{x_0}$ such that $B_{2\varrho_{x_0}}(x_0) \Subset \Omega$, estimate (2.25) holds, and we have

$$(2.30) \quad [Du]_{C^{0,\beta}(B_{R_{x_0}}(x_0))} \leq C_1,$$

where C_1 depends on $\text{data}, \max\{\nu, 1\}, \beta, \bar{r}_{x_0}, r_{x_0}, \varrho_{x_0}$, and on any upper bound for $\|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))}, \text{Tail}(u - (u)_{B_{4R_{x_0}}(x_0)}; B_{4R_{x_0}}(x_0)), \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}$.

We conclude this section by emphasizing that, in the u -independent setting $A(x, u) = A(x)$, we obtain full-regularity and the quantitative estimates in Theorems 2.4–2.8 no longer exhibit the implicit dependence on u through the radii of the balls on which the estimates are stated. Instead, the solution u only enters the estimates explicitly through its L^1 -norm and its tail, as specified in the statement below. This feature provides a powerful tool in compactness arguments for the family of operators (1.1), since it yields uniform local regularity estimates that are well suited for passing to the limit along sequences of solutions with controlled L^1 and tail bounds. We decided to regroup all the results in the following statement:

Theorem 2.9 (Full regularity in the u -independent case). *Assume that $A(x, u) \equiv A(x)$ does not depend on u . Then each of the preceding partial regularity results, namely Theorems 2.1 to 2.8, holds under the same assumptions with the same conclusions, but with $\Omega_u = \Omega$, and all the corresponding local estimates are also valid on arbitrary compact subsets of Ω . The same full-regularity conclusion also applies to the forthcoming Theorem 6.1.*

Moreover, the radii involved in the local estimates of the corresponding theorems can be chosen independently of u . More precisely, for every $x_0 \in \Omega$, the initial radii ϱ_{x_0} and r_{x_0} can be chosen depending only on $d = \text{dist}(x_0, \partial\Omega)$, data , $\max\{\nu, 1\}$ and

- $r_A(\cdot)$, β_0 , and any upper bound on $\|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_{d/2}(x_0))}$ for Theorem 2.4;
- $r_A(\cdot)$, β_0 , and any upper bound on $\sup_{\substack{x \in B_{d/2}(x_0) \\ 0 < t \leq d/4}} t^{2-\beta_0} \left(\int_{B_t(x)} |f|^\chi dx \right)^{1/\chi}$ for Theorem 6.1 ;
- $\omega_\chi(\cdot)$ and any upper bound on $\|\mathbf{I}_{1,\chi}^f(\cdot, d/4)\|_{L^\infty(B_{d/2}(x_0))}$ for Theorem 2.5;
- β and any upper bound on $\|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{3d/4}(x_0))}$ for Theorems 2.7 and 2.8.

The radii $\bar{r}_{x_0}$ (for Theorems 2.5, 2.7 and 2.8) and R_{x_0} (for Theorem 2.8) can also be chosen independently of u . Since they are chosen not larger than ϱ_{x_0} , they inherit the corresponding dependences of ϱ_{x_0} listed above; in Theorem 2.8 one may take $R_{x_0} := \bar{r}_{x_0}/32$.

3. NOTATION, FUNCTIONS SPACES AND PRELIMINARY RESULTS

In the following section we collect a few basic definitions, and we recall a few standard inequalities and properties which we shall use throughout the paper.

3.1. Notation. We denote by c a generic constant larger or equal than one, possibly varying from line to line, whose relevant dependencies will be emphasized using parentheses: for instance, $c = c(n, s)$ will mean that c depends on n and s . In order to simplify the notation for various dependencies, we shall denote by data the set of given by (2.3). On the other hand, we shall often keep explicit the dependence on the constant ν due to scaling reasons (see Proposition 4.4 for instance). Special occurrences of constants will be denoted by c_1 , c_* , $\bar{c}$ or the like. We also write $\omega_n = |B_1|$ the Lebesgue measure of the unit ball on $\mathbb{R}^n$; $\mathbb{N} = \{1, 2, 3, \dots\}$, while $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

As we deal with vector-valued functions and want to avoid componentwise notation as much as possible, we distinguish the scalar product in $\mathbb{R}^N$ from that in $\mathbb{R}^{n \times N} \simeq \mathbb{R}^{nN}$. Namely, for $\xi, \mu \in \mathbb{R}^N$ and $p, q \in \mathbb{R}^{nN}$, we set

$$\xi \cdot \mu = \sum_{\alpha=1}^N \xi^\alpha \mu^\alpha, \quad \langle p, q \rangle = \sum_{i=1}^n \sum_{\alpha=1}^N p_i^\alpha q_i^\alpha.$$

Vectors in $\mathbb{R}^n$ are regarded as column vectors, while vectors in $\mathbb{R}^N$ are regarded as row vectors. With this convention, gradients of scalar functions are matrices with n rows and one column, whereas gradients of vector-valued functions are matrices with n rows and N columns. More precisely, if $u = (u^1, \dots, u^N)$, then the α -th column of Du is Du^α . Integrals and averages of vectors and matrices are taken componentwise.

For tensors $A \in \mathbb{R}^{nN \times nN}$ and matrices $\xi \in \mathbb{R}^{nN}$, typically gradients of vector-valued functions, we use the shorthand $A\xi$ for the matrix in $\mathbb{R}^{nN}$ whose components are

$$(A\xi)_i^\alpha = \sum_{j=1}^n \sum_{\beta=1}^N A_{ij}^{\alpha\beta} \xi_j^\beta, \quad i = 1, \dots, n, \quad \alpha = 1, \dots, N.$$

With $E \subseteq \mathbb{R}^n$ being a measurable set with positive, finite measure $0 < |E| < +\infty$, and with $F \in L^1(E; \mathbb{R}^\ell)$, $\ell \geq 1$, we shall denote by

$$(F)_E \equiv \int_E F(x) dx := \frac{1}{|E|} \int_E F(x) dx$$

its integral average. For $p \in (1, \infty)$, we write $p' = p/(p-1)$ for its Hölder conjugate.

We denote by $W^{1,2}(\Omega; \mathbb{R}^N)$ the Sobolev space of vector-valued functions $W^{1,2}(\Omega; \mathbb{R}^N) := \{u \in L^2(\Omega; \mathbb{R}^N) : Du \in L^2(\Omega; \mathbb{R}^{nN})\}$, where Du denotes the distributional gradient of u . Accordingly, $W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N)$ denotes the class of functions u such that $u \in W^{1,2}(A; \mathbb{R}^N)$ for every open set A compactly contained in Ω : $A \Subset \Omega$, while $W_c^{1,2}(\Omega; \mathbb{R}^N)$ denotes the subspace of functions with compact support in Ω . We will also denote by $[w]_{W^{s,2}(B_r)}$ the s -fractional seminorm

$$[w]_{W^{s,2}(B_r)}^2 := \int_{B_r} \int_{B_r} \frac{|w(x) - w(y)|^2}{|x - y|^{n+2s}} dx dy.$$

3.2. Function spaces. In this section we are going to consider measurable maps $F : U \rightarrow \mathbb{R}^\ell$, $\ell \in \mathbb{N} \setminus \{0\}$ and $U \subseteq \mathbb{R}^n$ measurable, not necessarily bounded, $n \in \mathbb{N}$, $n \geq 2$, and define several regularity function spaces where F can belong to.

The (L^2) -excess of $F \in L^2(U; \mathbb{R}^\ell)$ on balls is classically defined as

$$(3.1) \quad E(F; B_\varrho(x_0)) := \left(\int_{B_\varrho(x_0)} |F - (F)_{B_\varrho(x_0)}|^2 dx \right)^{1/2}$$

if $B_\varrho(x_0) \subseteq U$. It is very standard to see that

$$(3.2) \quad \left(\int_{B_\varrho(x_0)} |F - (F)_{B_\varrho(x_0)}|^2 dx \right)^{1/2} \leq \left(\int_{B_\varrho(x_0)} |F - z|^2 dx \right)^{1/2}$$

holds true for any $z \in \mathbb{R}^\ell$. As an immediate consequence of (3.2), for every $B_\varrho(x_1) \subseteq B_r(x_2)$ not necessarily concentric balls, we deduce

$$(3.3) \quad E(F; B_\varrho(x_1)) \leq \left(\frac{r}{\varrho} \right)^{n/2} E(F; B_r(x_2)).$$

As we are only interested in local results, we localize the classic definition of the BMO space in the following way: for $F \in L^2(U; \mathbb{R}^\ell)$ as above and a threshold $r_0 \in (0, +\infty]$, we set

$$E_{r_0}(F; x_0) := \sup_{\substack{\varrho \in (0, r_0]: \\ B_\varrho(x_0) \subseteq U}} E(F; B_\varrho(x_0));$$

notice that for $0 < r_0 < r_1 \leq +\infty$, $E_{r_0}(F; x_0) < +\infty$ if and only if $E_{r_1}(F; x_0) < +\infty$. Motivated by this fact, we say that $F(\cdot)$ is:

- BMO in U ($F \in BMO(U; \mathbb{R}^\ell)$) if these exist $c \geq 0$ and $r_0 > 0$ such that $E_{r_0}(F; x_0) \leq c$ for a.e. $x_0 \in U$; we set in this case

$$[F]_{BMO(U)} = \sup_{x_0 \in U} E_{r_0}(F; x_0) = \sup_{\substack{\varrho \in (0, r_0]: \\ B_\varrho(x_0) \subseteq U}} E(F; B_\varrho(x_0))$$

where sup is the essential supremum;

- locally BMO in U ($F \in BMO_{\text{loc}}(U; \mathbb{R}^\ell)$) if it is BMO in K for every compact subset $K \Subset U$;
- VMO in U if

$$\lim_{r \rightarrow 0^+} E_r(F; x_0) = 0;$$

- locally VMO in U ($F \in VMO_{\text{loc}}(U; \mathbb{R}^\ell)$) if $E_r(\cdot)$ converges to zero locally uniformly in U , that is, for every $K \Subset U$, $E_r(\cdot)$ converges to zero uniformly in K .

Remark 3.1. Our definition of BMO uses the L^2 -excess for simplicity, but it is equivalent to the more classical one employing the L^1 -excess, see [44, Section 1.3]. Note that, as a consequence of our definition, $F \in BMO_{\text{loc}}(U; \mathbb{R}^\ell)$ if and only if for every $K \Subset U$ there exists a threshold $r_K \in (0, \text{dist}(K, \partial U))$ such that $E_{r_K}(F; \cdot)$ is bounded in K . Moreover notice that $F \in VMO_{\text{loc}}(U; \mathbb{R}^\ell)$ if and only if for every $K \Subset U$ and $\delta > 0$, there exists a non-negative $\bar{r} = \bar{r}(\delta, K) < \text{dist}(K, \partial U)$ such that $\sup_K E_{\bar{r}}(\cdot) \leq \delta$.

Let $F \in C^0(U; \mathbb{R}^\ell)$; we say that $\omega : [0, \infty) \rightarrow [0, 1]$ is a modulus of continuity for F if $\omega(\cdot)$ is a concave, non-decreasing function such that $\omega(0) = \lim_{r \rightarrow 0^+} \omega(r) = 0$, and

$$|F(x) - F(y)| \leq \Lambda \omega(|x - y|), \quad x, y \in U,$$

for some $\Lambda > 0$. We recall that F is Dini-continuous if $\omega(\cdot)$ satisfies (2.19). Notice that if $\delta \in (0, 1/2]$ and $R_0 > 0$, using the monotonicity of $\omega(\cdot)$

$$\begin{aligned} \sum_{i=0}^{\infty} \omega(\delta^i R_0) &\leq \frac{1}{\log(1/\delta)} \sum_{i=1}^{\infty} \omega(\delta^i R_0) \int_{\delta^i R_0}^{\delta^{i-1} R_0} \frac{d\rho}{\rho} + \frac{1}{\log 2} \omega(R_0) \int_{R_0}^{2R_0} \frac{d\rho}{\rho} \\ (3.4) \quad &\leq \frac{1}{\log(1/\delta)} \left[\sum_{i=1}^{\infty} \int_{\delta^i R_0}^{\delta^{i-1} R_0} \omega(\rho) \frac{d\rho}{\rho} + \int_{R_0}^{2R_0} \omega(\rho) \frac{d\rho}{\rho} \right] = \frac{1}{\log(1/\delta)} \int_0^{2R_0} \omega(\rho) \frac{d\rho}{\rho}. \end{aligned}$$

Suppose that F belongs to the Marcinkiewicz, or weak-Lebesgue, space $\mathcal{M}^m(U; \mathbb{R}^\ell)$, i.e. (2.13) holds. By prescribing the decay of the measure of the super-level set, cake-layers formula implies a precise decay of the averages of F . Let indeed F be in $\mathcal{M}^m(U; \mathbb{R}^\ell)$ for some $m \geq 1$; then for every $\chi \in [1, m)$ and every ball $B_R(x_0) \subseteq U$, there holds

$$(3.5) \quad \left(\int_{B_R(x_0)} |F|^\chi dx \right)^{1/\chi} \leq \left[\frac{\chi}{m - \chi} \right]^{1/m} |B_R|^{-1/m} \|F\|_{\mathcal{M}^m(B_R(x_0); \mathbb{R}^\ell)}.$$

Notice that assumptions (2.15) and (2.24) imply (2.1) for a suitable exponent $\chi > 2_*$: in the case of (2.15) we can choose $2_* < \chi < \frac{n}{2 - \beta_0} < n$ and in the case of (2.24) we can choose $2_* < \chi < \frac{n}{1 - \beta} < n$. In dimension $n = 2$, whenever a Marcinkiewicz assumption with exponent depending on β_0 is used, the auxiliary exponent $2_* \in (1, 2)$ is chosen after β_0 , sufficiently close to 1, so that the interval $(2_*, 2/(2 - \beta_0))$ is non-empty. We conclude this paragraph with two elementary inequalities.

Lemma 3.2 (Fractional embedding). *Let $s \in (0, 1)$ and $B_R(x_0) \subseteq \mathbb{R}^n$. Assume that $w \in W^{1,2}(B_R(x_0); \mathbb{R}^N)$; then $w \in W^{s,2}(B_R(x_0); \mathbb{R}^N)$ with*

$$(3.6) \quad \frac{1}{|B_R(x_0)|} [w]_{W^{s,2}(B_R(x_0))}^2 = \int_{B_R(x_0)} \int_{B_R(x_0)} \frac{|w(x) - w(y)|^2}{|x - y|^{n+2s}} dx dy \leq \frac{c(n, N)}{1 - s} R^{2(1-s)} \int_{B_R(x_0)} |Dw|^2 dx.$$

Proof. By a scaling argument, we may assume that $x_0 = 0, R = 1$. We first extend w from B_1 to $\mathbb{R}^n$ via the spherical inversion map. Specifically, let

$$\tilde{w}(x) = \begin{cases} w(x) & x \in B_1 \\ w\left(\frac{x}{|x|^2}\right) & x \in \mathbb{R}^n \setminus B_1, \end{cases}$$

and with some work one can show that

$$(3.7) \quad \|D\tilde{w}\|_{L^2(B_4)} \leq c(n, N) \|Dw\|_{L^2(B_1)},$$

see for instance [14, Proposition 3.9.1]. Via a change of variables, the fundamental theorem of calculus, Fubini's theorem and (3.7), we obtain

$$\begin{aligned} \int_{B_1} \int_{B_1} \frac{|\tilde{w}(x) - \tilde{w}(y)|^2}{|x - y|^{n+2s}} dx dy &\leq \int_{B_1} \int_{B_2} \frac{|\tilde{w}(x) - \tilde{w}(z+x)|^2}{|z|^{n+2s}} dz dx \\ &\leq \int_{B_1} \int_{B_2} \int_0^1 |D\tilde{w}(x + tz)|^2 dt \frac{dz}{|z|^{n+2(s-1)}} dx \\ &= \int_{B_2} \int_0^1 \int_{B_1} |D\tilde{w}(x + tz)|^2 dx dt \frac{dz}{|z|^{n+2(s-1)}} \\ &\leq \int_{B_1} \frac{dz}{|z|^{n+2(s-1)}} \int_{B_3} |D\tilde{w}|^2 dx \leq \frac{c(n, N)}{1 - s} \int_{B_1} |Dw|^2 dx, \end{aligned}$$

which proves (3.6). □

The following interpolation inequality can be found in [42, Theorem 6.23].

Lemma 3.3. *Let $s \in (0, 1)$, $R > 0$ and $B_R(x_0) \subseteq \mathbb{R}^n$. If $u \in W^{s,2}(B_R(x_0))$ and $\varphi \in W^{1,\infty}(B_R(x_0))$, then $u\varphi \in W^{s,2}(B_R(x_0))$ and (if φ is not identically zero)*

$$(3.8) \quad [u\varphi]_{W^{s,2}(B_R(x_0))} \leq c(n, s) \|\varphi\|_{L^\infty(B_R(x_0))} \left[\|\varphi\|_{L^\infty(B_R(x_0))}^s \|D\varphi\|_{L^\infty(B_R(x_0))} \|u\|_{L^p(B_R(x_0))} + [u]_{W^{s,2}(B_R(x_0))} \right].$$

3.3. Nonlinear potentials of Riesz-type. We shall repeatedly use the following elementary comparison between the χ -averaged truncated Riesz-type potential introduced in (2.11),

$$\mathbf{I}_{\alpha,\chi}^f(x_0, R) = \int_0^R \varrho^\alpha \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} \frac{d\varrho}{\varrho},$$

and its dyadic counterpart. In particular, the continuous potential controls the sum of the corresponding quantities over a geometric sequence of radii, as well as the related one-scale supremum.

Lemma 3.4. *Let $f \in L^\chi(B_{2R}(x_0); \mathbb{R}^N)$ for some $\chi \geq 1$ and some ball $B_{2R}(x_0) \subseteq \mathbb{R}^n$; let moreover $\alpha \in (0, n/\chi)$ and $\tau \in (0, 1/2]$ be a parameter. Then, setting $R_j = \tau^j R$ for $j \in \mathbb{N}_0$, $B_j = B_{R_j}(x_0)$, and $R_{-1} = 2R$, we have*

$$(3.9) \quad \sum_{j=\bar{j}}^{\infty} R_j^\alpha \left(\int_{B_j} |f|^\chi dx \right)^{1/\chi} \leq c \mathbf{I}_{\alpha,\chi}^f(x_0, R_{\bar{j}-1}) \quad \text{for all } \bar{j} \in \mathbb{N}_0$$

and

$$(3.10) \quad \sup_{t \in (0, R_{\bar{j}}]} t^\alpha \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \leq c \mathbf{I}_{\alpha,\chi}^f(x_0, R_{\bar{j}-1}),$$

where both the constants depend only on n, χ and τ .

Proof. Recalling that $R_j = \tau^j R$, we have

$$\begin{aligned} \sum_{j=\bar{j}}^{\infty} R_j^\alpha \left(\int_{B_j} |f|^\chi dx \right)^{1/\chi} &\leq \frac{1}{\log(1/\tau)} \sum_{j=\bar{j}}^{\infty} R_j^\alpha \left(\int_{B_j} |f|^\chi dx \right)^{1/\chi} \int_{R_j}^{R_{j-1}} \frac{d\varrho}{\varrho} \\ &\leq \frac{\tau^{-n/\chi}}{\log(1/\tau)} \sum_{j=\bar{j}}^{\infty} \int_{R_j}^{R_{j-1}} \varrho^\alpha \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} \frac{d\varrho}{\varrho} = \frac{\tau^{-n/\chi}}{\log(1/\tau)} \mathbf{I}_{\alpha,\chi}^f(x_0, R_{\bar{j}-1}) \end{aligned}$$

which proves (3.9). For the proof of (3.10), we observe that if $0 < t \leq R_{\bar{j}}$, then we may find $j \geq \bar{j}$ such that $R_{j+1} < t \leq R_j$; we thus have

$$t^\alpha \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \leq \tau^{-n/\chi} R_j^\alpha \left(\int_{B_j(x_0)} |f|^\chi dx \right)^{1/\chi},$$

and the proof is completed by applying (3.9). $\square$

3.4. A-harmonic vectorial approximation. In the due course of the proof, we will exploit the quantitative harmonic type approximation; the version we are going to use is a small variation of the very clean and sharp version that can be found in [20]. We recall that, for a ball B and a tensor $\mathbf{A}$ as in the statement below, we say that a function $h \in W^{1,2}(B; \mathbb{R}^N)$ is $\mathbf{A}$ -harmonic if it solves

$$-\operatorname{div}(\mathbf{A}Dh) = 0 \quad \text{in } B, \quad \text{i.e.,} \quad \int_B \langle \mathbf{A}Dh, D\varphi \rangle dx = 0 \quad \text{for all } \varphi \in C_c^\infty(B).$$

Lemma 3.5 (Harmonic approximation Lemma). *Let $n \geq 2, N \geq 1$ and $\mathbf{A} \in \mathbb{R}^{nN \times nN}$ be a constant tensor satisfying for some $\Lambda \geq 1$*

$$(3.11) \quad \langle \mathbf{A}\xi, \xi \rangle \geq \Lambda^{-1} |\xi|^2 \quad \text{and} \quad |\mathbf{A}| \leq \Lambda \quad \text{for all } \xi \in \mathbb{R}^{nN}.$$

Let $B \subseteq \mathbb{R}^n$ be a ball and assume that $u \in W^{1,2}(B; \mathbb{R}^N)$ is a function satisfying

$$(3.12) \quad \int_B |Du|^{2+\zeta} dx \leq c_0,$$

for some $\zeta > 0$ and $c_0 \geq 1$. If u is also approximately $\mathbf{A}$ -harmonic, in the sense that it satisfies

$$\left| \int_B \langle \mathbf{A}Du, D\varphi \rangle dx \right| \leq \varepsilon M \|D\varphi\|_{L^\infty(B)} \quad \text{for all } \varphi \in C_c^\infty(B; \mathbb{R}^N)$$

for some $\varepsilon > 0$ and $M > 0$, there exists a (unique) $\mathbf{A}$ -harmonic function $h \in u + W_0^{1,2}(B; \mathbb{R}^N)$ such that

$$\int_B |Du - Dh|^2 dx \leq c \varepsilon^{\zeta/(1+\zeta)} [1 + M^2],$$

for a constant c depending only n, N, Λ, ζ and c_0 .

Proof. The Lemma follows from the main Theorem of [20] and the subsequent Remark, choosing $p = 2, q = 2 + \zeta$, using our assumption (3.12) and performing basic algebraic manipulations. $\square$

3.5. Nonlocal tail. An important quantity is the generalization of the so-called tail, namely for $\gamma > 0$, we shall use the γ -tail defined by

$$(3.13) \quad \mathbf{Tail}_\gamma(u; B_R(x_0)) := R^\gamma \int_{\mathbb{R}^n \setminus B_R(x_0)} \frac{|u(x)|}{|x - x_0|^{n+2s}} dx,$$

for $x_0 \in \mathbb{R}^n$ and $R > 0$. When $\gamma = 2s$, this is the classical Tail term introduced in [27, 38], and for brevity we set

$$(3.14) \quad \mathbf{Tail}_{2s}(u; B_R(x_0)) \equiv \mathbf{Tail}(u; B_R(x_0)) \implies \mathbf{Tail}_\gamma(u; B_R(x_0)) = R^{\gamma-2s} \mathbf{Tail}(u; B_R(x_0)).$$

Let us recall a few properties concerning the tail: clearly, whenever this is defined,

$$(3.15) \quad \mathbf{Tail}_\gamma(u + \lambda v; B_R(x_0)) \leq \mathbf{Tail}_\gamma(u; B_R(x_0)) + |\lambda| \mathbf{Tail}_\gamma(v; B_R(x_0))$$

for $\lambda \in \mathbb{R}$, but also

$$(3.16) \quad \mathbf{Tail}(1; B_R(x_0)) = \frac{n \omega_n}{1 + 2s}, \quad \mathbf{Tail}_\gamma(1; B_R(x_0)) = \frac{n \omega_n}{1 + 2s} R^{\gamma-2s}.$$

Note moreover that the γ -tail increases as the radius of the ball decreases:

$$(3.17) \quad 0 < R_1 \leq R_2 \implies \mathbf{Tail}_\gamma(u; B_{R_2}(x)) \leq \mathbf{Tail}_\gamma(u; B_{R_1}(x)).$$

The proof of the following Lemma is a consequence of [25, Lemma 3.2] and (3.14).

Lemma 3.6. *Let $w \in L_{2s}^1$. There exists $c = c(n, s) \geq 1$ such that*

$$(3.18) \quad \mathbf{Tail}_\gamma(w - (w)_{B_\varrho}; B_\varrho) \leq c \left(\frac{\varrho}{r}\right)^\gamma \mathbf{Tail}_\gamma(w - (w)_{B_r}; B_r) + c \frac{\varrho^\gamma}{r^{2s}} \int_{B_r} |w - (w)_{B_r}| dx \\ + c \int_\varrho^r \frac{\varrho^\gamma}{\mu^{2s}} \int_{B_\mu} |w - (w)_{B_\mu}| dx \frac{d\mu}{\mu}$$

if $B_\varrho \Subset B_r \subseteq \mathbb{R}^n$ are concentric balls. Moreover, if $B_\varrho(x_1) \subseteq B_r(x_2)$ are not necessarily concentric balls, then

$$(3.19) \quad \mathbf{Tail}_\gamma(w - (w)_{B_\varrho(x_1)}; B_\varrho(x_1)) \leq c \left(\frac{\varrho}{r}\right)^\gamma \left(\frac{r}{r - |x_1 - x_2|}\right)^{n+2s} \mathbf{Tail}_\gamma(w - (w)_{B_r(x_2)}; B_r(x_2)) \\ + c \varrho^{\gamma-2s} \left(\frac{r}{\varrho}\right)^n \int_{B_r(x_2)} |w - (w)_{B_r(x_2)}| dx,$$

and

$$(3.20) \quad \mathbf{Tail}_\gamma(w; B_\varrho(x_1)) \leq c \left(\frac{\varrho}{r}\right)^\gamma \left(\frac{r}{r - |x_1 - x_2|}\right)^{n+2s} \mathbf{Tail}_\gamma(w; B_r(x_2)) + c \varrho^{\gamma-2s} \left(\frac{r}{\varrho}\right)^n \int_{B_r(x_2)} |w| dx.$$

3.6. A modified excess. For our approach we need to introduce a modified excess, incorporating the non-locality in a weighted tail-type term: we set

$$(3.21) \quad \mathcal{E}_\gamma(w; B_\varrho(x_0)) := \left(\int_{B_\varrho(x_0)} |w - (w)_{B_\varrho(x_0)}|^2 dx \right)^{1/2} + \mathbf{Tail}_\gamma(w - (w)_{B_\varrho(x_0)}; B_\varrho(x_0)) \\ = E(w; B_\varrho(x_0)) + \mathbf{Tail}_\gamma(w - (w)_{B_\varrho(x_0)}; B_\varrho(x_0)),$$

where $\mathbf{Tail}_\gamma$ is defined in (3.13), and $E(\cdot)$ is the local L^2 -excess defined in (3.1).

Moreover, the map $x \mapsto \mathcal{E}_\gamma(w; B_\varrho(x))$ is continuous—see, for instance, [25, Proposition 3.4] for the continuity of the $\mathbf{Tail}$ -term, and therefore of $\mathbf{Tail}_\gamma$ by (3.14).

We now prove two properties; the first one shows that our modified excess enjoys, qualitatively, the same behavior as the standard excess under enlargement of the integration domain.

Lemma 3.7. *Let $w \in L^2_{\text{loc}}(\Omega; \mathbb{R}^N) \cap L^1_{2s}$ and suppose that $\gamma \geq 2s$. Then, for every couple of radii $0 < \varrho \leq r \leq 1$, there holds*

$$(3.22) \quad \mathcal{E}_\gamma(w; B_\varrho(x_0)) \leq c(n, s) \left(\frac{r}{\varrho}\right)^{n/2} \mathcal{E}_\gamma(w; B_r(x_0)).$$

Proof. Let us denote, for a moment, $w_r = w - (w)_{B_r(x_0)}$ for $r \in (0, \varrho]$. First observe that by (3.18), Hölder's inequality and (3.2), we get

$$(3.23) \quad \begin{aligned} \text{Tail}_\gamma(w_\varrho; B_\varrho(x_0)) &\leq c\left(\frac{\varrho}{r}\right)^\gamma \text{Tail}_\gamma(w_r; B_r(x_0)) + c\frac{\varrho^\gamma}{r^{2s}} \left(\int_{B_r(x_0)} |w_r|^2 dx \right)^{1/2} \\ &\quad + c\varrho^\gamma \int_\varrho^r \mu^{-2s} \left(\int_{B_\mu(x_0)} |w_\mu|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \\ &\leq c\left(\frac{\varrho}{r}\right)^\gamma \text{Tail}_\gamma(w_r; B_r(x_0)) + c\frac{\varrho^\gamma}{r^{2s}} \left(\int_{B_r(x_0)} |w_r|^2 dx \right)^{1/2} \\ &\quad + c\varrho^\gamma r^{n/2} \left(\int_{B_r(x_0)} |w_r|^2 dx \right)^{1/2} \int_\varrho^r \mu^{-2s-n/2} \frac{d\mu}{\mu} \\ &\leq c\left(\frac{\varrho}{r}\right)^\gamma \text{Tail}_\gamma(w_r; B_r(x_0)) + c\frac{\varrho^\gamma}{r^{2s}} \left[1 + \left(\frac{r}{\varrho}\right)^{2s+n/2} \right] \left(\int_{B_r(x_0)} |w_r|^2 dx \right)^{1/2} \end{aligned}$$

for some constant $c = c(n, s)$. Recalling that $\varrho \leq r \leq 1$ and $\gamma \geq 2s$, we get

$$\text{Tail}_\gamma(w_\varrho; B_\varrho(x_0)) \leq c\left(\frac{\varrho}{r}\right)^\gamma \text{Tail}_\gamma(w_r; B_r(x_0)) + c\varrho^{\gamma-2s} \left(\frac{r}{\varrho}\right)^{n/2} E_\gamma(w; B_r(x_0));$$

merging this inequality with the obvious estimate for the excess part of $\mathcal{E}_\gamma$ (using also (3.2)), we deduce (3.22). $\square$

The second property we prove justifies the fact that, qualitatively, the behavior of the modified excess $E_\gamma(\cdot)$ is dictated by the decay of the excess only.

Lemma 3.8. *Let $w \in L^1_{2s}$ and suppose that $\gamma \geq 2s$ as in the previous Lemma. Then*

$$\limsup_{\varrho \rightarrow 0^+} E(w; B_\varrho(x_0)) = 0 \quad \implies \quad \limsup_{\varrho \rightarrow 0^+} \text{Tail}_\gamma(w - (w)_{B_\varrho(x_0)}; B_\varrho(x_0)) = 0$$

so that, in particular

$$(3.24) \quad \limsup_{\varrho \rightarrow 0^+} \mathcal{E}_\gamma(w; B_\varrho(x_0)) > 0 \quad \implies \quad \limsup_{\varrho \rightarrow 0^+} E(w; B_\varrho(x_0)) > 0.$$

Proof. The proof the first implication follows again from (3.18), which implies, for any $\varrho \in (0, r)$ (see the first inequality in (3.23); here we use the same notation as there)

$$\begin{aligned} \text{Tail}_\gamma(w - (w)_{B_\varrho(x_0)}; B_\varrho(x_0)) &\leq c\varrho^\gamma \left[\frac{1}{r^\gamma} \text{Tail}_\gamma(w_r; B_r(x_0)) + \frac{1}{r^{2s}} \left(\int_{B_r(x_0)} |w - (w)_{B_r(x_0)}|^2 dx \right)^{1/2} \right] \\ &\quad + c\varrho^\gamma \int_\varrho^r \mu^{-2s} \left(\int_{B_\mu(x_0)} |w - (w)_{B_\mu(x_0)}|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \end{aligned}$$

and it is clear that we only need to check that the latter term is vanishing as $\varrho \rightarrow 0^+$. However, supposing the lim sup of the integral as $\varrho \rightarrow 0^+$ is $+\infty$ (otherwise the product, trivially, goes to zero), we can use de l'Hôpital rule and get, as $\gamma - 2s \geq 0$,

$$\begin{aligned} \limsup_{\varrho \rightarrow 0^+} \varrho^\gamma \int_\varrho^r \mu^{-2s} \left(\int_{B_\mu(x_0)} |w - (w)_{B_\mu(x_0)}|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \\ \leq \limsup_{\varrho \rightarrow 0^+} \frac{1}{\gamma \varrho^{\gamma-1}} \varrho^{-2s-1} \left(\int_{B_\varrho(x_0)} |w - (w)_{B_\varrho(x_0)}|^2 dx \right)^{1/2} \\ = \limsup_{\varrho \rightarrow 0^+} \frac{\varrho^{\gamma-2s}}{\gamma} \left(\int_{B_\varrho(x_0)} |w - (w)_{B_\varrho(x_0)}|^2 dx \right)^{1/2} = 0. \end{aligned}$$

The second implication is obvious by the definitions of $E(\cdot)$ and $\mathcal{E}_\gamma(\cdot)$. $\square$

4. CACCIOPPOLI'S INEQUALITY AND HIGHER INTEGRABILITY

The first main result of this section is a Caccioppoli-type inequality. For our future purposes, we prove it for a wider class of operators than (1.1). Specifically, let $\mathcal{A} : \Omega \times \mathbb{R}^N \times \mathbb{R}^{nN} \rightarrow \mathbb{R}^{nN}$, $\mathcal{A}(x, z, p) = \{\mathcal{A}_i^\alpha(x, z, p)\}_{i=1, \dots, n}^{\alpha=1, \dots, N}$ be a Carathéodory function satisfying the quadratic growth and ellipticity assumptions

$$(4.1) \quad |\mathcal{A}(x, z, p)| \leq \Lambda(1 + |p|) \quad \text{and} \quad \langle \mathcal{A}(x, z, p), p \rangle \geq \Lambda^{-1}(|p|^2 - 1),$$

for almost all $x \in \Omega$ and all $z \in \mathbb{R}^N, p \in \mathbb{R}^{nN}$, for a given constant $\Lambda > 0$. Moreover, let f be as in (2.1) and $G = \{G_i^\alpha\}_{i=1, \dots, n}^{\alpha=1, \dots, N}$ is a vector field in $L_{\text{loc}}^2(\Omega; \mathbb{R}^{nN})$.

We consider local weak solutions to

$$(4.2) \quad -\operatorname{div} \mathcal{A}(x, u, Du) + Q_{\Phi}^{\text{nl}} u = f + \operatorname{div} G \quad \text{in } \Omega$$

that is, functions $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ that satisfy

$$\begin{aligned} \int_{\Omega} \langle \mathcal{A}(x, u, Du), D\varphi \rangle dx + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \Phi(x, y, u(x), u(y), u(x) - u(y)) \cdot (\varphi(x) - \varphi(y)) \frac{dx dy}{|x - y|^{n+2s}} \\ = \int_{\Omega} f \cdot \varphi dx + \int_{\Omega} \langle G, D\varphi \rangle dx, \end{aligned}$$

for all test functions $\varphi \in W_c^{1,2}(\Omega; \mathbb{R}^N)$.

Lemma 4.1 (Caccioppoli's inequality). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (4.2) with the assumptions (1.4), (1.5) and (4.1) in force. Then there exists a radius $r_0 = r_0(\text{data}, \nu) \in (0, 1)$ such that for any ball $B_r \equiv B_r(x_0) \Subset \Omega$ with $0 < r \leq r_0$, and for all $\sigma \in (0, 1)$ and $u_0 \in \mathbb{R}^N$*

$$(4.3) \quad \begin{aligned} \sigma^n \int_{B_{\sigma r}} |Du|^2 dx \leq c \left(\frac{1}{(1-\sigma)^{2r^2}} + \frac{\nu}{(1-\sigma)^{2s r^{2s}}} \right) \int_{B_r} |u - u_0|^2 dx \\ + \frac{c}{(1-\sigma)^{n+2s}} \nu r^{-2s} \operatorname{Tail}(u - u_0; B_r) \int_{B_r} |u - u_0| dx + c r^2 \left(\int_{B_r} |f|^\chi dx \right)^{2/\chi} + c \int_{B_r} [|G|^2 + 1] dx \end{aligned}$$

holds true with a constant c depending on **data**.

Remark 4.2. We remark that the constant c appearing in (4.3) is independent of ν . The only implicit dependence on ν in Lemma 4.1 is in the radius r_0 , which has to satisfy

$$(4.4) \quad 0 < r_0 \leq \min \left\{ 1, \left(\frac{1}{c(\text{data}) \nu} \right)^{1/[2(1-s)]} \right\}.$$

In particular, for $\nu \ll 1$, we may take $r_0(\text{data}, \nu) \equiv 1$.

Proof of Lemma 4.1. It is sufficient to prove the estimate for $u_0 = 0$, for otherwise the function $u - u_0$ is a weak solution to $-\operatorname{div}(\mathcal{A}_0(x, v, Dv)) + Q_{\Phi_0}^{\text{nl}} v = f + \operatorname{div} G$, where $\mathcal{A}_0(x, z, p) = \mathcal{A}(x, z + u_0, p)$ still satisfies (4.1) and $\Phi_0(x, y, v, w, t) = \Phi(x, y, v + u_0, w + u_0, t)$ satisfies (1.4) and (1.5). We fix $\sigma \in (0, 1)$ and two intermediate radii $\sigma r \leq \tau_1 < \tau_2 \leq r$, and we define $\tilde{\tau}_1 = (\tau_1 + \tau_2)/2$ and $\tilde{\tau}_2 = (3\tau_2 + \tau_1)/4$, so that $\tau_1 < \tilde{\tau}_1 < \tilde{\tau}_2 < \tau_2$.

Let η be a cut-off function $\eta \in C_c^\infty(B_{\tilde{\tau}_2})$ such that $0 \leq \eta \leq 1$, $\eta \equiv 1$ in $B_{\tilde{\tau}_1}$, and $\|D\eta\|_\infty \leq c/(\tau_2 - \tau_1)$. Take $\varphi = (u - (u)_{B_{\tau_2}})\eta^2$ as a test function in (4.2). Using (1.5), we get

$$(4.5) \quad \begin{aligned} (O_1) + (O_2) &:= \int_{B_{\tau_2}} f \cdot \varphi dx + \int_{B_{\tau_2}} \langle G, D\varphi \rangle dx = \int_{B_{\tau_2}} \langle \mathcal{A}(x, u, Du), D\varphi \rangle dx \\ &+ \int_{B_{\tau_2}} \int_{B_{\tau_2}} \Phi(x, y, u(x), u(y), u(x) - u(y)) \cdot (\varphi(x) - \varphi(y)) \frac{dx dy}{|x - y|^{n+2s}} \\ &+ 2 \int_{\mathbb{R}^n \setminus B_{\tau_2}} \int_{B_{\tau_2}} \Phi(x, y, u(x), u(y), u(x) - u(y)) \cdot \varphi(x) \frac{dx dy}{|x - y|^{n+2s}} =: (I) + (II) + (III). \end{aligned}$$

We estimate separately each term in (4.5): via Hölder's inequality and, after recalling (2.1)–(2.2) that implies $\chi' \leq 2^*$, using Sobolev's inequality we obtain

$$\begin{aligned}
 (O_1) &\leq \left(\int_{B_{\tau_2}} |f\eta|^\chi dx \right)^{1/\chi} \left(\int_{B_{\tau_2}} |(u - (u)_{B_{\tau_2}})\eta|^{\chi'} dx \right)^{1/\chi'} \\
 &\leq c(n, N, \chi) \left(\int_{B_{\tau_2}} |f|^\chi dx \right)^{1/\chi} \tau_2^{1+n/2-n/\chi} \left(2 \int_{B_{\tau_2}} |Du|^2 \eta^2 dx + 4 \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}|^2 |D\eta|^2 dx \right)^{1/2} \\
 (4.6) \quad &\leq \delta \int_{B_{\tau_2}} |Du|^2 \eta^2 dx + c(n, N, \chi, \delta) r^{n+2(1-\frac{n}{\chi})} \left(\int_{B_r} |f|^\chi dx \right)^{2/\chi} + \frac{c}{(\tau_2 - \tau_1)^2} \int_{B_{\tau_2}} |u|^2 dx,
 \end{aligned}$$

where in the last inequality we used Young's inequality for $\delta > 0$ to be chosen later, and triangle inequality too. Then, by Young's inequality and the triangle inequality, we similarly compute

$$\begin{aligned}
 (O_2) &= \int_{B_{\tau_2}} \langle G, Du \rangle \eta^2 dx + 2 \int_{B_{\tau_2}} \langle G, D\eta \rangle (u - (u)_{B_{\tau_2}}) \eta dx \\
 &\leq \delta \int_{B_{\tau_2}} |Du|^2 \eta^2 dx + c(n, N, \delta) \int_{B_r} |G|^2 dx + \frac{c(n, N, \delta)}{(\tau_2 - \tau_1)^2} \int_{B_{\tau_2}} |u|^2 dx,
 \end{aligned}$$

again for $\delta > 0$ to be chosen. Next, by (4.1) we deduce

$$\begin{aligned}
 (I) &\geq \Lambda^{-1} \int_{B_{\tau_2}} (|Du|^2 - 1) \eta^2 dx - 2\Lambda \int_{B_{\tau_2}} (1 + |Du|) |D\eta| |u - (u)_{B_{\tau_2}}| \eta dx \\
 &\geq \Lambda^{-1} \int_{B_{\tau_2}} |Du|^2 \eta^2 dx - \Lambda^{-1} - \frac{\Lambda^{-1}}{2} \int_{B_{\tau_2}} |Du|^2 \eta^2 dx - c(\Lambda) \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}|^2 |D\eta|^2 dx \\
 &\quad - 2\Lambda \int_{B_{\tau_2}} |D\eta| |u - (u)_{B_{\tau_2}}| \eta dx \\
 &\geq \frac{\Lambda^{-1}}{2} \int_{B_{\tau_2}} |Du|^2 \eta^2 dx - \frac{c}{(\tau_2 - \tau_1)^2} \int_{B_{\tau_2}} |u|^2 dx - c,
 \end{aligned}$$

where we have used Young's inequality several times, the assumptions on η and the triangle inequality. Using (1.4), Hölder's inequality, the product rule (3.8) and (3.6), we infer

$$\begin{aligned}
 -(II) &\leq \nu \int_{B_{\tau_2}} \int_{B_{\tau_2}} \frac{|u(x) - u(y)| |\varphi(x) - \varphi(y)|}{|x - y|^{n+2s}} dx dy \\
 &\leq \nu [u]_{s,2;B_{\tau_2}} [\varphi]_{s,2;B_{\tau_2}} \\
 &\leq \frac{\nu}{2} [u]_{s,2;B_{\tau_2}}^2 + c\nu \left[\|\eta^2\|_{L^\infty}^{2(1-s)} \|D\eta^2\|_{L^\infty}^{2s} \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}|^2 dx + \|\eta^2\|_{L^\infty}^2 [u]_{s,2;B_{\tau_2}}^2 \right] \\
 &\leq c\nu [u]_{s,2;B_{\tau_2}}^2 + \frac{c\nu}{(\tau_2 - \tau_1)^{2s}} \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}|^2 dx \\
 &\leq c\nu r^{2(1-s)} \int_{B_{\tau_2}} |Du|^2 dx + \frac{c\nu}{(\tau_2 - \tau_1)^{2s}} \int_{B_{\tau_2}} |u|^2 dx,
 \end{aligned}$$

with $c = c(n, s)$. Finally, using (1.4), by using

$$\mathbf{Tail}(u - (u)_{B_{\tau_2}}; B_{\tau_2}) \leq \mathbf{Tail}(u; B_{\tau_2}) + c(n, s) |(u)_{B_{\tau_2}}|$$

and the fact that, for $x \in B_{\tilde{\tau}_2} = B_{\tilde{\tau}_2}(x_0)$ and $y \in \mathbb{R}^n \setminus B_{\tau_2}(x_0)$, one has

$$(4.7) \quad |y - x| \geq \frac{\tau_2 - |x - x_0|}{\tau_2} |y - x_0| \geq \frac{\tau_2 - \tilde{\tau}_2}{\tau_2} |y - x_0| \approx \frac{\tau_2 - \tau_1}{\tau_2} |y - x_0|$$

(as $\tau_2 |y - x| \leq \tau_2 |y - x_0| + \tau_2 |x - x_0| \leq \tau_2 |y - x_0| + |y - x_0| |x - x_0|$), we estimate

$$-(III) \leq c\nu \int_{\mathbb{R}^n \setminus B_{\tau_2}} \int_{B_{\tilde{\tau}_2}} \left[|u(x) - (u)_{B_{\tau_2}}| |u(y) - (u)_{B_{\tau_2}}| + |u(x) - (u)_{B_{\tau_2}}|^2 \right] \frac{dx dy}{|x - y|^{n+2s}}$$

$$\begin{aligned}
& \stackrel{(4.7)}{\leq} c\nu \left(\frac{\tau_2}{\tau_2 - \tau_1} \right)^{n+2s} \tau_2^{-2s} \left[\mathbf{Tail}(u - (u)_{B_{\tau_2}}; B_{\tau_2}) \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}| dx \right. \\
& \quad \left. + \mathbf{Tail}(1; B_{\tau_2}) \int_{B_{\tau_2}} |u - (u)_{B_{\tau_2}}|^2 dx \right] \\
& \stackrel{(3.16)}{\leq} \frac{c\nu\tau_2^n}{(\tau_2 - \tau_1)^{n+2s}} \left[\mathbf{Tail}(u; B_{\tau_2}) \int_{B_{\tau_2}} |u| dx + \left(\int_{B_{\tau_2}} |u| dx \right)^2 + \int_{B_{\tau_2}} |u|^2 dx \right] \\
& \stackrel{\text{Jensen}}{\leq} \frac{c\nu\tau_2^n}{(\tau_2 - \tau_1)^{n+2s}} \left[\mathbf{Tail}(u; B_{\tau_2}) \int_{B_{\tau_2}} |u| dx + \int_{B_{\tau_2}} |u|^2 dx \right] \\
& \stackrel{(3.20)}{\leq} \frac{c\nu\tau_2^n}{(\tau_2 - \tau_1)^{n+2s}} \left[\mathbf{Tail}(u; B_r) + \frac{r^n}{\tau_2^n} \int_{B_r} |u| dx \right] \int_{B_r} |u| dx + \int_{B_r} |u|^2 dx \\
(4.8) \quad & \leq \frac{c\nu r^n}{(\tau_2 - \tau_1)^{n+2s}} \left[\mathbf{Tail}(u; B_r) \int_{B_r} |u| dx + \int_{B_r} |u|^2 dx \right].
\end{aligned}$$

Inserting (4.6)–(4.8) into (4.5), choosing $\delta = \Lambda^{-1}/8$, multiplying both members by r^n and recalling that $\eta \equiv 1$ on B_{τ_1} and $r \leq r_0$, we find

$$\begin{aligned}
\int_{B_{\tau_1}} |Du|^2 dx & \leq \bar{c}\nu r_0^{2(1-s)} \int_{B_{\tau_2}} |Du|^2 dx + c \left(\frac{1}{(\tau_2 - \tau_1)^2} + \frac{\nu}{(\tau_2 - \tau_1)^{2s}} \right) \int_{B_r} |u|^2 dx \\
& \quad + \frac{c\nu r^n}{(\tau_2 - \tau_1)^{n+2s}} \left[\mathbf{Tail}(u; B_r) \int_{B_r} |u| dx + \int_{B_r} |u|^2 dx \right] \\
& \quad + c r^{n+2(1-\frac{n}{\chi})} \left(\int_{B_r} |f|^\chi dx \right)^{2/\chi} + c \int_{B_r} [|G|^2 + 1] dx
\end{aligned}$$

with c and $\bar{c}$ depending only on **data**; now we choose $r_0 \leq 1$ such that

$$\bar{c}\nu r_0^{2(1-s)} \leq \frac{1}{2}$$

(which yields (4.4)) and we apply the well-known “simple, but fundamental Lemma” by Giaquinta & Giusti [35, Lemma 1] (see also [37, Lemma 6.1]) to reabsorb the first integral on the right-hand side, getting

$$\begin{aligned}
\int_{B_{\sigma r}} |Du|^2 dx & \leq c \left(\frac{1}{(1-\sigma)^2 r^2} + \frac{\nu}{(1-\sigma)^{2s} r^{2s}} \right) \int_{B_r} |u|^2 dx \\
& \quad + \frac{c\nu}{(1-\sigma)^{n+2s} r^{2s}} \mathbf{Tail}(u; B_r) \int_{B_r} |u| dx + c r^{n+2(1-\frac{n}{\chi})} \left(\int_{B_r} |f|^\chi dx \right)^{2/\chi} + c \int_{B_r} [|G|^2 + 1] dx.
\end{aligned}$$

After taking averages, this is the desired estimate. $\square$

Our next goal is to show that, under suitable integrability assumptions on f, G , we have higher integrability of the gradient of solutions to (4.2). Before stating such result, we anticipate a useful lemma to rewrite the right-hand side of (4.2) in divergence form by using classical potential theory for the Poisson equation. More precisely, we have

Lemma 4.3. *Let $f \in L^\chi(B_R(x_0))$, with $\chi > 2_*$. Then there exist an exponent $q > 2$ and a vector field $F \in L^q(B_R(x_0); \mathbb{R}^n)$ such that $f = \operatorname{div} F$ in distributional sense, and satisfying*

$$(4.9) \quad \left(\int_{B_R(x_0)} |F|^q dx \right)^{1/q} \leq c(n, \chi, q) R \left(\int_{B_R(x_0)} |f|^\chi dx \right)^{1/\chi}.$$

If $2_* < \chi < n$, one may take

$$q = \chi^* := \frac{n\chi}{n - \chi} > 2,$$

while if $\chi \geq n$, one may take any number $q > 2$ (actually, for $\chi > n$ one might even take $q = +\infty$).

Proof. We extend f to the whole space by setting $f \equiv 0$ in $\mathbb{R}^n \setminus B_R(x_0)$. Consider the function

$$w = \int_{\mathbb{R}^n} \Gamma(x - y) f(y) dy,$$

where Γ is the fundamental solution of the Laplace equation in $\mathbb{R}^n$. It follows that $-\Delta w = f$, and classical potential estimates ([36, Section 7.8], [39, Section 2]) give $|Dw| \leq c(n) \mathbf{I}_1(|f|)$. The assertion follows now by setting $F = -Dw$, and taking into account the Hardy–Littlewood–Sobolev inequality and its standard localized consequences. If $1 < \chi < n$, this is immediate as $\chi^* > 2$ by $\chi > 2_*$. If $\chi \geq n$, we fix $q > 2$ and choose

$$\chi_0 = q_* = \frac{nq}{n+q} \in (2_*, n), \quad q = \chi_0^*$$

and using the previous case with exponent χ_0 and Hölder's inequality gives (4.9). Finally, if $\chi > n$, one can use Hölder's inequality applied to the kernel $|x - y|^{1-n}$. $\square$

The main result of this section is the following:

Proposition 4.4 (Higher integrability of the gradient). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (4.2), under assumptions (1.4)–(1.5), (4.1) in force, and $f = \text{div } F$ as above. Then there exist constants*

$$\zeta = \zeta(\text{data}, \max\{\nu, 1\}) > 0 \quad \text{and} \quad c = c(\text{data}, \max\{\nu, 1\}) \geq 1$$

such that, if $F, G \in L_{\text{loc}}^{2+\zeta}(\Omega; \mathbb{R}^{nN})$, then $|Du| \in L_{\text{loc}}^{2+\zeta}(\Omega)$, with quantitative estimate

$$(4.10) \quad \left(\int_{B_{R_0/2}} |Du|^{2+\zeta} dx \right)^{1/(2+\zeta)} \leq c \left(\int_{B_{2R_0}} |Du|^2 dx \right)^{1/2} + c \left(\int_{B_{2R_0}} [|F| + |G| + 1]^{2+\zeta} dx \right)^{1/(2+\zeta)} + c \nu R_0^{1-2s} \text{Tail}(u - (u)_{B_{2R_0}}; B_{2R_0}).$$

This holds for all $B_{2R_0} \equiv B_{2R_0}(x_0) \Subset \Omega$ with $R_0 \leq r_0(\nu) \leq 1$ satisfying (4.4).

The proof of Proposition 4.4 is very similar to that of [15, Proposition 3.4], our starting point being (4.3). For the sake of completeness, we sketch the full proof.

We start by proving a reverse Hölder inequality, in the same spirit as [15, Lemma 3.3]. To this end, we fix a parameter $\tau > 0$, depending only on s , satisfying

$$(4.11) \quad \max\{s, 1/2\} < \tau < 1.$$

Proposition 4.5 (Reverse Hölder's inequality). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (4.2), under assumptions (1.4)–(1.5), (4.1) in force. Suppose that $f = \text{div } F$, with $F, G \in L_{\text{loc}}^2(\Omega; \mathbb{R}^{nN})$. Let $r < r_0(\nu)$, where $r_0(\nu)$ is the threshold appearing in Lemma 4.1, $x_0 \in \Omega$, and let $\ell \in \mathbb{N}$ be such that*

$$B_{2^{\ell+1}r}(x_0) \Subset \Omega \quad \text{and} \quad 2^{\ell+1}r \leq 1.$$

Then there exists a constant $c = c(\text{data})$ such that

$$(4.12) \quad \int_{B_{r/2}} |Du|^2 dx \leq c \left(1 + \frac{\nu^2}{\delta} \right) \left(\int_{B_r} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} + \delta \left[\left(\int_{B_{2^\ell r}} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} + \int_r^{2^\ell r} \mu^{1-2\tau} \left(\int_{B_\mu} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} \frac{d\mu}{\mu} \right] + \delta \left[\nu r (2^\ell r)^{-2s} \text{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r}) \right]^2 + c \int_{B_r} [|F|^2 + |G|^2 + 1] dx$$

holds for all $\delta \in (0, 1]$.

Proof. In view of Lemma 4.3 and (4.9), we may assume that $f \equiv 0$ by replacing G with $F + G$, where $f = \text{div } F$; we omit the center of the various balls under consideration. We want to estimate the terms on the right-hand side of (4.3) with $\sigma = 1/2$ and $u_0 = (u)_{B_r}$. By Sobolev and Hölder inequalities, we have

$$(4.13) \quad \int_{B_r(x_0)} |u - (u)_{B_r(x_0)}| dx \leq \left(\int_{B_r(x_0)} |u - (u)_{B_r(x_0)}|^2 dx \right)^{1/2} \leq c(n, N) r \left(\int_{B_r(x_0)} |Du|^{2n/(n+2)} dx \right)^{(n+2)/[2n]};$$

using (3.18), we can estimate the tail term as follows:

$$r^{-2s} \text{Tail}(u - (u)_{B_r}; B_r) \leq c (2^\ell r)^{-2s} \text{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r})$$

$$+ c(2^\ell r)^{-2s} \int_{B_{2^\ell r}} |u - (u)_{B_{2^\ell r}}| dx + c \int_r^{2^\ell r} \mu^{-2s} \int_{B_\mu} |u - (u)_{B_\mu}| dx \frac{d\mu}{\mu}.$$

Using the previous estimate, Poincaré's inequality and then repeated use of weighted Young's inequality, we get

$$\begin{aligned}
& \nu r^{-2s} \mathbf{Tail}(u - (u)_{B_r}; B_r) \int_{B_r} |u - (u)_{B_r}| dx \\
& \leq c \nu \left[r(2^\ell r)^{-2s} \mathbf{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r}) \right. \\
& \quad \left. + r(2^\ell r)^{1-2s} \int_{B_{2^\ell r}} |Du| dx + r \int_r^{2^\ell r} \mu^{1-2s} \int_{B_\mu} |Du| dx \frac{d\mu}{\mu} \right] \int_{B_r} |Du| dx \\
& \leq \delta_1 \left[\nu r(2^\ell r)^{-2s} \mathbf{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r}) \right]^2 + \delta_2 \left(\int_{B_{2^\ell r}} |Du| dx \right)^2 \\
(4.14) \quad & + \delta_3 \left[\int_r^{2^\ell r} \mu^{1-2s} \int_{B_\mu} |Du| dx \frac{d\mu}{\mu} \right]^2 + c \left[\frac{1}{\delta_1} + \frac{\nu^2 r^2 (2^\ell r)^{2(1-2s)}}{\delta_2} + \frac{\nu^2 r^2}{\delta_3} \right] \left(\int_{B_r} |Du| dx \right)^2
\end{aligned}$$

for $\delta_i \in (0, 1]$, $i = 1, 2, 3$. Moreover, for $\mu \leq 2^\ell r \leq 1$ we may bound $\mu^{1-2s} \leq \mu^{1-2\tau}$ thanks to (4.11), so by Hölder's inequality applied with measure $d\lambda(\mu) = \mu^{1-2\tau} \frac{d\mu}{\mu}$, we compute

$$\begin{aligned}
& \left(\int_r^{2^\ell r} \mu^{1-2s} \int_{B_\mu} |Du| dx \frac{d\mu}{\mu} \right)^2 \leq \int_r^{2^\ell r} d\lambda(\mu) \int_r^{2^\ell r} \left(\int_{B_\mu} |Du| dx \right)^2 d\lambda(\mu) \\
(4.15) \quad & \leq \frac{r^{1-2\tau}}{2\tau-1} \int_r^{2^\ell r} \mu^{1-2\tau} \left(\int_{B_\mu} |Du| dx \right)^2 \frac{d\mu}{\mu};
\end{aligned}$$

in the last inequality we used that $\tau \geq 1/2$ by (4.11). Combining (4.14)–(4.15), choosing $\delta_1 = \delta_2 = \delta$ and $\delta_3 = (2\tau-1)^2 r^{2(2\tau-1)} \delta$ and finally using the facts $r^2 (2^\ell r)^{2(1-2s)} \leq (2^\ell r)^{2(2-2s)} \leq 1$ and $r^{1-\tau} \leq 1$ by (4.11), we infer

$$\begin{aligned}
& \nu r^{-2s} \mathbf{Tail}(u - (u)_{B_r}; B_r) \int_{B_r} |u - (u)_{B_r}| dx \leq \delta \left[\nu r(2^\ell r)^{-2s} \mathbf{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r}) \right]^2 \\
& + \delta \left(\int_{B_{2^\ell r}} |Du| dx \right)^2 + \delta \int_r^{2^\ell r} \mu^{1-2\tau} \left(\int_{B_\mu} |Du| dx \right)^2 \frac{d\mu}{\mu} + \frac{c(1+\nu^2)}{\delta} \left(\int_{B_r} |Du| dx \right)^2
\end{aligned}$$

with c depending on n, N and s . To conclude, it suffices to use Caccioppoli's inequality (4.3) with the choice $\sigma = 1/2$, $u_0 = (u)_{B_r}$, which, together with (4.13), (4.15) and Hölder's inequality gives

$$\begin{aligned}
& \int_{B_{r/2}} |Du|^2 dx \leq c \left(1 + \frac{\nu^2}{\delta} + \frac{\nu}{r^{2(s-1)}} \right) \left(\int_{B_r} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} + c \int_{B_r} [|F|^2 + |G|^2 + 1] dx \\
& + \delta \left[\left(\int_{B_{2^\ell r}} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} + \int_r^{2^\ell r} \mu^{1-2\tau} \left(\int_{B_\mu} |Du|^{2n/(n+2)} dx \right)^{(n+2)/n} \frac{d\mu}{\mu} \right] \\
& + \delta \left[\nu r (2^\ell r)^{-2s} \mathbf{Tail}(u - (u)_{B_{2^\ell r}}; B_{2^\ell r}) \right]^2,
\end{aligned}$$

and using that $\nu/r^{2(s-1)} \leq c(\mathbf{data})$ by (4.4), we get the desired outcome. $\square$

We are now in the position to prove Proposition 4.4, whose proof is based on the reverse Hölder's inequality (4.12) and an exit time argument as in [19, 40] and [15, Proposition 3.4].

Proof of Proposition 4.4. In the following, we set

$$|\tilde{F}| := |F| + |G| + 1,$$

and we fix two constants $\zeta > 0$ and $M \geq 1$ depending only on **data** and $\max\{\nu, 1\}$, to be determined later. Given R_0 and x_0 as in the statement of the proposition, for all R and $\bar{x} \in \Omega$ such that $B_R(\bar{x}) \subseteq B_{R_0}(x_0)$, we set

$$\begin{aligned}\Upsilon(\bar{x}, R) &:= \left(\int_{B_R(\bar{x})} |\tilde{F}|^{2+\zeta} dx \right)^{1/(2+\zeta)} \\ \Psi_M(\bar{x}, R) &:= \left(\int_{B_R(\bar{x})} |Du|^2 dx \right)^{1/2} + M \left(\int_{B_R(\bar{x})} |\tilde{F}|^2 dx \right)^{1/2},\end{aligned}$$

and

$$(4.16) \quad \mathcal{T}(\bar{x}, R) := \left[\int_R^{2^\ell R} \mu^{1-2\tau} \left(\int_{B_\mu(\bar{x})} |Du|^{2\gamma} dx \right)^{1/\gamma} \frac{d\mu}{\mu} \right]^{1/2} + \nu R \int_{\mathbb{R}^n \setminus B_{2^\ell R}(\bar{x})} \frac{|u(y) - (u)_{B_{2^\ell R}(\bar{x})}|}{|y - \bar{x}|^{n+2s}} dy,$$

where $\gamma = n/(n+2) \in [1/2, 1)$ and $\ell = \ell(R) \in \mathbb{N}$ is such that

$$(4.17) \quad \frac{R_0}{2} \leq 2^\ell R < R_0;$$

we also set

$$(4.18) \quad \Xi(\bar{x}, R) := \Upsilon(\bar{x}, R) + \Psi_M(\bar{x}, R) + \mathcal{T}(\bar{x}, R).$$

We fix two radii $r_1, r_2 > 0$ such that

$$(4.19) \quad \frac{R_0}{2} \leq r_1 < r_2 \leq R_0$$

and define the quantity

$$(4.20) \quad \lambda_0 := \sup \left\{ \Xi(\bar{x}, R) : \bar{x} \in B_{r_1}(x_0), \frac{r_2 - r_1}{16} < R \leq \frac{R_0}{2} \right\}.$$

Step 1: Upper bound for λ_0 . Let

$$(4.21) \quad \bar{x} \in B_{r_1}(x_0) \quad \text{and} \quad \frac{r_2 - r_1}{16} < R \leq \frac{R_0}{2};$$

we have $B_R(\bar{x}) \subseteq B_{2R_0}(x_0)$, and for this reason

$$(4.22) \quad \Upsilon(\bar{x}, R) \leq \left(\frac{2R_0}{R} \right)^{n/(2+\zeta)} \Upsilon(x_0, 2R_0) \leq \left(\frac{32R_0}{r_2 - r_1} \right)^{n/2} \Upsilon(x_0, 2R_0).$$

Moreover, by (4.17)–(4.21), we have $B_{2^{\ell+1}R}(\bar{x}) \subseteq B_{2R_0}(x_0)$, so that for all $R \leq \mu \leq 2^\ell R$, we similarly have

$$\left(\int_{B_\mu(\bar{x})} |Du|^{2\gamma} dy \right)^{1/\gamma} \leq c(n) \left(\frac{32R_0}{r_2 - r_1} \right)^{n/\gamma} \left(\int_{B_{2R_0}(x_0)} |Du|^{2\gamma} dy \right)^{1/\gamma}.$$

In particular, by (4.11) and $R \leq R_0 \leq 1$, we get

$$\int_R^{2^\ell R} \mu^{1-2\tau} \left(\int_{B_\mu(\bar{x})} |Du|^{2\gamma} dy \right)^{1/\gamma} \frac{d\mu}{\mu} \leq c(n, s) \left(\frac{32R_0}{r_2 - r_1} \right)^{n+2} \int_{B_{2R_0}(x_0)} |Du|^2 dy,$$

where we also estimated the last integral using Hölder's inequality, since $\gamma = n/(n+2) < 1$. Again, using the inclusion $B_{2^{\ell+1}R}(\bar{x}) \subseteq B_{2R_0}(x_0)$, we may use (3.19), (4.21) and Poincaré's inequality to deduce

$$\begin{aligned}\nu R \int_{\mathbb{R}^n \setminus B_{2^\ell R}(\bar{x})} \frac{|u(y) - (u)_{B_{2^\ell R}(\bar{x})}|}{|y - \bar{x}|^{n+2s}} dy &= \nu R (2^\ell R)^{-2s} \text{Tail}(u - (u)_{B_{2^\ell R}(\bar{x})}; B_{2^\ell R}(\bar{x})) \\ &\leq c \nu R (2R_0)^{-2s} \left(\frac{2R_0}{2R_0 - |\bar{x} - x_0|} \right)^{n+2s} \text{Tail}(u - (u)_{B_{2R_0}(x_0)}; B_{2R_0}(x_0)) \\ &\quad + c \nu R (2^\ell R)^{-2s} \left(\frac{2R_0}{2^\ell R} \right)^n \int_{B_{2R_0}(x_0)} |u - (u)_{B_{2R_0}(x_0)}| dx \\ &\leq c \nu R \left(\frac{2R_0}{r_2 - r_1} \right)^{n+2s} (2R_0)^{-2s} \text{Tail}(u - (u)_{B_{2R_0}(x_0)}; B_{2R_0}(x_0)) \\ &\quad + c \nu R (2^\ell R)^{-2s} \frac{(2R_0)^{n+1}}{(2^\ell R)^n} \left(\int_{B_{2R_0}(x_0)} |Du|^2 dx \right)^{1/2};\end{aligned}$$

here $c = c(n, N, s)$. Since $2^\ell R \leq R_0 \leq 1$, by (4.11) we estimate $(2^\ell R)^{-2s} \leq (2^\ell R)^{-2\tau} \leq R^{-2\tau}$, whence from the above inequality we find

$$\begin{aligned}
 \nu R \int_{\mathbb{R}^n \setminus B_{2^\ell R}(\bar{x})} \frac{|u(y) - (u)_{B_{2^\ell R}(\bar{x})}|}{|y - \bar{x}|^{n+2s}} dy &\leq c \nu R \left(\frac{2R_0}{r_2 - r_1} \right)^{n+2s} (2R_0)^{-2s} \mathbf{Tail}(u - (u)_{B_{2R_0}(x_0)}; B_{2R_0}(x_0)) \\
 &\quad + c \nu \frac{(2R_0)^{n+1}}{R^{n+2\tau-1}} \left(\int_{B_{2R_0}(x_0)} |Du|^2 dx \right)^{1/2} \\
 &\leq c \nu \left(\frac{2R_0}{r_2 - r_1} \right)^{n+2s} (2R_0)^{1-2s} \mathbf{Tail}(u - (u)_{B_{2R_0}(x_0)}; B_{2R_0}(x_0)) \\
 &\quad + c \nu \left(\frac{32R_0}{r_2 - r_1} \right)^{n+2\tau-1} \left(\int_{B_{2R_0}(x_0)} |Du|^2 dx \right)^{1/2},
 \end{aligned} \tag{4.23}$$

with $c = c(n, N, s)$, where in the last inequality we also used (4.21) and $n+1 > n + (2\tau - 1) > n$ by (4.11). Combining (4.22)–(4.23), we thus find

$$\lambda_0 \leq c M \left(\frac{32R_0}{r_2 - r_1} \right)^{n+2} \Xi_0, \tag{4.24}$$

where $c = c(n, N, s)$, and we have set

$$\begin{aligned}
 \Xi_0 := \left(\int_{B_{2R_0}(x_0)} |\tilde{F}|^{2+\zeta} dx \right)^{1/(2+\zeta)} &+ \left(\int_{B_{2R_0}(x_0)} |Du|^2 dx \right)^{1/2} \\
 &+ \nu (2R_0)^{1-2s} \mathbf{Tail}(u - (u)_{B_{2R_0}(x_0)}; B_{2R_0}(x_0)).
 \end{aligned} \tag{4.25}$$

Step 2: Exit time and Vitali covering. Let $\lambda \geq \lambda_0$, and define

$$E_\lambda := \left\{ \bar{x} \in B_{r_1}(x_0) : \sup_{0 \leq R \leq \frac{r_2 - r_1}{16}} \Psi_M(\bar{x}, R) > \lambda \right\}.$$

By definition (4.20), we have

$$x \in B_{r_1}(x_0), R \in \left[\frac{r_2 - r_1}{16}, \frac{R_0}{2} \right] \implies \Psi_M(x, R) \leq \lambda_0 \leq \lambda.$$

On the other hand, for every $x \in E_\lambda$, there exists an exit radius $R_x \in (0, \frac{r_2 - r_1}{16})$ such that

$$\Psi_M(x, R_x) = \lambda \quad \text{and} \quad \Psi_M(x, R) \leq \lambda \text{ for all } R \in \left(R_x, \frac{R_0}{2} \right). \tag{4.26}$$

By Vitali's covering argument, we may find a countable collection of pairwise disjoint balls $\{B_{2R_j}(x_j)\}_{j \in \mathcal{J}}$, with cardinality $\#\mathcal{J} \leq \#\mathbb{N}$, with centers $\{x_j\}_{j \in \mathcal{J}} \subseteq E_\lambda$ and radii $\{R_j\}_{j \in \mathcal{J}}$ such that

$$E_\lambda \subseteq \bigcup_{j \in \mathcal{J}} B_{10R_j}(x_j) \cup \mathcal{N}, \tag{4.27}$$

with $\mathcal{N}$ negligible. For notational simplicity, we set $B_j = B_{R_j}(x_j)$, and denote by $KB_j = B_{KR_j}(x_j)$ for $K > 0$. For all $j \in \mathcal{J}$, let us pick $\ell_j = \ell(R_j) \in \mathbb{N}$ satisfying

$$\frac{R_0}{2} \leq 2^{\ell_j} R_j \leq R_0.$$

Since $R_j < R_0/8$, it follows that $\ell_j \geq 3$ for all $j \in \mathcal{J}$, and $2^{\ell_j+1} B_j = B_{2^{\ell_j+1} R_j}(x_j) \subseteq B_{2R_0}(x_0)$.

We aim to estimate the measure $|B_j|$. By (4.26), we have that either

$$\left(\int_{B_j} |Du|^2 dx \right)^{1/2} \geq \frac{\lambda}{2} \quad \text{or} \quad M \left(\int_{B_j} |\tilde{F}|^2 dx \right)^{1/2} \geq \frac{\lambda}{2} \tag{4.28}$$

holds true. Assume that the first condition in (4.28) is in force. This information coupled with (4.12) (with $r = 2R_j$) gives

$$\lambda^2 \leq c \left(1 + \frac{\nu^2}{\delta} \right) \left(\int_{2B_j} |Du|^{2\gamma} dx \right)^{1/\gamma} + \delta \left[\nu 2R_j (2^{\ell_j+1} R_j)^{-2s} \mathbf{Tail}(u - (u)_{B_{2^{\ell_j+1} R_j}(x_j)}; B_{2^{\ell_j+1} R_j}(x_j)) \right]^2 \tag{4.29}$$

$$+ \delta \left[\left(\int_{2^{\ell_j} B_j} |Du|^{2\gamma} dx \right)^{1/\gamma} + \int_{2R_j}^{2^{\ell_j+1} R_j} \mu^{1-2\tau} \left(\int_{B_\mu(x_j)} |Du|^{2\gamma} dx \right)^{1/\gamma} \frac{d\mu}{\mu} \right] + c \int_{2B_j} |\tilde{F}|^2 dx;$$

furthermore, thanks to (4.26), for all $\mu \geq R_j$ we have

$$\int_{B_\mu(x_j)} |Du|^2 dy \leq \lambda^2,$$

while, by (4.16), (4.18), (4.20) and the fact that $\lambda \geq \lambda_0$, we have

$$\begin{aligned} \nu 2R_j (2^{\ell_j+1} R_j)^{-2s} \mathbf{Tail}(u - (u)_{2^{\ell_j+1} B_j}; 2^{\ell_j+1} B_j) &= \nu 2R_j \int_{\mathbb{R}^n \setminus 2^{\ell_j+1} B_j} \frac{|u(y) - (u)_{2^{\ell_j+1} B_j}|}{|y - x_j|^{n+2s}} dy \leq \mathcal{T}(x_j, 2R_j) \\ &\leq \Xi(x_j, 2R_j) \leq \lambda, \end{aligned}$$

and

$$\int_{2B_j} |\tilde{F}|^2 dx \leq \frac{[\Xi(x_j, 2R_j)]^2}{M^2} \leq \frac{\lambda^2}{M^2}.$$

Using the last three estimates in (4.29) together with Hölder's inequality, yields

$$\lambda^2 \leq c \left(1 + \frac{\nu^2}{\delta} \right) \left(\int_{2B_j} |Du|^{2\gamma} dx \right)^{1/\gamma} + c \frac{\lambda^2}{M^2} + c \delta \lambda^2.$$

Hence we fix δ small enough and M large enough, both depending only on $\mathbf{data}, \max\{\nu, 1\}$, so as to obtain

$$\lambda^2 \leq c(1 + \nu)^2 \left(\int_{2B_j} |Du|^{2\gamma} dx \right)^{1/\gamma},$$

from which it follows, for any $\kappa > 0$,

$$(4.30) \quad |B_j| \leq \frac{c(1 + \nu)^{2\gamma}}{\lambda^{2\gamma}} \int_{2B_j} |Du|^{2\gamma} dx \leq \frac{c(1 + \nu)^{2\gamma}}{\lambda^{2\gamma}} \int_{2B_j \cap \{|Du| > \kappa \lambda\}} |Du|^{2\gamma} dx + c(1 + \nu)^{2\gamma} \kappa^{2\gamma} |B_j|;$$

by choosing $\kappa = \kappa(\mathbf{data}, \max\{\nu, 1\}) \in (0, 1)$ so that $c(1 + \nu)^{2\gamma} \kappa^{2\gamma} \leq 1/2$, from (4.30) we get

$$(4.31) \quad |B_j| \leq \frac{c(1 + \nu)^{2\gamma}}{\lambda^{2\gamma}} \int_{2B_j \cap \{|Du| > \kappa \lambda\}} |Du|^{2\gamma} dx$$

with $c \equiv c(\mathbf{data})$. If, on the other hand, the second condition in (4.28) holds true, then trivially (recall that now M is fixed depending on the $\mathbf{data}$)

$$|B_j| \leq \frac{c}{\lambda^2} \int_{B_j} |\tilde{F}|^2 dx,$$

and arguing as before we find $\kappa = \kappa(\mathbf{data}, \max\{\nu, 1\}) \in (0, 1)$ small enough such that

$$(4.32) \quad |B_j| \leq \frac{c}{\lambda^2} \int_{B_j \cap \{|\tilde{F}| > \kappa \lambda\}} |\tilde{F}|^2 dx$$

(clearly we can suppose that the two small constants κ in (4.31) and (4.32) are the same). In any case, from (4.31) and (4.32) we deduce

$$(4.33) \quad |B_j| \leq \frac{c(1 + \nu)^{2\gamma}}{\lambda^{2\gamma}} \int_{2B_j \cap \{|Du| > \kappa \lambda\}} |Du|^{2\gamma} dx + \frac{c}{\lambda^2} \int_{2B_j \cap \{|\tilde{F}| > \kappa \lambda\}} |\tilde{F}|^2 dx.$$

From this point the proof is standard, thus we only sketch the main points; for the details see [2], [8, Section 8], [15, Section 5.2] with the appropriate modifications. We observe that (4.27) implies

$$(4.34) \quad B_{r_1}(x_0) \cap \{|Du| > \lambda\} \subseteq E_\lambda \subseteq \bigcup_{j \in \mathcal{J}} 10B_j \quad \text{up to a negligible set,}$$

and this can be easily proven considering Lebesgue points of the gradient in $B_{r_1}(x_0) \cap \{|Du| > \lambda\}$.

Step 3: The λ -inequality. From now on we denote B_{r_1}, B_{r_2} for short, omitting their center x_0 . As $\{2B_j\}$ is a pairwise disjoint family and $\bigcup_{j \in \mathcal{J}} 2B_j \subseteq B_{r_2}$, summing (4.33) over $j \in \mathcal{J}$ yields

$$(4.35) \quad \sum_{j \in \mathcal{J}} |B_j| \leq \frac{c(1+\nu)^{2\gamma}}{\lambda^{2\gamma}} \int_{B_{r_2} \cap \{|Du| > \kappa\lambda\}} |Du|^{2\gamma} dx + \frac{c}{\lambda^2} \int_{B_{r_2} \cap \{|\tilde{F}| > \kappa\lambda\}} |\tilde{F}|^2 dx.$$

On the other hand, by (4.26), we have $\Psi_M(x_j, R) \leq \lambda$ if $R \geq R_j$, hence

$$(4.36) \quad \int_{10B_j} |Du|^2 dx \leq \lambda^2 |10B_j| \quad \text{for all } j \in \mathcal{J};$$

from (4.34)–(4.36) we infer

$$\int_{B_{r_1} \cap \{|Du| > \lambda\}} |Du|^2 dx \leq \sum_{j \in \mathcal{J}} \int_{10B_j} |Du|^2 dx \leq 10^n \lambda^2 \sum_{j \in \mathcal{J}} |B_j|$$

and combining the latter inequality with (4.35), we obtain

$$\int_{B_{r_1} \cap \{|Du| > \lambda\}} |Du|^2 dx \leq \frac{c(1+\nu)^2}{\lambda^{2\gamma-2}} \int_{B_{r_2} \cap \{|Du| > \kappa\lambda\}} |Du|^{2\gamma} dx + c \int_{B_{r_2} \cap \{|\tilde{F}| > \kappa\lambda\}} |\tilde{F}|^2 dx,$$

that is

$$(4.37) \quad \int_{\mathcal{U}_\lambda(r_1)} |Du|^2 dx \leq c \lambda^{2-2\gamma} \int_{\mathcal{U}_{\kappa\lambda}(r_2)} |Du|^{2\gamma} dx + c \int_{\mathcal{F}_{\kappa\lambda}(r_2)} |\tilde{F}|^2 dx$$

(where now c may also depend on $\max\{\nu, 1\}$) where we have introduced the sets

$$\mathcal{U}_\lambda(r) = \{x \in B_r(x_0) : |Du|(x) > \lambda\}, \quad \mathcal{F}_\lambda(r) = \{x \in B_r(x_0) : |\tilde{F}| > \lambda\}.$$

Step 4: Integration and conclusion. To ensure that all the quantities we are considering are finite, for any $t \geq 0$ we should define the truncated functions

$$|Du|_t = \min\{|Du|, t\} \quad \text{and} \quad |\tilde{F}|_t = \min\{|\tilde{F}|, t\}$$

and proceed as in [8]; we shall instead proceed formally and let the reader follow [8, Section 8] to implement the necessary changes needed to make the proof rigorous. Let us multiply (4.37) by $\lambda^{\zeta-1}$, where $\zeta > 0$ is as in the statement, and then integrate it with respect to λ over $(\lambda_0, +\infty)$ to obtain, after changing variables in the two integrals on the right-hand side,

$$\begin{aligned} & \int_{\lambda_0}^{\infty} \lambda^{\zeta} \int_{\mathcal{U}_\lambda(r_1)} |Du|^2 dx \frac{d\lambda}{\lambda} \\ & \leq \frac{c}{\kappa^{\zeta+2-2\gamma}} \int_{\kappa\lambda_0}^{\infty} \lambda^{\zeta+2(1-\gamma)} \int_{\mathcal{U}_\lambda(r_2)} |Du|^{2\gamma} dx \frac{d\lambda}{\lambda} + \frac{c}{\kappa^{\zeta}} \int_{\kappa\lambda_0}^{\infty} \lambda^{\zeta} \int_{\mathcal{F}_\lambda(r_2)} |\tilde{F}|^2 dx \frac{d\lambda}{\lambda} \\ & \leq \frac{c}{[\zeta + 2(1-\gamma)]\kappa^{\zeta}} \int_{B_{r_2}} |Du|^{2+\zeta} dx + \frac{c}{\zeta\kappa^{\zeta}} \int_{B_{r_2}} |\tilde{F}|^{2+\zeta} dx \end{aligned}$$

with c, κ depending on **data** and $\max\{\nu, 1\}$; in the last estimate we used Fubini's layer formula twice. On the other hand, recalling (4.24)–(4.25), we have

$$\int_0^{\lambda_0} \lambda^{\zeta} \int_{\mathcal{U}_\lambda(r_1)} |Du|^2 dx \frac{d\lambda}{\lambda} \leq \frac{\lambda_0^{\zeta}}{\zeta} \int_{B_{r_2}} |Du|^2 dx \leq \frac{c}{\zeta} \left(\frac{R_0}{r_2 - r_1} \right)^{\zeta(n+2)} |B_{R_0}| \Xi_0^{\zeta+2}.$$

Combining these two estimates and again using Fubini's formula yields

$$\begin{aligned} & \int_{B_{r_1}} |Du|^{2+\zeta} dx = \zeta \int_0^{+\infty} \lambda^{\zeta} \int_{\mathcal{U}_\lambda(r_1)} |Du|^2 dx \frac{d\lambda}{\lambda} \\ & \leq \frac{c\zeta}{[\zeta + 2(1-\gamma)]\kappa^{\zeta}} \int_{B_{r_2}} |Du|^{2+\zeta} dx + \frac{c}{\kappa^{\zeta}} \int_{B_{R_0}} |\tilde{F}|^{2+\zeta} dx + c \left(\frac{R_0}{r_2 - r_1} \right)^{\zeta(n+2)} |B_{R_0}| \Xi_0^{\zeta+2}; \end{aligned}$$

this inequality is valid for any r_1, r_2 fulfilling (4.19). Finally we conclude by choosing $\zeta > 0$ depending on **data**, $\max\{\nu, 1\}$ such that

$$\frac{c\zeta}{[\zeta + 2(1-\gamma)]\kappa^{\zeta}} \leq \frac{1}{2},$$

as we are in the position to apply the aforementioned [37, Lemma 6.1] to reabsorb $\int_{B_{r_2}} |Du|^{2+\zeta} dx$; this gives, after dividing by $|B_{R_0}|$, and recalling the definition of Ξ_0 in (4.25), the desired estimate. $\square$

5. LINEARIZATION AND EXCESS DECAY

In this section, we fix $B_\varrho(x_0) \Subset \Omega$ and we assume that u satisfies

$$(5.1) \quad \mathcal{E}_\gamma(u; B_\varrho(x_0)) \leq \varepsilon_0$$

for some $\varepsilon_0 = \varepsilon_0(\mathbf{data}, \omega_u(\cdot)) \in (0, 1)$ to be determined later, where $\mathcal{E}_\gamma(\cdot)$ is the γ -excess defined in (3.21). We also assume that $A(\cdot, \cdot)$ fulfills the *BMO*-condition

$$(5.2) \quad E_{r_A}(A; x_0) \leq \delta,$$

for some radius $r_A \in (0, 1)$, and $\delta = \delta(\mathbf{data}) \in (0, 1)$ to be chosen later, where $E_r(A; x_0)$ is defined in (2.6).

We carry out a blow-up procedure at a point $x_0 \in \Omega$ within the ball $B_\varrho(x_0)$, where $\varrho \leq \min\{r_0, r_A\}$. Here r_A is the radius from (5.2) and $r_0 = r_0(\mathbf{data}, \nu) \in (0, 1)$ satisfies (4.4); r_0 will be further reduced in due course of the proof.

In this section, unlike in the previous ones, we will always make the center of the balls explicit when it is x_0 and we will use the notation $B_{1/4}, B_1$, etc., for balls centered at the origin.

From now on, we fix a parameter

$$(5.3) \quad \gamma = \gamma(s) \text{ such that } \max\{2s, 1\} < \gamma < 2,$$

and for $\mathcal{E}_\gamma \equiv \mathcal{E}_\gamma(u; B_\varrho(x_0))$, we set

$$(5.4) \quad u_\varrho(x) := \frac{u(x_0 + \varrho x) - (u)_{B_\varrho(x_0)}}{\mathcal{E}_\gamma}, \quad f_\varrho(x) := \frac{\varrho^2 f(x_0 + \varrho x)}{\mathcal{E}_\gamma}.$$

Note that

$$(5.5) \quad (u_\varrho)_{B_1} = 0, \quad \int_{B_1} |u_\varrho|^2 dx \leq 1$$

and, by a change of variables, we have

$$(5.6) \quad \text{Tail}_\gamma(u_\varrho; B_1) = \text{Tail}(u_\varrho; B_1) = \frac{1}{\mathcal{E}_\gamma} \text{Tail}(u - (u)_{B_\varrho(x_0)}; B_\varrho(x_0)) \leq \varrho^{2s-\gamma};$$

the last inequality is due to the definition of γ -excess in (3.21) and (3.14). By setting

$$A_\varrho(x, z) = A(x_0 + \varrho x, (u)_{B_\varrho(x_0)} + \mathcal{E}_\gamma z)$$

and

$$\Phi_\varrho(x, y, v, w, t) = \frac{\Phi(x_0 + \varrho x, x_0 + \varrho y, (u)_{B_\varrho(x_0)} + \mathcal{E}_\gamma v, (u)_{B_\varrho(x_0)} + \mathcal{E}_\gamma w, \mathcal{E}_\gamma t)}{\mathcal{E}_\gamma}, \quad Q_\varrho^{\text{nl}} \equiv Q_{\Phi_\varrho}^{\text{nl}},$$

straightforward computations (via change of variables) show that u_ϱ is a weak solution to

$$(5.7) \quad -\text{div}(A_\varrho(\cdot, u_\varrho) Du_\varrho) + \varrho^{2(1-s)} Q_\varrho^{\text{nl}} u_\varrho = f_\varrho \quad \text{in } B_1.$$

Furthermore, it is straightforward to verify that $A_\varrho(\cdot, \cdot)$ still satisfies (1.3), and that Φ_ϱ continues to satisfy (1.4)–(1.5), that is

$$(5.8) \quad \langle A_\varrho(x, z) \xi, \xi \rangle \geq \Lambda^{-1} |\xi|^2, \quad |A_\varrho(x, z)| \leq \Lambda \quad \text{for a.e. } x \in B_1 \text{ and } z \in \mathbb{R}^N,$$

and, for a.e. $x, y \in \mathbb{R}^n$, all $u, v \in \mathbb{R}^N$ and $t \in \mathbb{R}^N$

$$(5.9) \quad \begin{cases} \Phi_\varrho(x, y, u, v, t) = \Phi_\varrho(y, x, v, u, t) \\ \Phi_\varrho(x, y, u, v, -t) = -\Phi_\varrho(x, y, u, v, t) \end{cases}, \quad |\Phi_\varrho(x, y, v, w, t)| \leq \nu |t|.$$

In addition, if $A(\cdot, \cdot)$ satisfies (5.2), then a simple change of variables shows that $A_\varrho(\cdot, \cdot)$ also fulfills

$$(5.10) \quad \left(\int_{B_\sigma} |A_\varrho(x, z) - (A_\varrho)_{B_\sigma}(z)|^2 dx \right)^{1/2} \leq \delta, \quad \text{for all } \sigma \leq r_A/\varrho \text{ and } z \in \mathbb{R}^N.$$

Moreover, by (2.4), we have that $u \mapsto A_\varrho(x, u)$ is uniformly continuous (uniformly in x), with modulus of continuity given by

$$(5.11) \quad \omega_{u, \varrho}(t) := \omega_u(\mathcal{E}_\gamma t) \quad \text{for } t \geq 0.$$

Remark 5.1. From now on, possibly reducing r_0 , we will always assume that

$$(5.12) \quad c(\mathbf{data})r_0^{2(1-s)}\nu \leq 1 \quad \text{and} \quad r_0^{2-\gamma}\nu \leq 1,$$

where $c(\mathbf{data}) \geq 1$ is the constant appearing in (4.4). This choice is possible due to the bound in (5.3). Consequently, all quantitative constants from the previous sections (in particular, those in the Caccioppoli inequality of Lemma 4.1 and the gradient higher integrability of Proposition 4.4) applied to u_ϱ will be independent of ϱ . Indeed, (5.8) holds, and the nonlocal term Φ_ϱ in (5.7) satisfies (5.9) while being multiplied by $\varrho^{2(1-s)}$, yielding

$$(5.13) \quad \varrho^{2(1-s)}|\Phi_\varrho(x, y, u, v, t)| \leq r_0^{2(1-s)}\nu|t| \leq |t|, \quad \text{for all } t \in \mathbb{R}.$$

Moreover, for the rescaled equation, the threshold radius appearing in (4.4) effectively becomes 1 for any choice of $\varrho \leq r_0$; this allows us to apply (4.10) directly to the function u_ϱ .

5.1. Small potential regime. In this paragraph, we further assume that we are in the so-called *small potential regime*, which means that

$$(5.14) \quad \varrho^2 \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} < \varepsilon^* \mathcal{E}_\gamma(u; B_\varrho(x_0))$$

for some constant $\varepsilon^* \in (0, 1)$ depending on $\mathbf{data}$ to be determined later. In particular, by definition of f_ϱ in (5.4), this implies

$$(5.15) \quad \left(\int_{B_1} |f_\varrho|^\chi dx \right)^{1/\chi} \leq \varepsilon^* < 1.$$

We claim that, under assumption (5.14), there exists $\delta_1 = \delta_1(\mathbf{data})$ such that the higher integrability property

$$(5.16) \quad \int_{B_{1/2}} |Du_\varrho|^2 dx + \int_{B_{1/2}} |Du_\varrho|^{2+\delta_1} dx \leq c(\mathbf{data})$$

holds true if $\varrho \leq r_0$. To prove it, let us fix $y \in B_{1/2}$; since u_ϱ solves (5.7) with (5.8) and (5.9)–(5.13) in force, from Caccioppoli inequality (4.3) with $u_0 = 0$ and ν replaced by $\varrho^{2(1-s)}\nu \leq 1$, we obtain the energy inequality

$$(5.17) \quad \begin{aligned} & \int_{B_{1/8}(y)} |Du_\varrho|^2 dx \\ & \leq c \int_{B_{1/4}(y)} |u_\varrho|^2 dx + c \varrho^{2(1-s)} \nu \mathbf{Tail}(u_\varrho; B_{1/4}(y)) \int_{B_{1/4}(y)} |u_\varrho| dx + c \left(\int_{B_{1/4}(y)} |f_\varrho|^\chi dx \right)^{2/\chi} + c, \end{aligned}$$

with $c = c(\mathbf{data})$ due to Remark 5.1. Next observe that, for any $\sigma \in (1/32, 1/4)$ and all $y \in B_{1/2}$, we have $B_\sigma(y) \subseteq B_1$. It then follows by (3.20) and (5.5)–(5.6) that

$$\begin{aligned} \varrho^{2(1-s)} \nu \mathbf{Tail}(u_\varrho; B_\sigma(y)) & \leq c(n, s) \varrho^{2(1-s)} \nu \left[\mathbf{Tail}(u_\varrho; B_1) + \int_{B_1} |u_\varrho| dx \right] \\ & \leq c(n, s) \varrho^{2-\gamma} \nu \leq c(n, s) r_0^{2-\gamma} \nu \leq c(n, s), \end{aligned}$$

where the last inequality due to (5.12); a similar estimate holds if we replace u_ϱ with $u_\varrho - (u_\varrho)_{B_{1/8}(y)}$, using this time (3.19) and (3.15)–(3.16). Therefore, using Hölder's inequality, (5.15), (5.5) and the last estimate, from (5.17) we infer

$$\int_{B_{1/8}(y)} |Du_\varrho|^2 dx \leq c(\mathbf{data}).$$

Then, using (4.10) (with ν replaced by $\varrho^{2(1-s)}\nu$ by (5.13)) and (4.9), if we set $\delta_1 = \min\{\zeta, \epsilon\}$ (which only depends on $\mathbf{data}$ thanks to Remark 5.1), we have

$$\left(\int_{B_{1/32}(y)} |Du_\varrho|^{2+\delta_1} dx \right)^{1/(2+\delta_1)} \leq c \left(\int_{B_{1/8}(y)} |Du_\varrho|^2 dx \right)^{1/2} + c + c \left(\int_{B_1} |f_\varrho|^\chi dx \right)^{1/\chi}$$

$$+ c \varrho^{2(1-s)} \nu \mathbf{Tail}(u_\varrho - (u_\varrho)_{B_{1/8}(y)}; B_{1/8}(y)),$$

which, estimating as before, yields

$$(5.18) \quad \int_{B_{1/32}(y)} |Du_\varrho|^{2+\delta_1} dx \leq c(\mathbf{data}).$$

By the arbitrariness of $y \in B_{1/2}$, (5.16)) can be obtained from (5.18) and a standard covering argument.

It is useful to introduce the following notation: for Φ a kernel as in (1.4), and $\varphi \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^N)$, we write

$$(5.19) \quad \langle\langle Q_\Phi^{\text{nl}} u, \varphi \rangle\rangle_{\text{nl}} := \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \Phi(x, y, u(x), u(y), u(x) - u(y)) \cdot (\varphi(x) - \varphi(y)) \frac{dx dy}{|x - y|^{n+2s}}.$$

In view of the quantitative approximation lemma, the next lemma shows the core of our paper: namely, it shows why the nonlocal term can be treated as a perturbative term, and it can be made arbitrarily small by taking ϱ suitably small.

Lemma 5.2. *There exists a constant c , depending only on $\mathbf{data}$, such that if $\gamma = \gamma(s)$ is given by (5.3), and (5.15) is in force, then*

$$(5.20) \quad \left| \varrho^{2(1-s)} \langle\langle Q_\varrho^{\text{nl}} u_\varrho, \varphi \rangle\rangle_{\text{nl}} \right| \leq c \nu \varrho^{2-\gamma} \|D\varphi\|_{L^\infty(B_{1/4})},$$

for all functions $\varphi \in C_c^\infty(B_{1/4}; \mathbb{R}^N)$.

Proof. Let φ be as in the statement: by the symmetry property in (1.5),

$$\begin{aligned} \langle\langle Q_\varrho^{\text{nl}} u_\varrho, \varphi \rangle\rangle_{\text{nl}} &= \int_{B_{1/4}} \int_{B_{1/4}} \Phi_\varrho(x, y, u_\varrho(x), u_\varrho(y), u_\varrho(x) - u_\varrho(y)) \cdot (\varphi(x) - \varphi(y)) \frac{dx dy}{|x - y|^{n+2s}} \\ &\quad + 2 \int_{\mathbb{R}^n \setminus B_{1/4}} \int_{B_{1/4}} \Phi_\varrho(x, y, u_\varrho(x), u_\varrho(y), u_\varrho(x) - u_\varrho(y)) \cdot \varphi(x) \frac{dx dy}{|x - y|^{n+2s}} := (I) + (II). \end{aligned}$$

Using (5.9), the triangle inequality and (3.6), we have

$$|(I)| \leq \nu [u_\varrho]_{W^{s,2}(B_{1/4})} [\varphi]_{W^{s,2}(B_{1/4})} \leq c(n, s) \nu \|Du_\varrho\|_{L^2(B_{1/4})} \|D\varphi\|_{L^2(B_{1/4})} \leq c \nu \|D\varphi\|_{L^\infty(B_{1/4})},$$

where the last inequality follows from (5.16) and the constant c depends only on $\mathbf{data}$. Next we observe that for $x \in B_{1/4}$ and $y \in \mathbb{R}^n \setminus B_{1/2}$, we have $2|x - y| \geq |y|$; by using this information, the triangle inequality and again (5.9), we get, for a constant c ultimately depending only on $\mathbf{data}$,

$$\begin{aligned} |(II)| &\leq c \nu \left[\int_{\mathbb{R}^n \setminus B_{1/2}} \frac{dy}{|y|^{n+2s}} \int_{B_{1/4}} |u_\varrho(x)| |\varphi(x)| dx + \int_{\mathbb{R}^n \setminus B_{1/2}} \frac{|u_\varrho(y)|}{|y|^{n+2s}} dy \int_{B_{1/4}} |\varphi(x)| dx \right] \\ &\leq c \nu \left[\left(\int_{B_1} |u_\varrho|^2 dx \right)^{1/2} \left(\int_{B_{1/4}} |\varphi|^2 dx \right)^{1/2} + \int_{\mathbb{R}^n \setminus B_{1/2}} \frac{|u_\varrho(y)|}{|y|^{n+2s}} dy \int_{B_{1/4}} |\varphi(x)| dx \right] \\ &\leq c \nu \left[\left(\int_{B_1} |u_\varrho|^2 dx \right)^{1/2} + \int_{\mathbb{R}^n \setminus B_{1/2}} \frac{|u_\varrho(y)|}{|y|^{n+2s}} dy \right] \|D\varphi\|_{L^\infty(B_{1/4})}. \end{aligned}$$

We conclude the proof of (5.20) by estimating the first term between parentheses by (5.5) and the second one as follows, using (5.6):

$$\begin{aligned} \int_{\mathbb{R}^n \setminus B_{1/2}} \frac{|u_\varrho(y)|}{|y|^{n+2s}} dy &= \mathbf{Tail}(u_\varrho; B_1) + \int_{B_1 \setminus B_{1/2}} \frac{|u_\varrho(y)|}{|y|^{n+2s}} dy \\ &\leq \varrho^{2s-\gamma} + 2^{n+2s} \int_{B_1} |u_\varrho(y)| dy \leq c(n) (\varrho^{2s-\gamma} + 1). \end{aligned}$$

Multiplying the estimates above by $\varrho^{2(1-s)}$ and recalling that $\varrho^{2s-\gamma} > 1$, $\varrho^{2(1-s)} < \varrho^{2-\gamma}$ by (5.3) yields (5.20). $\square$

5.2. Approximate harmonicity. With Lemma 5.2 at our disposal, we now want to apply Lemma 3.5 with

$$(5.21) \quad \mathbf{A} \equiv \bar{A}_\varrho := (A_\varrho)_{B_1}(0),$$

where $(A_\varrho)_{B_\sigma}(z)$ was defined in (2.5); $\bar{A}_\varrho$ clearly satisfies (3.11) by (5.8). To this end, we first prove the following

Proposition 5.3 (Quantified approximate harmonicity). *Let u_ϱ, A_ϱ be given by (5.4) and (5.21), respectively. Assume that (5.1), (5.3) and (5.14) are in force. Then we have*

$$(5.22) \quad \left| \int_{B_{1/4}} \langle \bar{A}_\varrho Du_\varrho, D\varphi \rangle dx \right| \leq c s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta) \|D\varphi\|_{L^\infty(B_{1/4})}$$

for all $\varphi \in C_c^\infty(B_{1/4}; \mathbb{R}^N)$, where $c = c(\mathbf{data})$ and

$$(5.23) \quad s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta) = \nu \varrho^{2-\gamma} + \varepsilon^* + [\omega_u(\varepsilon_0)]^{1/2} + \delta.$$

Proof. Given $\varphi \in C_c^\infty(B_{1/4}; \mathbb{R}^N)$, we have for $X \in B_{1/4}$

$$\langle \bar{A}_\varrho Du_\varrho(x), D\varphi(x) \rangle = \langle A_\varrho(x, u_\varrho(x)) Du_\varrho(x), D\varphi(x) \rangle + \langle (\bar{A}_\varrho - A_\varrho(x, u_\varrho(x))) Du_\varrho(x), D\varphi(x) \rangle,$$

so that, integrating over $B_{1/4}$ and using the weak formulation (5.7), we obtain

$$\begin{aligned} \left| \int_{B_{1/4}} \langle \bar{A}_\varrho Du_\varrho, D\varphi \rangle dx \right| &\leq \left| \varrho^{2(1-s)} \langle Q_\varrho^{\text{nl}} u_\varrho, \varphi \rangle_{\text{nl}} \right| + \left| \int_{B_{1/4}} f_\varrho \cdot \varphi dx \right| \\ &\quad + \int_{B_{1/4}} |\bar{A}_\varrho - A_\varrho(x, u_\varrho)| |Du_\varrho| |D\varphi| dx = (I) + (II) + (III). \end{aligned}$$

The term (I) is estimated using Lemma 5.2. Moreover, by Hölder's inequality, (5.15), and since $\|\varphi\|_{L^\infty(B_{1/4})} \leq \|D\varphi\|_{L^\infty(B_{1/4})}$ being φ compactly supported, we have

$$(5.24) \quad (II) \leq \|f_\varrho\|_{L^X(B_1)} \|\varphi\|_{L^\infty(B_{1/4})} \leq c(n) \varepsilon^* \|D\varphi\|_{L^\infty(B_{1/4})};$$

for the last term we use the triangle inequality and split

$$(III) \leq \int_{B_{1/4}} |\bar{A}_\varrho - A_\varrho(x, 0)| |Du_\varrho| |D\varphi| dx + \int_{B_{1/4}} |A_\varrho(x, 0) - A_\varrho(x, u_\varrho)| |Du_\varrho| |D\varphi| dx =: (IV) + (V).$$

By Hölder's inequality, (5.10) and (5.21) (recalling that $1/4 \leq r_A/\varrho$ as $\varrho \leq r_0(\mathbf{data}, \nu, r_A) \leq r_A$), and by (5.16) we infer

$$\begin{aligned} (IV) &\leq \left(\int_{B_{1/4}} |\bar{A}_\varrho - A_\varrho(x, 0)|^2 dx \right)^{1/2} \|Du_\varrho\|_{L^2(B_{1/4})} \|D\varphi\|_{L^\infty(B_{1/4})} \\ &\leq c \left(\int_{B_{1/4}} |\bar{A}_\varrho - A_\varrho(x, 0)|^2 dx \right)^{1/2} \|D\varphi\|_{L^\infty(B_{1/4})} \leq c \delta \|D\varphi\|_{L^\infty(B_{1/4})}, \end{aligned}$$

with $c = c(\mathbf{data})$. Then, by the uniform continuity of A_ϱ (5.11), Hölder's and Jensen's inequalities, (5.16), the boundedness of $\omega_{u,\varrho}(\cdot) \leq 1$, (5.5) and then (5.1), we deduce

$$\begin{aligned} (V) &\leq c \|D\varphi\|_{L^\infty(B_{1/4})} \int_{B_{1/4}} \omega_{u,\varrho}(|u_\varrho|) |Du_\varrho| dx \\ &\leq c \left(\int_{B_1} \omega_{u,\varrho}(|u_\varrho|)^2 dx \right)^{1/2} \|Du_\varrho\|_{L^2(B_{1/4})} \|D\varphi\|_{L^\infty(B_{1/4})} \\ &\leq c \left[\omega_{u,\varrho} \left(\int_{B_1} |u_\varrho| dy \right) \right]^{1/2} \|D\varphi\|_{L^\infty(B_{1/4})} \\ &\leq c \left[\omega_u \left(\mathcal{E}_\gamma(u; B_\varrho(x_0)) \right) \right]^{1/2} \|D\varphi\|_{L^\infty(B_{1/4})} \leq c [\omega_u(\varepsilon_0)]^{1/2} \|D\varphi\|_{L^\infty(B_{1/4})}, \end{aligned}$$

where $c = c(\mathbf{data})$. Combining all these estimates, we get the desired estimate. $\square$

By the previous proposition, we are ready to apply the approximate harmonicity lemma 3.5: owing to (5.16) and (5.22), there exist a small exponent $\vartheta = \vartheta(\mathbf{data}) \in (0, 1]$ and $c = c(\mathbf{data})$ such that

$$(5.25) \quad \int_{B_{1/4}} |Du_\varrho - Dv_\varrho|^2 dx \leq c[s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^\vartheta,$$

for all $0 < \varrho < r_0(\mathbf{data}, \nu, r_A)$, where $v_\varrho \in u_\varrho + W_0^{1,2}(B_{1/4}; \mathbb{R}^N)$ is the unique solution to

$$(5.26) \quad \begin{cases} -\operatorname{div}(\bar{A}_\varrho Dv_\varrho) = 0 & \text{in } B_{1/4} \\ v_\varrho = u_\varrho & \text{on } \partial B_{1/4} \end{cases};$$

here $\bar{A}_\varrho$ is given by (5.21). By standard elliptic regularity theory (see [34, Remark 2.2, Chapter III]) and (5.16), for all $k \in \mathbb{N}$ we have

$$(5.27) \quad \|D^k v_\varrho\|_{L^\infty(B_{1/8})} \leq c(k, \Lambda) \|Dv_\varrho\|_{L^2(B_{1/4})} \leq c(k, \Lambda) \|Du_\varrho\|_{L^2(B_{1/4})} \leq c(k, \mathbf{data});$$

the second inequality comes from testing the weak formulation of (5.26)₁ with $u_\varrho - v_\varrho$, and performing standard computations, using the ellipticity and growth assumptions (5.8).

Now let $\tau \in (0, 1/64)$ to be fixed later, and estimate $\mathcal{E}_\gamma(u_\varrho; B_\tau)$ by considering separately its two components. By (5.25) and (5.27), we find

$$(5.28) \quad \begin{aligned} \left(\int_{B_\tau} |Du_\varrho|^2 dx \right)^{1/2} &\leq \left(\int_{B_\tau} |Du_\varrho - Dv_\varrho|^2 dx \right)^{1/2} + \left(\int_{B_\tau} |Dv_\varrho|^2 dx \right)^{1/2} \\ &\leq c(n) \tau^{-n/2} \left(\int_{B_{1/4}} |Du_\varrho - Dv_\varrho|^2 dx \right)^{1/2} + \|Dv_\varrho\|_{L^\infty(B_{1/8})} \\ &\leq c \tau^{-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} + c, \end{aligned}$$

with $c = c(\mathbf{data})$. Moreover, by Poincaré's inequality and a change of variables,

$$\frac{1}{\mathcal{E}_\gamma(u; B_\varrho(x_0))} \left(\int_{B_{\tau\varrho}(x_0)} |u - (u)_{B_{\tau\varrho}(x_0)}|^2 dx \right)^{1/2} = \left(\int_{B_\tau} |u_\varrho - (u_\varrho)_{B_\tau}|^2 dx \right)^{1/2} \leq c \tau \left(\int_{B_\tau} |Du_\varrho|^2 dx \right)^{1/2},$$

$c = c(n)$, which, coupled with (5.28) implies

$$(5.29) \quad \left(\int_{B_{\tau\varrho}(x_0)} |u - (u)_{B_{\tau\varrho}(x_0)}|^2 dx \right)^{1/2} \leq c \tau \left\{ \tau^{-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} + 1 \right\} \mathcal{E}_\gamma(u; B_\varrho(x_0)),$$

again with c depending on $\mathbf{data}$. Next, using (3.18) with $w = u_\varrho$ (remember that $(u_\varrho)_{B_1} = 0$), (5.5)–(5.6), we may estimate the γ -tail term as follows:

$$(5.30) \quad \begin{aligned} \mathbf{Tail}_\gamma(u_\varrho - (u_\varrho)_{B_\tau}; B_\tau) &\leq c \tau^\gamma \mathbf{Tail}_\gamma(u_\varrho; B_1) + c \tau^\gamma \int_{B_1} |u_\varrho| dx + c \int_\tau^1 \frac{\tau^\gamma}{\mu^{2s}} \int_{B_\mu} |u_\varrho - (u_\varrho)_{B_\mu}| dx \frac{d\mu}{\mu} \\ &\leq c \tau^\gamma \varrho^{2s-\gamma} + c \tau^\gamma \left[1 + \int_\tau^1 \mu^{-2s} \int_{B_\mu} |u_\varrho - (u_\varrho)_{B_\mu}| dx \frac{d\mu}{\mu} \right]. \end{aligned}$$

We estimate separately the last integral, thanks to the triangle, Hölder's and Poincaré's inequalities, (5.5), (5.25) and (5.27):

$$\begin{aligned} \int_\tau^1 \mu^{-2s} \int_{B_\mu} |u_\varrho - (u_\varrho)_{B_\mu}| dx \frac{d\mu}{\mu} &\leq \int_\tau^{1/16} \mu^{-2s} \int_{B_\mu} |u_\varrho - (u_\varrho)_{B_\mu}| dx \frac{d\mu}{\mu} + 2 \int_{1/16}^1 \mu^{-2s} \int_{B_\mu} |u_\varrho| dx \frac{d\mu}{\mu} \\ &\leq c(n) \int_\tau^{1/16} \frac{1}{\mu^{2s-1}} \left(\int_{B_\mu} |Du_\varrho - Dv_\varrho|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \\ &\quad + c(n) \int_\tau^{1/16} \frac{1}{\mu^{2s-1}} \left(\int_{B_\mu} |Dv_\varrho|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \\ &\quad + 2 \int_{1/16}^1 \mu^{-2s} \left(\int_{B_\mu} |u_\varrho|^2 dx \right)^{1/2} \frac{d\mu}{\mu} \end{aligned}$$

$$\begin{aligned}
&\leq c [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} \int_{\tau}^{1/16} \mu^{1-2s-n/2} \frac{d\mu}{\mu} \\
&\quad + c \|Dv_{\varrho}\|_{L^{\infty}(B_{1/8})} \int_{\tau}^{1/16} \mu^{1-2s} \frac{d\mu}{\mu} + c \int_{1/16}^1 \mu^{-2s-n/2} \frac{d\mu}{\mu} \\
(5.31) \quad &\leq c(\mathbf{data}) \left\{ \tau^{1-2s-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} + \tau^{1-2s} + 1 \right\}.
\end{aligned}$$

On the other hand, using (3.14) and a change of variables, we have

$$\begin{aligned}
\mathbf{Tail}_{\gamma}(u_{\varrho} - (u_{\varrho})_{B_{\tau}}; B_{\tau}) &= \tau^{\gamma-2s} \mathbf{Tail}(u_{\varrho} - (u_{\varrho})_{B_{\tau}}; B_{\tau}) = \frac{\tau^{\gamma-2s}}{\mathcal{E}_{\gamma}(u; B_{\varrho}(x_0))} \mathbf{Tail}(u - (u)_{B_{\tau\varrho}(x_0)}; B_{\tau\varrho}(x_0)) \\
(5.32) \quad &= \frac{\varrho^{2s-\gamma}}{\mathcal{E}_{\gamma}(u; B_{\varrho}(x_0))} \mathbf{Tail}_{\gamma}(u - (u)_{B_{\tau\varrho}(x_0)}; B_{\tau\varrho}(x_0)).
\end{aligned}$$

Plugging (5.31) into (5.30) and in turn merging with (5.32) yields

$$\begin{aligned}
\mathbf{Tail}_{\gamma}(u - (u)_{B_{\tau\varrho}(x_0)}; B_{\tau\varrho}(x_0)) &= \frac{\mathcal{E}_{\gamma}(u; B_{\varrho}(x_0))}{\varrho^{2s-\gamma}} \mathbf{Tail}_{\gamma}(u_{\varrho} - (u_{\varrho})_{B_{\tau}}; B_{\tau}) \\
&\leq c \frac{\mathcal{E}_{\gamma}(u; B_{\varrho}(x_0))}{\varrho^{2s-\gamma}} \left[\tau^{\gamma} \varrho^{2s-\gamma} + \tau^{\gamma} \left\{ \tau^{1-2s-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} + \tau^{1-2s} + 1 \right\} \right] \\
(5.33) \quad &= c \tau^{\gamma} \left\{ 1 + \tau^{1-2s-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} \varrho^{\gamma-2s} + \tau^{1-2s} \varrho^{\gamma-2s} \right\} \mathcal{E}_{\gamma}(u; B_{\varrho}(x_0)),
\end{aligned}$$

recalling that $1 \leq \varrho^{\gamma-2s}$ from (5.3). Finally, combining (5.29) and (5.33), and using that $\tau^{\gamma} \leq \tau$ by (5.3) and $\tau \leq 1$, we arrive at

$$(5.34) \quad \mathcal{E}_{\gamma}(u; B_{\tau\varrho}(x_0)) \leq c \left\{ \tau^{1-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} (1 + \tau^{\gamma-2s} \varrho^{\gamma-2s}) + \tau + \tau^{\gamma+1-2s} \varrho^{\gamma-2s} \right\} \mathcal{E}_{\gamma}(u; B_{\varrho}(x_0)),$$

where $c = c(\mathbf{data})$. As a consequence of the above estimate, we can prove the following

Proposition 5.4 (Excess decay in the small potential regime). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (1.1) under assumptions (1.3) to (1.5), let f satisfy (2.1) and let $B_{\varrho}(x_0) \Subset \Omega$. Then for every $\beta \in (0, 1)$, there exist small constants*

$$(5.35) \quad \varepsilon^* = \varepsilon^*(\mathbf{data}, \beta), \quad \delta_0 = \delta_0(\mathbf{data}, \beta), \quad \varepsilon_0 = \varepsilon_0(\mathbf{data}, \beta, \omega_u(\cdot)), \quad r_1 = r_1(\mathbf{data}, \beta, \nu) \in (0, 1), \quad \tau = \tau(\mathbf{data}, \beta)$$

such that if

- $A(\cdot, \cdot)$ satisfies (2.4) and the BMO-condition (5.2) for some $\delta \leq \delta_0$ and radius $r_A > 0$,
- the ball has radius $\varrho \leq \min\{r_0, r_1, r_A\}$, with r_0 satisfying (4.4) and (5.12),
- the excess is small at the initial scale, that is (5.1) holds,
- the system is in the small potential regime, that is (5.14) holds;

then there holds the γ -excess decay estimate

$$(5.36) \quad \mathcal{E}_{\gamma}(u; B_{\tau\varrho}(x_0)) \leq \tau^{\beta} \mathcal{E}_{\gamma}(u; B_{\varrho}(x_0)) \leq \frac{1}{2} \mathcal{E}_{\gamma}(u; B_{\varrho}(x_0)).$$

Proof. We start from (5.34): factoring out τ^{β} and grouping the terms appropriately, we obtain

$$(5.37) \quad \mathcal{E}_{\gamma}(u; B_{\tau\varrho}(x_0)) \leq \tau^{\beta} \left\{ \bar{c} \tau^{1-\beta} \left(\tau^{-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} + 1 \right) \left(1 + \tau^{\gamma-2s} \varrho^{\gamma-2s} \right) \right\} \mathcal{E}_{\gamma}(u; B_{\varrho}(x_0))$$

with $\bar{c}$ depending on $\mathbf{data}$. We now fix $\tau = \tau(\mathbf{data}, \beta) > 0$ so small that

$$(5.38) \quad \tau^{1-\beta} \leq \frac{1}{4\bar{c}} \quad \tau^{\beta} \leq \frac{1}{2}, \quad \tau \leq \frac{1}{64}$$

then, given the dependence $\gamma = \gamma(s)$ and $\vartheta = \vartheta(\mathbf{data}) \in (0, 1)$, we may choose a small radius $r_1 = r_1(\mathbf{data}, \beta, \nu) \in (0, 1)$ so that

$$(5.39) \quad \tau^{\gamma-2s} r_1^{\gamma-2s} + \tau^{-n/2} (\nu r_1^{2-\gamma})^{\vartheta/2} \leq \frac{1}{4}$$

then we choose $\varepsilon_0 = \varepsilon_0(\mathbf{data}, \beta, \omega_u(\cdot))$ small enough such that

$$(5.40) \quad \tau^{-n/2} \omega_u(\varepsilon_0)^{\vartheta/4} \leq \frac{1}{4},$$

and finally we pick small parameters ε^*, δ_0 as in (5.35) fulfilling

$$(5.41) \quad \tau^{-n/2} [\varepsilon^* + \delta_0]^{\vartheta/2} \leq \frac{1}{2}.$$

Combining the content of (5.39) to (5.41) into (5.23) yields, by sub-additivity,

$$\tau^{-n/2} [s_1(\varrho, \varepsilon^*, \varepsilon_0, \delta)]^{\vartheta/2} \leq 1$$

for all $\varrho \leq r_1$ and $\delta \leq \delta_0$. Inserting this estimate and (5.38)₁ into (5.37), yields (5.36). $\square$

5.3. Large potential regime. Let us now assume that (5.14) is violated, that is

$$(5.42) \quad \varrho^2 \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} > \varepsilon^* \mathcal{E}_\gamma(u; B_\varrho(x_0))$$

holds; for $\beta \in (0, 1)$ now ε^* is the fixed constant, depending on **data** and β , given by Proposition 5.4. As a simple application of (3.22), we infer the following

Proposition 5.5 (Excess decay in the large potential regime). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be as in Proposition 5.4, and assume that $B_\varrho(x_0) \Subset \Omega$, $\varrho \leq r_0$, r_0 as in the beginning of Section 5 and satisfying (5.12), are such that (5.42) is in force. Then, for every $\tau \in (0, 1]$, we have*

$$(5.43) \quad \mathcal{E}_\gamma(u; B_{\tau\varrho}(x_0)) \leq \frac{c(n, s)\tau^{-n/2}}{\varepsilon^*} \varrho^2 \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi}.$$

5.4. Conclusion. We now combine Propositions 5.4 and 5.5 and implement a standard iteration procedure to deduce the following conditional excess decay estimate:

Proposition 5.6 (Excess decay). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a weak solution to (1.1) under assumptions (1.3) to (1.5) and let f satisfy (2.1). Fix $\beta \in (0, 1)$, γ as in (5.3), and let $B_\varrho(x_0) \Subset \Omega$. There exists a smallness parameter $\varepsilon_s = \varepsilon_s(\mathbf{data}, \omega_u(\cdot), \beta)$, a small radius $r_1 = r_1(\mathbf{data}, \nu, \beta) \in (0, 1)$, a small constant $\delta_0 = \delta_0(\mathbf{data}, \beta)$ and a large constant $c = c(\mathbf{data}, \beta) \geq 1$ such that if*

- $A(\cdot, \cdot)$ satisfies (2.4) and the BMO condition (5.2) for some $\delta \leq \delta_0$ and $r_A > 0$;
- the ball has radius $\varrho \leq \min\{r_0, r_1, r_A\}$, with r_0 satisfying (4.4) and (5.12);
- the following smallness condition at the initial scale holds:

$$(5.44) \quad \mathcal{E}_\gamma(u; B_\varrho(x_0)) + \sup_{t \leq \varrho} t^2 \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \leq \varepsilon_s;$$

then there exists $\tau = \tau(\mathbf{data}, \beta)$ such that $\tau^\beta \leq 1/2$ and

$$(5.45) \quad \mathcal{E}_\gamma(u; B_{\tau^{k+1}\varrho}(x_0)) \leq \tau^\beta \mathcal{E}_\gamma(u; B_{\tau^k\varrho}(x_0)) + c(\tau^k\varrho)^2 \left(\int_{B_{\tau^k\varrho}(x_0)} |f|^\chi dx \right)^{1/\chi}$$

for all $k \in \mathbb{N}_0$, and

$$(5.46) \quad \mathcal{E}_\gamma(u; B_\sigma(x_0)) \leq c\left(\frac{\sigma}{\varrho}\right)^\beta \mathcal{E}_\gamma(u; B_\varrho(x_0)) + c \sup_{\sigma \leq t \leq \varrho} t^2 \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi}$$

for every $0 < \sigma \leq \varrho$. Moreover

$$(5.47) \quad \sup_{t \leq \varrho} \mathcal{E}_\gamma(u; B_t(x_0)) \leq c(\mathbf{data}, \beta) \varepsilon_s.$$

Proof. For $\beta \in (0, 1)$ as in the statement, we take $\varepsilon^*(\mathbf{data}, \beta)$, $\varepsilon_0(\mathbf{data}, \beta, \omega_u(\cdot))$, $\tau(\mathbf{data}, \beta)$, $r_1(\mathbf{data}, \beta, \nu)$ the corresponding constants from Proposition 5.4 and we fix ε_s as

$$\varepsilon_s = \varepsilon_s(\mathbf{data}, \omega_u(\cdot), \beta) = \frac{\varepsilon_0}{4} \left[\frac{c(n, s)\tau^{-n/2}}{\varepsilon^*} \right]^{-1},$$

where $c(n, s)$ is the constant in (5.43); notice that $\varepsilon_s \leq \varepsilon_0/4$. Now, for $\varrho \leq \min\{r_0, r_1, r_A\}$, combining the estimates from Propositions 5.4 and 5.5, we get

$$(5.48) \quad \mathcal{E}_\gamma(u; B_{\tau\varrho}(x_0)) \leq \tau^\beta \mathcal{E}_\gamma(u; B_\varrho(x_0)) + \frac{c\tau^{-n/2}}{\varepsilon^*} \varrho^2 \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi}$$

for c depending on the **data**. Using (5.44), we can estimate the right-hand side as

$$\tau^\beta \mathcal{E}_\gamma(u; B_\varrho(x_0)) + \frac{c \tau^{-n/2}}{\varepsilon^*} \varrho^2 \left(\int_{B_\varrho(x_0)} |f|^\chi dx \right)^{1/\chi} \leq \left[\tau^\beta + \frac{c \tau^{-n/2}}{\varepsilon^*} \right] \varepsilon_s \leq \frac{\varepsilon_0}{4} + \frac{\varepsilon_0}{4} \leq \frac{\varepsilon_0}{2},$$

we are in a position to use again the decay estimate (5.48) with $\tau \varrho$ replacing ϱ , as (5.1) holds in $B_{\tau \varrho}(x_0)$ and, repeating the argument, we get that for every $k \in \mathbb{N}_0$,

$$\begin{cases} \mathcal{E}_\gamma(u; B_{\tau^k \varrho}(x_0)) \leq \frac{\varepsilon_0}{2} \\ \mathcal{E}_\gamma(u; B_{\tau^{k+1} \varrho}(x_0)) \leq \tau^\beta \mathcal{E}_\gamma(u; B_{\tau^k \varrho}(x_0)) + c(\tau^k \varrho)^2 \left(\int_{B_{\tau^k \varrho}(x_0)} |f|^\chi dx \right)^{1/\chi} \end{cases},$$

so that (5.45) is proved; moreover, by iterating this estimate we deduce that, for every $i, j \in \mathbb{N}_0$ with $j \geq i$,

$$\begin{aligned} \mathcal{E}_\gamma(u; B_{\tau^j \varrho}(x_0)) &\leq \tau^{\beta(j-i)} \mathcal{E}_\gamma(u; B_{\tau^i \varrho}(x_0)) + c \sum_{k=i}^{j-1} \tau^{\beta(j-1-k)} (\tau^k \varrho)^2 \left(\int_{B_{\tau^k \varrho}(x_0)} |f|^\chi dx \right)^{1/\chi} \\ &\leq \tau^{\beta(j-i)} \mathcal{E}_\gamma(u; B_{\tau^i \varrho}(x_0)) + \frac{c}{1 - \tau^\beta} \sup_{\tau^{j-1} \varrho \leq t \leq \tau^i \varrho} t^2 \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \end{aligned}$$

(if $i = j$ we agree that the summation and the supremum are both zero). Now the conclusion is standard: to prove (5.46), we first consider when $\sigma < \tau \rho$, and in this case there are $j, k \in \mathbb{N}$, with $j \geq k + 1$, such that $\tau^{j+1} \varrho < \sigma \leq \tau^j \varrho \leq \tau^{k+1} \varrho < \rho \leq \tau^k \varrho$: at this point by (3.22), using the previous estimate with $i = k + 1$, with the same agreement on the supremum,

$$\begin{aligned} \mathcal{E}_\gamma(u; B_\sigma(x_0)) &\leq c(n, s) \tau^{-\frac{n}{2}} \mathcal{E}_\gamma(u; B_{\tau^j \varrho}(x_0)) \\ &\leq c(n, s) \tau^{-\frac{n}{2}} \left[\tau^{\beta(j-k-1)} \mathcal{E}_\gamma(u; B_{\tau^{k+1} \varrho}(x_0)) + \frac{c}{1 - \tau^\beta} \sup_{\tau^{j-1} \varrho \leq t \leq \tau^{k+1} \varrho} t^2 \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \right] \\ &\leq c \left[\left(\frac{\sigma}{\rho} \right)^\beta \mathcal{E}_\gamma(u; B_\rho(x_0)) + \sup_{\sigma \leq t \leq \rho} t^2 \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} \right], \end{aligned}$$

with now $c = c(\mathbf{data}, \beta)$, since $\tau^{j-k-1} \leq \tau^{-2}(\sigma/\rho) \leq c(\sigma/\rho)$. In the case $\sigma \in [\tau \rho, \rho]$ the conclusion is again standard by (3.22). Using (5.44) in (5.46) immediately leads to (5.47). $\square$

6. REGULAR SET AND THEREIN REGULARITY OF SOLUTIONS

For $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ a fixed local solution to (1.1), under the assumptions of Theorem 2.1, and γ as in (5.3), let us define the γ -singular set

$$\Omega_{\text{sing}, \gamma} = \{x_0 \in \Omega : \limsup_{\varrho \rightarrow 0^+} \mathcal{E}_\gamma(u; B_\varrho(x_0)) > 0\}.$$

Taking into account (3.24) from Lemma 3.24, with ζ defined in Proposition 4.4, we have

$$\Omega_{\text{sing}, \gamma} \subseteq \{x_0 \in \Omega : \limsup_{\varrho \rightarrow 0^+} E(u; B_\varrho(x_0)) > 0\}$$

and

$$\Omega_{\text{sing}, \gamma} \subseteq \left\{ x_0 \in \Omega : \limsup_{\varrho \rightarrow 0^+} \varrho^{-[n-(2+\zeta)]} \int_{B_\varrho(x_0)} |Du|^{2+\zeta} dx > 0 \right\}.$$

The latter inclusion follows from Poincaré's inequality

$$\left(\int_{B_\varrho(x_0)} |u - (u)_{B_\varrho(x_0)}|^2 dx \right)^{(2+\zeta)/2} \leq c(n, \zeta) \varrho^{2+\zeta-n} \int_{B_\varrho(x_0)} |Du|^{2+\zeta} dx.$$

Therefore $\Omega_{\text{sing}, \gamma} = \emptyset$ in dimension $n = 2$, while, if $n \geq 3$, the classical Giusti lemma [34, Chapter IV, Theorem 2.2] gives

$$\dim_{\mathcal{H}}(\Omega_{\text{sing}, \gamma}) \leq n - 2 - \zeta.$$

6.1. A general construction. We now record a construction that will be used with different choices of the parameter $\beta \in (0, 1)$. Fix $\beta \in (0, 1)$ and let $\varepsilon_s = \varepsilon_s(\mathbf{data}, \omega_u(\cdot), \beta)$, $r_1 = r_1(\mathbf{data}, \beta, \nu)$, $\delta_0 = \delta_0(\mathbf{data})$ and $\tau = \tau(\mathbf{data}, \beta)$ be the quantities provided by Proposition 5.6. We assume that the smallness hypotheses (2.7) and (2.8) hold with this value of δ_0 and with $\varepsilon_b := \varepsilon_s/2$. Let also r_0 satisfy (4.4) and (5.12).

We fix a compact exhaustion $(K_i)_{i \in \mathbb{N}}$ of Ω such that $K_i \Subset \text{int}(K_{i+1})$ for every $i \in \mathbb{N}$. For every $i \in \mathbb{N}$ there exist radii $r_A(K_i), r_f(K_i) \in (0, \text{dist}(K_i, \partial\Omega))$ such that (2.7) and (2.8) hold on K_i ; we set

$$(6.1) \quad \bar{r}_\beta(K_i) := \min \left\{ r_0, r_1, r_f(K_i), r_A(K_i), \frac{1}{4} \text{dist}(K_i, \partial\Omega) \right\},$$

and then

$$(6.2) \quad K_{u,i,\beta} := \left\{ x \in \text{int}(K_i) : \exists \varrho_x \in (0, \bar{r}_\beta(K_i)) \text{ such that } B_{\varrho_x}(x) \Subset \Omega \text{ and } \mathcal{E}_\gamma(u; B_{\varrho_x}(x)) < \varepsilon_s/2 \right\}$$

and

$$(6.3) \quad \Omega_{u,\beta} := \bigcup_{i \in \mathbb{N}} K_{u,i,\beta}.$$

By the continuity of $x \mapsto \mathcal{E}_\gamma(u; B_\varrho(x))$ for fixed ϱ (see the discussion after (3.21)) the set $\Omega_{u,\beta}$ is open. Moreover, if $x_0 \in \Omega \setminus \Omega_{u,\beta}$, then

$$\limsup_{\varrho \rightarrow 0^+} \mathcal{E}_\gamma(u; B_\varrho(x_0)) \geq \frac{\varepsilon_s}{2}.$$

Indeed, otherwise, since $x_0 \in \text{int}(K_i)$ for some i , one could choose $\varrho < \bar{r}_\beta(K_i)$ sufficiently small so that $B_\varrho(x_0) \Subset \Omega$ and $\mathcal{E}_\gamma(u; B_\varrho(x_0)) < \varepsilon_s/2$, which would imply $x_0 \in K_{u,i,\beta} \subset \Omega_{u,\beta}$, a contradiction. Hence $\Omega \setminus \Omega_{u,\beta} \subseteq \Omega_{\text{sing},\gamma}$; consequently, $\Omega \setminus \Omega_{u,\beta} = \emptyset$ if $n = 2$ while, if $n \geq 3$,

$$(6.4) \quad \dim_{\mathcal{H}}(\Omega \setminus \Omega_{u,\beta}) \leq n - 2 - \zeta.$$

The next proofs rely on the following observation. Let $x_0 \in \Omega_{u,\beta}$; then, there exist $i = i(x_0) \in \mathbb{N}$ and a radius $\varrho_{x_0} \in (0, \bar{r}_\beta(K_i))$ such that $B_{\varrho_{x_0}}(x_0) \Subset \Omega$ and

$$(6.5) \quad \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x_0)) < \frac{\varepsilon_s}{2}.$$

By the already mentioned continuity of the map $x \mapsto \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x))$, there exists a small radius $r_{x_0} \leq \varrho_{x_0}/8$ such that $B_{r_{x_0}}(x_0) \Subset \text{int}(K_i)$, $B_{\varrho_{x_0}}(\bar{x}) \Subset \Omega$ for every $\bar{x} \in B_{r_{x_0}}(x_0)$ and

$$(6.6) \quad \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(\bar{x})) < \frac{\varepsilon_s}{2} \quad \text{for every } \bar{x} \in B_{r_{x_0}}(x_0);$$

since $\varrho_{x_0} < r_f(K_i)$ and $\bar{x} \in K_i$, assumption (2.8), applied with $K = K_i$, gives

$$(6.7) \quad \sup_{\substack{\bar{x} \in B_{r_{x_0}}(x_0) \\ 0 < t \leq \varrho_{x_0}}} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} < \frac{\varepsilon_s}{2}.$$

Therefore the assumption (5.44) of Proposition 5.6 holds on every ball $B_{\varrho_{x_0}}(\bar{x})$ with $\bar{x} \in B_{r_{x_0}}(x_0)$. Moreover, (2.7) gives the required coefficient smallness and, since also $\varrho_{x_0} < \min\{r_0, r_1\}$, all assumptions of Proposition 5.6 are fulfilled on such balls. Therefore Proposition 5.6 yields

$$(6.8) \quad \sup_{0 < t \leq \varrho_{x_0}} \mathcal{E}_\gamma(u; B_t(\bar{x})) \leq c(\mathbf{data}, \beta) \varepsilon_s,$$

$$(6.9) \quad \mathcal{E}_\gamma(u; B_{\tau^{k+1}\varrho_{x_0}}(\bar{x})) \leq \tau^\beta \mathcal{E}_\gamma(u; B_{\tau^k\varrho_{x_0}}(\bar{x})) + c(\tau^k\varrho_{x_0})^2 \left(\int_{B_{\tau^k\varrho_{x_0}}(\bar{x})} |f|^\chi dx \right)^{1/\chi} \quad \text{for all } k \in \mathbb{N}_0,$$

and

$$(6.10) \quad \mathcal{E}_\gamma(u; B_\sigma(\bar{x})) \leq c \left(\frac{\sigma}{\rho} \right)^\beta \mathcal{E}_\gamma(u; B_\rho(\bar{x})) + c \sup_{t \leq \rho} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi}$$

for all $0 < \sigma \leq \rho \leq \varrho_{x_0}$ and all $\bar{x} \in B_{r_{x_0}}(x_0)$; all the constants c depend on the $\mathbf{data}$ and β , but not on $\bar{x}$.

Proof of Theorem 2.1. The proof is essentially a consequence of Proposition 5.6. However, due to the local nature of assumptions (2.7)–(2.8), the argument is slightly more delicate than in the standard setting, and we therefore provide the details.

We use the construction above with $\beta = 1/2$. Thus $\varepsilon_s = \varepsilon_s(\mathbf{data}, \omega_u(\cdot))$, $r_1 = r_1(\mathbf{data}, \nu)$, $\delta_0 = \delta_0(\mathbf{data})$, and the smallness parameter in the statement of the theorem is chosen as $\varepsilon_b := \varepsilon_s/2$; we also set $\Omega_u := \Omega_{u,1/2}$. By the construction above, Ω_u is open, $\Omega \setminus \Omega_u = \emptyset$ if $n = 2$, while, if $n \geq 3$, estimate (6.4) holds. It remains to prove (2.9); the case $A(x, u) \equiv A(x)$ will be treated separately in Section 8.

Let now $K \Subset \Omega_u$. For every $x_0 \in K$ we have radii $\varrho_{x_0} > 0$ and $r_{x_0} \leq \varrho_{x_0}/8$ satisfying (6.8) for every $\bar{x} \in B_{r_{x_0}}(x_0)$. The radii r_{x_0} may depend on x_0 and may degenerate along K ; we therefore use compactness: there exist points $x_1, \dots, x_m \in K$ such that the union of $B_{r_{x_j}}(x_j)$ covers K ; we set then $r_K := \min_{1 \leq j \leq m} r_{x_j} > 0$. We claim that

$$E_{r_K}(u; \bar{x}) \leq c(\mathbf{data}) \varepsilon_s \quad \text{for every } \bar{x} \in K.$$

Indeed, if $\bar{x} \in K$, then $\bar{x} \in B_{r_{x_j}}(x_j)$ for some $j \in \{1, \dots, m\}$. Since $r_K \leq r_{x_j} \leq \varrho_{x_j}$, estimate (6.8), applied with x_j in place of x_0 , gives

$$E_{r_K}(u; \bar{x}) = \sup_{0 < t \leq r_K} E(u; B_t(\bar{x})) \leq \sup_{0 < t \leq r_K} \mathcal{E}_\gamma(u; B_t(\bar{x})) \leq c(\mathbf{data}) \varepsilon_s;$$

thus, by the definition in Paragraph 3.2, $u \in BMO(K; \mathbb{R}^N)$. Since $K \Subset \Omega_u$ was arbitrary, this proves (2.9). $\square$

Proof of Theorem 2.2. We again fix $\beta \in (0, 1)$, say $\beta = 1/2$, and apply the construction described at the beginning of the section. Indeed, assumption (2.10) implies the validity of (2.8) on every compact subset of Ω , for some radius $r_f > 0$ depending on the compact set. In particular, estimate (6.10) holds for all $0 < \sigma \leq \rho \leq \varrho_{x_0}$, with a constant $c = c(\mathbf{data})$ independent of $\bar{x} \in B_{r_{x_0}}(x_0)$. Hence

$$\mathcal{E}_\gamma(u; B_\sigma(\bar{x})) \leq c \left(\frac{\sigma}{\rho} \right)^\beta \mathcal{E}_\gamma(u; B_\rho(\bar{x})) + c \sup_{t \leq \rho} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi}.$$

Taking the supremum with respect to $\bar{x} \in B_{r_{x_0}}(x_0)$ and then the limsup as $\sigma \rightarrow 0^+$, with ρ fixed, we obtain

$$\limsup_{\sigma \rightarrow 0^+} \sup_{\bar{x} \in B_{r_{x_0}}(x_0)} \mathcal{E}_\gamma(u; B_\sigma(\bar{x})) \leq c \sup_{\substack{\bar{x} \in B_{r_{x_0}}(x_0), \\ t \leq \rho}} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi}.$$

Letting $\rho \rightarrow 0^+$ and using (2.10), we obtain

$$(6.11) \quad \lim_{\sigma \rightarrow 0^+} \mathcal{E}_\gamma(u; B_\sigma(\cdot)) = 0 \quad \text{uniformly in } B_{r_{x_0}}(x_0).$$

This implies $u \in VMO_{\text{loc}}(\Omega_u; \mathbb{R}^N)$ by a covering argument analogous to the one used in the proof of Theorem 2.1. $\square$

Proof of Theorem 2.3. Again, we fix $\beta \in (0, 1)$, say $\beta = 1/2$. Thanks to (3.10), we see that (2.12) implies

$$\lim_{\varrho \rightarrow 0^+} \sup_{t \leq \varrho} t^2 \left(\int_{B_t(\cdot)} |f|^\chi dx \right)^{1/\chi} = 0 \quad \text{locally uniformly in } \Omega,$$

that in turn implies the validity of (2.8) for some radius $r_f > 0$, so that (6.9) holds true.

Now, let $\tau = \tau(\mathbf{data}) \in (0, 1)$ be as given at the beginning of the Section for the choice $\beta = 1/2$. We set $\varrho_j = \tau^j \varrho_{x_0}$, and we write, for short, $B_j(\bar{x}) = B_{\varrho_j}(\bar{x})$, $\mathcal{E}_j(\bar{x}) = \mathcal{E}_\gamma(u; B_{\varrho_j}(\bar{x}))$ for $\bar{x} \in B_{r_{x_0}}(x_0)$; then by using (6.9) and recalling the bound before (5.45) we deduce

$$\mathcal{E}_{j+1}(\bar{x}) \leq \frac{1}{2} \mathcal{E}_j(\bar{x}) + c \varrho_j^2 \left(\int_{B_j(\bar{x})} |f|^\chi dx \right)^{1/\chi}$$

for all $j \in \mathbb{N}_0$ and all $\bar{x} \in B_{r_{x_0}}(x_0)$. For $i, k \in \mathbb{N}$, $k > i$, summing over $j = i, i+1, \dots, k$, changing the summation variable and enlarging the summation on the right-hand side, we obtain

$$\sum_{j=i}^{k+1} \mathcal{E}_j(\bar{x}) \leq \mathcal{E}_i(\bar{x}) + \frac{1}{2} \sum_{j=i}^{k+1} \mathcal{E}_j(\bar{x}) + c \sum_{j=i}^k \varrho_j^2 \left(\int_{B_j(\bar{x})} |f|^\chi dx \right)^{1/\chi}$$

and, after reabsorbing,

$$\sum_{j=i}^{k+1} \mathcal{E}_j(\bar{x}) \leq 2\mathcal{E}_i(\bar{x}) + c \sum_{j=i}^{\infty} \varrho_j^2 \left(\int_{B_j(\bar{x})} |f|^\chi dx \right)^{1/\chi};$$

the constant c depends here only on **data**. Now, for the same indices, by telescopic summation, using the above estimate and (3.9), we get

$$(6.12) \quad \begin{aligned} |(u)_{B_k(\bar{x})} - (u)_{B_i(\bar{x})}| &\leq \sum_{j=i}^k |(u)_{B_{j+1}(\bar{x})} - (u)_{B_j(\bar{x})}| \leq \tau^{-n} \sum_{j=i}^{k+1} \left(\int_{B_j(\bar{x})} |u - (u)_{B_j(\bar{x})}|^2 dx \right)^{1/2} \\ &\leq c \mathcal{E}_i(\bar{x}) + c \sum_{j=i}^k \varrho_j^2 \left(\int_{B_j(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq c \mathcal{E}_i(\bar{x}) + c \mathbf{I}_{2,\chi}^f(\bar{x}, 2\varrho_{i-1}). \end{aligned}$$

As the right-hand side of (6.12) vanishes uniformly in $B_{r_{x_0}}(x_0)$ thanks to (6.11) and (2.12), it follows that the sequence of maps $\{\bar{x} \mapsto (u)_{B_i(\bar{x})}\}_{i \in \mathbb{N}}$ satisfies the Cauchy condition with respect to the uniform convergence in $B_{r_{x_0}}(x_0)$. As $(u)_{B_i(\bar{x})} \rightarrow u(\bar{x})$ for $i \rightarrow +\infty$, for a.e. $\bar{x}$ by Lebesgue's point theorem, this implies that u agrees a.e. with a continuous function. $\square$

The next result is a slightly more general version of Theorem 2.4.

Theorem 6.1 (Partial Hölder continuity). *Let $\beta_0 \in (0, 1)$, and assume that for every compact set $K \Subset \Omega$, there exists a radius $R_f = R_f(K) > 0$ such that f satisfies*

$$(6.13) \quad \sup_{\substack{x_0 \in K, \\ 0 < t \leq R_f}} t^{2-\beta_0} \left(\int_{B_t(x_0)} |f|^\chi dx \right)^{1/\chi} < \infty;$$

suppose moreover that $A(\cdot, \cdot)$ satisfies (2.4). There exists a small constant $\delta_0 = \delta_0(\mathbf{data}, \beta_0)$ such that if for every compact $K \Subset \Omega$, there exists a radius $r_A = r_A(K) > 0$ for which (5.2) holds uniformly for $x_0 \in K$, then

$$u \in C_{\text{loc}}^{0, \beta_0}(\Omega_u; \mathbb{R}^N),$$

where Ω_u is the open set given by (6.2)–(6.3) for the choice $\beta = (1 + \beta_0)/2$. Moreover, for every $x_0 \in \Omega_u$, there exist radii $0 < r_{x_0} \leq \varrho_{x_0}$ such that $B_{r_{x_0}}(x_0) \Subset \Omega_u$, $B_{2\varrho_{x_0}}(x_0) \Subset \Omega$, and

$$(6.14) \quad \begin{aligned} [u]_{C^{0, \beta_0}(B_{r_{x_0}}(x_0))} &\leq c \varrho_{x_0}^{-\beta_0} \left(\int_{B_{\varrho_{x_0}}(x_0)} |u - (u)_{B_{\varrho_{x_0}}(x_0)}|^2 dx \right)^{1/2} \\ &\quad + c \varrho_{x_0}^{\gamma-2s-\beta_0} \mathbf{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)) + c \|\mathbf{M}_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))}, \end{aligned}$$

where $c = c(\mathbf{data}, \beta_0)$, $\gamma = \gamma(s)$ is fixed as in (5.3), and we set

$$(6.15) \quad \mathbf{M}_{\varrho, \chi}^{2-\beta_0}[f](\bar{x}) := \sup_{t \in (0, \varrho]} t^{2-\beta_0} \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi}.$$

Assumption (6.13) is verified if $f \in \mathcal{M}_{\text{loc}}^{n/(2-\beta_0)}(\Omega; \mathbb{R}^N)$, in which case the quantitative estimate (2.17) holds true.

Proof. We let $\beta = (1 + \beta_0)/2$ and define $\Omega_u = \Omega_{u, (1+\beta_0)/2}$ as in (6.2)–(6.3); we fix, according to this value of β , $\varepsilon_s = \varepsilon_s(\mathbf{data}, \omega_u(\cdot), \beta_0)$, $r_1 = r_1(\mathbf{data}, \beta_0, \nu)$ and $\delta_0 = \delta_0(\mathbf{data}, \beta_0)$ as at the beginning of this Section. We then fix a generic $x_0 \in \Omega_u$; as described in Paragraph 6.1, there exists an index $i = i(x_0)$ such that $x_0 \in \text{int}(K_i)$ and we let now r_A to be the constant $r_A(K_i)$ given from the assumption of the Theorem. Moreover, from (6.13) it follows immediately that there exists a radius $0 < r_f \leq R_f$ such that (2.8) is satisfied on K_i with $\varepsilon_b = \varepsilon_s/2$; moreover we can suppose $r_f \leq \text{dist}(K_i, \partial\Omega)/4$. Using the general construction of Paragraph 6.1, we may find radii $r_{x_0} \leq \varrho_{x_0}/8$ such that (6.10) holds and $B_{\varrho_{x_0}}(\bar{x}) \subset B_{2\varrho_{x_0}}(x_0) \Subset \Omega$ for all $\bar{x} \in B_{r_{x_0}}(x_0)$.

We now choose $\vartheta = \vartheta(\mathbf{data}, \beta_0) \in (0, 1/16)$ such that $c\vartheta^{\beta-\beta_0} = c\vartheta^{(1-\beta_0)/2} \leq 1/2$, with $c = c(\mathbf{data}, \beta_0)$ exactly the constant appearing in (6.10); thus, by choosing $\sigma = \vartheta\rho$ for $0 < \rho \leq \varrho_{x_0}$ in (6.10), we find

$$(6.16) \quad \mathcal{E}_\gamma(u; B_{\vartheta\rho}(\bar{x})) \leq \frac{1}{2} \vartheta^{\beta_0} \mathcal{E}_\gamma(u; B_\rho(\bar{x})) + c\rho^{\beta_0} \sup_{t \leq \rho} t^{2-\beta_0} \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} \quad \text{for all } \bar{x} \in B_{r_{x_0}}(x_0),$$

with $c = c(\mathbf{data}, \beta_0)$. Now, for $0 < \tilde{\rho} \leq \rho$, we define the fractional maximal operators

$$(6.17) \quad \begin{cases} M_{\tilde{\rho}, \rho}^{\beta_0}(\bar{x}) := \sup_{t \in (\tilde{\rho}, \rho]} t^{-\beta_0} \mathcal{E}_\gamma(u; B_t(\bar{x})) \\ M_{\rho}^{\beta_0}(\bar{x}) := \sup_{t \in (0, \rho]} t^{-\beta_0} \mathcal{E}_\gamma(u; B_t(\bar{x})) \end{cases},$$

and $M_{\vartheta, \chi}^{2-\beta_0}[f](\bar{x})$ is defined by (6.15). Multiplying (6.16) by $(\vartheta\rho)^{-\beta_0}$, and by taking the supremum for $\rho \in (\varrho_{x_0}/[2k], \varrho_{x_0}/2]$ for $k \in \mathbb{N} \setminus \{0, 1\}$, we deduce

$$\begin{aligned} M_{\vartheta \varrho_{x_0}/[2k], \vartheta \varrho_{x_0}/2}^{\beta_0}(\bar{x}) &\leq \frac{1}{2} M_{\varrho_{x_0}/[2k], \varrho_{x_0}/2}^{\beta_0}(\bar{x}) + c \vartheta^{-\beta_0} M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f](\bar{x}) \\ &\leq \frac{1}{2} M_{\vartheta \varrho_{x_0}/[2k], \vartheta \varrho_{x_0}/2}^{\beta_0}(\bar{x}) + \frac{1}{2} M_{\vartheta \varrho_{x_0}/2, \varrho_{x_0}/2}^{\beta_0}(\bar{x}) + c M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f](\bar{x}), \end{aligned}$$

for all $\bar{x} \in B_{r_{x_0}}(x_0)$, where we also used that ϑ depends on $\mathbf{data}$ and β_0 . By reabsorbing terms, we get

$$M_{\vartheta \varrho_{x_0}/[2k], \vartheta \varrho_{x_0}/2}^{\beta_0}(\bar{x}) \leq M_{\vartheta \varrho_{x_0}/2, \varrho_{x_0}/2}^{\beta_0}(\bar{x}) + c M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f](\bar{x});$$

since, using the definition in (6.17) and the fact that $\vartheta = \vartheta(\mathbf{data}, \beta_0)$, we have

$$M_{\vartheta \varrho_{x_0}/2, \varrho_{x_0}/2}^{\beta_0}(\bar{x}) \leq c(\varrho_{x_0}/2)^{-\beta_0} \mathcal{E}_\gamma(u; B_{\varrho_{x_0}/2}(\bar{x})) = c \varrho_{x_0}^{-\beta_0} \mathcal{E}_\gamma(u; B_{\varrho_{x_0}/2}(\bar{x})),$$

combining the two estimates above, we obtain

$$M_{\vartheta \varrho_{x_0}/[2k], \varrho_{x_0}/2}^{\beta_0}(\bar{x}) \leq c \varrho_{x_0}^{-\beta_0} \mathcal{E}_\gamma(u; B_{\varrho_{x_0}/2}(\bar{x})) + c \|M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))},$$

with $c = c(\mathbf{data}, \beta_0)$, for all $\bar{x} \in B_{r_{x_0}}(x_0)$. On the other hand, recalling that $|\bar{x} - x_0| \leq r_{x_0} \leq \varrho_{x_0}/8 \leq 1/8$, we have $B_{\varrho_{x_0}/2}(\bar{x}) \subseteq B_{\varrho_{x_0}}(x_0)$; hence we may use (3.3) and (3.19) in (3.21) and get

$$\mathcal{E}_\gamma(u; B_{\varrho_{x_0}/2}(\bar{x})) \leq c \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x_0)) \quad \text{for all } \bar{x} \in B_{r_{x_0}}(x_0)$$

– recall that $\gamma - 2s \geq 0$. Combining the two estimates above, letting $k \rightarrow \infty$, and expanding the definition of $\mathcal{E}_\gamma(\cdot)$, we get

$$\begin{aligned} M_{\varrho_{x_0}/2}^{\beta_0}(\bar{x}) &\leq c \varrho_{x_0}^{-\beta_0} \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x_0)) + c \|M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))} \\ &\leq c \varrho_{x_0}^{-\beta_0} \left(\int_{B_{\varrho_{x_0}}(x_0)} |u - (u)_{B_{\varrho_{x_0}}(x_0)}|^2 dx \right)^{1/2} \\ &\quad + c \varrho_{x_0}^{\gamma-2s-\beta_0} \mathbf{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)) + c \|M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))}, \end{aligned}$$

which is valid for every $\bar{x} \in B_{r_{x_0}}(x_0)$, with $c = c(\mathbf{data}, \beta_0)$.

This estimate implies the desired β_0 -Hölder continuity of u and the local estimate (6.14) via Campanato's characterization of Hölder continuity [34, Chapter 3.1] since, by the very definition of $M_{\varrho}^{\beta_0}$ and $\mathcal{E}_\gamma(\cdot)$ in (6.17) and (3.21), we have

$$M_{\varrho_{x_0}/2}^{\beta_0}(\bar{x}) \geq \sigma^{-\beta_0} \left(\int_{B_\sigma(\bar{x})} |u - (u)_{B_\sigma(\bar{x})}|^2 dx \right)^{1/2}$$

for every $t \in (0, \varrho_{x_0}/2]$, and $\bar{x} \in B_{r_{x_0}}(x_0)$. Finally, thanks to (3.5), if (2.15) holds, then we have, for χ as chosen after (3.5)

$$t^{2-\beta_0} \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq c(n, \beta_0) \|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_t(\bar{x}))} \leq c(n, \beta_0) \|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_{2\varrho_{x_0}}(x_0))}$$

for all $\bar{x} \in B_{r_{x_0}}(x_0)$ and all $0 < t \leq \varrho_{x_0}$, hence (6.13) is valid, and we have

$$\|M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))} \leq c(n, \beta_0) \|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_{2\varrho_{x_0}}(x_0))}.$$

From this inequality, (6.14) and using that $\varrho_{x_0}^\gamma \leq \varrho_{x_0}$ by (5.3) as $\varrho_{x_0} \leq 1$, estimate (2.17) follows; local Hölder regularity follows using a covering argument similar to that employed in the proof of Theorem 2.1. $\square$

7. GRADIENT REGULARITY OF SOLUTIONS

We start with some preliminary remarks; in particular, let us verify the validity of Theorem 6.1 whenever assumptions (2.18) and (2.20) are in force, and consequently derive a Hölder estimate in the reference ball $B_{r_{x_0}}(x_0)$ for every exponent $\beta_0 \in (0, 1)$.

Let us now fix $\beta_0 = \beta_0(s) \in (0, 1) \setminus \{2s\}$ such that

$$(7.1) \quad 2s < 1 + \beta_0 \quad \Longleftrightarrow \quad 1 + \beta_0 - 2s > 0,$$

and let $x_0 \in \Omega_u$, where Ω_u is the open set given by (6.2)–(6.3) for the choice $\beta = (1 + \beta_0)/2$. Let $\delta_0 = \delta_0(\mathbf{data}, \beta_0)$ be the small constant appearing in Theorem 6.1. By assumption (2.18) and the fact that $\omega_x(\cdot)$ is a modulus of continuity, we may choose $r_A = r_A(\mathbf{data}, \beta_0, \omega_x(\cdot), \text{dist}(K, \partial\Omega)) \in (0, 1)$ such that (5.2) holds uniformly for $x_0 \in K$, as

$$(7.2) \quad \sup_{x_0 \in K} E_r(A; x_0) = \sup_{x_0 \in K} \sup_{\varrho \in (0, r]: B_{\varrho}(x_0) \subseteq \Omega, \substack{z \in \mathbb{R}^N, z \neq 0}} \int_{B_{\varrho}(x_0)} |A(x, z) - (A)_{B_{\varrho}(x_0)}(z)| dx \leq \Lambda \omega_x(r) \xrightarrow{r \rightarrow 0} 0.$$

Moreover, by (2.20) and (3.10), condition (6.13) is satisfied, with $R_f = \min\{1, \text{dist}(K, \partial\Omega)\}/2$; hence all the hypotheses of Theorem 6.1 are fulfilled. Let now r_{x_0} and ϱ_{x_0} be the radii provided by Theorem 6.1, with $0 < r_{x_0} \leq \varrho_{x_0}/8$ such that $B_{r_{x_0}}(x_0) \subseteq \Omega_u$, $B_{2\varrho_{x_0}}(x_0) \subseteq \Omega$. Using again (3.10), we have

$$\|M_{\varrho_{x_0}, \chi}^{2-\beta_0}[f]\|_{L^\infty(B_{r_{x_0}}(x_0))} \leq c(n, \beta_0, \chi) \|\mathbf{I}_{1, \chi}^f(\cdot, 2\varrho_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))}.$$

Therefore, estimate (6.14) gives

$$(7.3) \quad \begin{aligned} [u]_{C^{0, \beta_0}(B_{r_{x_0}}(x_0))} &\leq c \varrho_{x_0}^{-\beta_0} \left(\int_{B_{\varrho_{x_0}}(x_0)} |u - (u)_{B_{\varrho_{x_0}}(x_0)}|^2 dx \right)^{1/2} + c \varrho_{x_0}^{\gamma - \beta_0 - 2s} \mathbf{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)) \\ &\quad + c \|\mathbf{I}_{1, \chi}^f(\cdot, 2\varrho_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))} \\ &\leq c \varrho_{x_0}^{-n/2} \|Du\|_{L^2(B_{\varrho_{x_0}}(x_0))} + c \varrho_{x_0}^{\gamma - \beta_0 - 2s} \mathbf{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)) \\ &\quad + c \|\mathbf{I}_{1, \chi}^f(\cdot, 2\varrho_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))}, \end{aligned}$$

where $c = c(\mathbf{data}, \beta_0)$. In the last inequality we used Poincaré's inequality and the fact that $\varrho_{x_0}^{1-\beta_0} \leq 1$.

7.1. A gradient comparison estimate. Let us take a point $\bar{x} \in B_{r_{x_0}/2}(x_0)$ and a radius $0 < r \leq r_{x_0}/4$. Accordingly, we define

$$(7.4) \quad \bar{A} = A(\bar{x}, u(\bar{x})).$$

Here and in what follows, u denotes its continuous representative in Ω_u , so that $A(\bar{x}, u(\bar{x}))$ is well-defined for every $\bar{x} \in B_{r_{x_0}/2}(x_0)$, thanks to what proved at the beginning of the Section. Accordingly, with a slight abuse of notation, we shall write $[u]_{C^{0, \beta_0}}$ for the Hölder seminorm of this representative. In particular, the pointwise values $u(x)$ and $u(\bar{x})$ appearing below are well defined for every $x, \bar{x} \in B_{r_{x_0}}(x_0)$.

Consider $h \in W^{1,2}(B_r(\bar{x}); \mathbb{R}^N)$ solving

$$(7.5) \quad \begin{cases} -\text{div}(\bar{A} Dh) = 0 & \text{in } B_r(\bar{x}) \\ h = u & \text{on } \partial B_r(\bar{x}) \end{cases},$$

and we extend h outside $B_r(\bar{x})$ by setting

$$(7.6) \quad h \equiv u \quad \text{in } \mathbb{R}^n \setminus B_r(\bar{x}).$$

Let us now prove the following comparison estimate.

Proposition 7.1 (Comparison estimate). *Let $u \in W_{\text{loc}}^{1,2}(\Omega; \mathbb{R}^N) \cap L_{2s}^1$ be a local weak solution to (1.1) under the assumptions (1.3) and (1.4)–(1.5), and let $x_0, \bar{x}, r_{x_0}, r$ and $h \in W^{1,2}(B_r(\bar{x}); \mathbb{R}^N)$ be as above. Then the following comparison estimate holds:*

$$(7.7) \quad \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \leq c \left[\|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + r^{2(1-s)} \right] \left(\int_{B_{2r}(\bar{x})} |Du|^2 dx \right)^{1/2}$$

$$+ c r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c r^{1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})),$$

where $c = c(\mathbf{data}, \max\{\nu, 1\})$, and $\bar{A}$ is given by (7.4).

Proof. Using the coercivity and growth assumptions (1.3) (which are also inherited by $\bar{A}$), and testing (1.1) and (7.5) with $u - h$, we get

$$\begin{aligned} \frac{1}{\Lambda} \int_{B_r(\bar{x})} |Du - Dh|^2 dx &\leq \int_{B_r(\bar{x})} \langle \bar{A}(Du - Dh), Du - Dh \rangle dx = \int_{B_r(\bar{x})} \langle \bar{A}Du, Du - Dh \rangle dx \\ &= \int_{B_r(\bar{x})} \langle [\bar{A} - A(x, u(x))] Du, Du - Dh \rangle dx + \int_{B_r(\bar{x})} \langle A(x, u(x)) Du, Du - Dh \rangle dx \\ &= \int_{B_r(\bar{x})} \langle [\bar{A} - A(x, u(x))] Du, Du - Dh \rangle dx + \int_{B_r(\bar{x})} f \cdot (u - h) dx - \langle Q_\Phi^{\text{nl}} u, u - h \rangle_{\text{nl}} \\ (7.8) \quad &=: (I) + (II) + (III); \end{aligned}$$

we recall the notation introduced in (5.19). By Hölder's inequality, we get

$$\begin{aligned} |(I)| &\leq \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} \int_{B_r(\bar{x})} |Du| |Du - Dh| dx \\ (7.9) \quad &\leq c(n) \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} r^{n/2} \left(\int_{B_r(\bar{x})} |Du|^2 dx \right)^{1/2} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2}, \end{aligned}$$

and by Hölder's and Sobolev's inequalities

$$\begin{aligned} |(II)| &\leq c(n) r^{n/2^*} \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} \left(\int_{B_r(\bar{x})} |u - h|^{2^*} dx \right)^{1/2^*} \\ (7.10) \quad &\leq c(n) r^{n/2^*} \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2}. \end{aligned}$$

Then, by using (1.5) and recalling that $u - h \equiv 0$ in $\mathbb{R}^n \setminus B_r(\bar{x})$ by (7.6), we get

$$\begin{aligned} (7.11) \quad (III) &= - \int_{B_{2r}(\bar{x})} \int_{B_{2r}(\bar{x})} \Phi(y, z, u(y), u(z), u(y) - u(z)) \cdot [(u - h)(y) - (u - h)(z)] \frac{dy dz}{|y - z|^{n+2s}} \\ &\quad - 2 \int_{B_r(\bar{x})} \int_{\mathbb{R}^n \setminus B_{2r}(\bar{x})} \Phi(y, z, u(y), u(z), u(y) - u(z)) \cdot (u - h)(z) \frac{dy dz}{|y - z|^{n+2s}} =: (III)_1 + (III)_2. \end{aligned}$$

Using (1.4), Hölder's inequality and the fractional Sobolev inequality (3.6), we have

$$\begin{aligned} |(III)_1| &\leq c(n) \nu [u]_{W^{s,2}(B_{2r}(\bar{x}))} [u - h]_{W^{s,2}(B_{2r}(\bar{x}))} \\ (7.12) \quad &\leq c(n, s) \nu r^{2(1-s)} \left(\int_{B_{2r}(\bar{x})} |Du|^2 dx \right)^{1/2} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2}, \end{aligned}$$

where in the last line we also used that $Du = Dh$ in $B_{2r}(\bar{x}) \setminus B_r(\bar{x})$ by (7.6). Then, by using the elementary inequalities

$$(7.13) \quad |y - \bar{x}| \leq 4|y - z|, \quad \text{for } y \in \mathbb{R}^n \setminus B_{2r}(\bar{x}) \text{ and } z \in B_r(\bar{x}),$$

together with (1.4), (3.16), Hölder's and Poincaré's inequalities, we get

$$\begin{aligned} |(III)_2| &\leq c \nu \int_{B_r(\bar{x})} |(u - h)(z)| \left\{ \int_{\mathbb{R}^n \setminus B_{2r}(\bar{x})} (|u(y) - (u)_{B_r(\bar{x})}| + |u(z) - (u)_{B_r(\bar{x})}|) \frac{dy}{|y - \bar{x}|^{n+2s}} \right\} dz \\ &= c r^{-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_{2r}(\bar{x})) \int_{B_r(\bar{x})} |(u - h)(z)| dz + c r^{-2s} \int_{B_r(\bar{x})} |u - h| |u - (u)_{B_r(\bar{x})}| dz \\ &\leq c r^{n/2-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_{2r}(\bar{x})) \left(\int_{B_r(\bar{x})} |u - h|^2 dz \right)^{1/2} \\ &\quad + c r^{-2s} \left(\int_{B_r(\bar{x})} |u - h|^2 dz \right)^{1/2} \left(\int_{B_r(\bar{x})} |u - (u)_{B_r(\bar{x})}|^2 dz \right)^{1/2} \end{aligned}$$

$$\begin{aligned}
(7.14) \quad & \leq c r^{n/2+1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_{2r}(\bar{x})) \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dz \right)^{1/2} \\
& + c r^{n/2+2(1-s)} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dz \right)^{1/2} \left(\int_{B_{2r}(\bar{x})} |Du|^2 dz \right)^{1/2},
\end{aligned}$$

with $c = c(n, s, \max\{\nu, 1\})$. Combining (7.9)–(7.12) and (7.14) with (7.8), and dividing both sides of the resulting inequality by $r^{n/2} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2}$ (which may be assumed to be nonzero) yields our desired estimate

$$\begin{aligned}
\left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} & \leq c \left\{ \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + r^{2(1-s)} \right\} \left(\int_{B_{2r}(\bar{x})} |Du|^2 dz \right)^{1/2} \\
& + c r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c r^{1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})),
\end{aligned}$$

where $c = c(n, s, \max\{\nu, 1\}, \Lambda)$, where we also used (3.17), and that

$$\frac{n}{2_*} - \frac{n}{2} = \begin{cases} 1 & \text{for } n \geq 3 \\ \text{any number} > 1 & \text{for } n = 2 \end{cases};$$

this concludes the proof. $\square$

Before proving the above-stated theorems on the gradient, we establish a few preparatory lemmas. Using the comparison estimate (7.7), we can prove the following excess decay estimate for Du .

Proposition 7.2 (Excess decay estimate). *Under the notation and assumptions of Proposition 7.1, we have*

$$\begin{aligned}
(7.15) \quad E(Du; B_\varrho(\bar{x})) & \leq c \left[\frac{\varrho}{r} + \left(\frac{r}{\varrho} \right)^{n/2} \left[\|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + r^{2(1-s)} \right] \right] E(Du; B_{2r}(\bar{x})) \\
& + c \left(\frac{r}{\varrho} \right)^{n/2} \left[\|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + r^{2(1-s)} \right] |(Du)_{B_{2r}(\bar{x})}| \\
& + c \left(\frac{r}{\varrho} \right)^{n/2} r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \left(\frac{r}{\varrho} \right)^{n/2} r^{1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})),
\end{aligned}$$

for every $0 < \varrho \leq r \leq r_{x_0}/2$, where $c = c(\mathbf{data}, \max\{\nu, 1\})$, and $E(Du; B_r(x))$ is the excess functional defined by (3.1).

Proof. As $h \in W^{1,2}(B_r(\bar{x}))$ is a solution to a system with constant coefficients, a standard decay estimate is

$$\int_{B_\varrho(\bar{x})} |Dh - (Dh)_{B_\varrho(\bar{x})}|^2 dx \leq c \left(\frac{\varrho}{r} \right)^2 \int_{B_r(\bar{x})} |Dh - (Dh)_{B_r(\bar{x})}|^2 dx,$$

for every $0 < \varrho \leq r$, where $c = c(n, N, \Lambda)$, see [34, Chapter III, Theorem 2.1 & Remark 2.3]. Using standard algebraic manipulations taking into account (3.2) and the subadditivity several times, the previous decay estimate and standard manipulations yield

$$(7.16) \quad \int_{B_\varrho(\bar{x})} |Du - (Du)_{B_\varrho(\bar{x})}|^2 dx \leq c \left(\frac{\varrho}{r} \right)^2 \int_{B_r(\bar{x})} |Du - (Du)_{B_r(\bar{x})}|^2 dx + c \left(\frac{r}{\varrho} \right)^n \int_{B_r(\bar{x})} |Du - Dh|^2 dx,$$

with $c = c(n, N, \Lambda)$ (see, for instance, the proof of Proposition 3.2 for $\beta = 1$ in [7], although the problem considered there is different). Combining (7.16) with (7.7), we get

$$\begin{aligned}
E(Du; B_\varrho(\bar{x})) & \leq c \frac{\varrho}{r} \left(\int_{B_r(\bar{x})} |Du - (Du)_{B_r(\bar{x})}|^2 dx \right)^{1/2} \\
& + c \left(\frac{r}{\varrho} \right)^{n/2} \left[\|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + r^{2(1-s)} \right] \left(\int_{B_{2r}(\bar{x})} |Du|^2 dx \right)^{1/2} \\
& + c \left(\frac{r}{\varrho} \right)^{n/2} r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \left(\frac{r}{\varrho} \right)^{n/2} r^{1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})),
\end{aligned}$$

for all $0 < \varrho \leq r$; using (3.2) and the elementary inequality

$$\left(\int_{B_{2r}(\bar{x})} |Du|^2 dx \right)^{1/2} \leq c E(Du; B_{2r}(\bar{x})) + c |(Du)_{B_{2r}(\bar{x})}|$$

then yields our desired result (7.15). $\square$

Now we take advantage of the Hölder continuity of u in order to refine the tail estimate: this is the content of the next proposition.

Proposition 7.3 (Tail estimate for Hölder continuous functions). *Let $w \in C^{0,\beta}(B_R(x_0)) \cap L^1_{2s}$ for some $\beta \in (0, 1]$. Then for every $\bar{x} \in B_{R/2}(x_0)$, and every radius $0 < r \leq R/4$, we have*

$$\begin{aligned} r^{-2s} \mathbf{Tail}(w - (w)_{B_r(\bar{x})}; B_r(\bar{x})) &\leq c R^{-2s} \mathbf{Tail}(w - (w)_{B_R(x_0)}; B_R(x_0)) \\ &\quad + c [w]_{C^{0,\beta}(B_R(x_0))} \left[R^{\beta-2s} + \int_r^R \mu^{\beta-2s} \frac{d\mu}{\mu} \right], \end{aligned}$$

for a constant $c = c(n, s)$.

Proof. By (3.18), we have for $c = c(n, s)$

$$\begin{aligned} r^{-2s} \mathbf{Tail}(w - (w)_{B_r(\bar{x})}; B_r(\bar{x})) &\leq c R^{-2s} \mathbf{Tail}(w - (w)_{B_{R/4}(\bar{x})}; B_{R/4}(\bar{x})) \\ &\quad + c R^{-2s} \int_{B_{R/4}(\bar{x})} |w - (w)_{B_{R/4}(\bar{x})}| dx + c \int_r^{R/4} \mu^{-2s} \int_{B_\mu(\bar{x})} |w - (w)_{B_\mu(\bar{x})}| dx \frac{d\mu}{\mu}. \end{aligned}$$

By the β -Hölder continuity of w , and since $B_{R/4}(\bar{x}) \subseteq B_R(x_0)$, we have

$$\int_{B_{R/4}(\bar{x})} |w - (w)_{B_{R/4}(\bar{x})}| dx \leq [w]_{C^{0,\beta}(B_R(x_0))} R^\beta,$$

and

$$\int_r^{R/4} \mu^{-2s} \int_{B_\mu(\bar{x})} |w - (w)_{B_\mu(\bar{x})}| dx \frac{d\mu}{\mu} \leq [w]_{C^{0,\beta}(B_R(x_0))} \int_r^R \mu^{\beta-2s} \frac{d\mu}{\mu}.$$

As $B_{R/4}(\bar{x}) \subseteq B_R(x_0)$, and $R - |\bar{x} - x_0| \geq R/2$, we further estimate via (3.19)

$$\begin{aligned} \mathbf{Tail}(w - (w)_{B_{R/4}(\bar{x})}; B_{R/4}(\bar{x})) &\leq c \mathbf{Tail}(w - (w)_{B_R(x_0)}; B_R(x_0)) + c \int_{B_R(x_0)} |w - (w)_{B_R(x_0)}| dx \\ &\leq c \mathbf{Tail}(w - (w)_{B_R(x_0)}; B_R(x_0)) + c [w]_{C^{0,\beta}(B_R(x_0))} R^\beta; \end{aligned}$$

merging the content of these displays, we get the desired estimate. $\square$

7.2. Lipschitz continuity of solutions. We are now ready to prove the Lipschitz regularity of solutions to (1.1). The estimates established above provide the two ingredients needed for the iteration: the excess decay estimate (7.20), which controls the oscillation of Du at smaller scales in terms of its oscillation at larger scales, and the tail estimate of Proposition 7.3, which allows us to handle the nonlocal contribution with sufficient precision by using the local Hölder estimate already obtained in Theorem 6.1. We shall combine these estimates with a suitable choice of the Hölder exponent β_0 in order to absorb the lower-order terms and obtain a uniform bound for the gradient excess.

Proof of Theorem 2.5. As in the previous subsections, we shall consider $\bar{x} \in B_{r_{x_0}/2}(x_0)$ and a radius $0 < r \leq r_{x_0}/4$ so that $B_{2r}(\bar{x}) \subset B_{r_{x_0}}(x_0) \Subset \Omega_u$. As remarked at the beginning of this section (see the discussion preceding equation (7.3)), we are working under assumptions that are stronger than those leading to Theorem 6.1; thus by (2.4), (2.18) and the monotonicity of ω_u

$$\begin{aligned} \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} &\leq \sup_{x,y \in B_r(\bar{x})} |A(x, u(x)) - A(y, u(y))| \\ &\leq \sup_{x,y \in B_r(\bar{x})} |A(x, u(x)) - A(y, u(x))| + \sup_{x,y \in B_r(\bar{x})} |A(y, u(x)) - A(y, u(y))| \\ (7.17) \quad &\leq \Lambda \left[\omega_\times(r) + \omega_u \left(\sup_{x,y \in B_r(\bar{x})} |u(y) - u(x)| \right) \right], \end{aligned}$$

and then using that $u \in C^{0,\beta_0}(B_r(\bar{x}))$, we get

$$(7.18) \quad \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} \leq \Lambda \left[\omega_x(r) + \omega_u([u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} r^{\beta_0}) \right].$$

If we now define

$$(7.19) \quad \omega_{\beta_0}(r) := \omega_x(r) + \omega_u([u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} r^{\beta_0}) + r^{2(1-s)},$$

then combining (7.15) and (7.18) yields

$$(7.20) \quad E(Du; B_\varrho(\bar{x})) \leq c \left[\frac{\varrho}{r} + \left(\frac{r}{\varrho} \right)^{n/2} \omega_{\beta_0}(r) \right] E(Du; B_{2r}(\bar{x})) + c \left(\frac{r}{\varrho} \right)^{n/2} \omega_{\beta_0}(r) |(Du)_{B_{2r}(\bar{x})}| \\ + c \left(\frac{r}{\varrho} \right)^{n/2} r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \left(\frac{r}{\varrho} \right)^{n/2} r^{1-2s} \mathbf{Tail}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})),$$

for all $0 < \varrho \leq r \leq r_{x_0}/4$, where $c = c(\mathbf{data}, \max\{\nu, 1\})$.

Let us now consider a radius $\bar{r} \leq r_{x_0}/4$ (that we use as smallness parameter) and a parameter $\tau \in (0, 1/16)$ to be chosen later and we specialize the previous estimate to points $\bar{x} \in B_{\bar{r}}(x_0)$. We build the shrinking sequence of radii $\varrho_i = \tau^i \bar{r}$ for $r \leq \bar{r}$ and $i \in \mathbb{N}_0$, and subsequently set

$$(7.21) \quad B_i \equiv B_{\varrho_i}(\bar{x}), \quad k_i := |(Du)_{B_i}|.$$

First we observe that for $m \in \mathbb{N}_0$, we have by telescopic summation (see [7, Equation (4.5)])

$$(7.22) \quad k_{m+1} \leq k_0 + \sum_{i=0}^m \int_{B_{i+1}} |Du - (Du)_{B_i}| dx \leq k_0 + \tau^{-n} \sum_{i=0}^m E(Du; B_i),$$

enlarging the domain of integration, using Hölder's inequality and the definition of $E(Du; B_i)$. Next, by applying (7.20) with the choice $\varrho = \varrho_{i+1}$ and $r = \varrho_i/2$, we obtain

$$(7.23) \quad E(Du; B_{i+1}) \leq c \left[\tau + \tau^{-n/2} \omega_{\beta_0}(\bar{r}) \right] E(Du; B_i) \\ + c \tau^{-n/2} \left[\omega_{\beta_0}(\varrho_i) k_i + \varrho_i \left(\int_{B_i} |f|^{2^*} dx \right)^{1/2^*} + \varrho_i^{1-2s} \mathbf{Tail}(u - (u)_{B_{\varrho_i/2}(\bar{x})}; B_{\varrho_i/2}(\bar{x})) \right]$$

where $c = c(\mathbf{data}, \max\{\nu, 1\})$. We now first choose $\tau = \tau(\mathbf{data}, \max\{\nu, 1\}) \in (0, 1/16)$, and then choose the radius $\bar{r} \leq r_{x_0}/4$, depending on $\mathbf{data}, \max\{\nu, 1\}$ and also on $\omega_x(\cdot)$, $\omega_u(\cdot)$, $[u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))}$ sufficiently small so that

$$(7.24) \quad c \left[\tau + \tau^{-n/2} \omega_{\beta_0}(\bar{r}) \right] \leq \frac{1}{2}.$$

In particular, recalling that $\beta_0 = \beta_0(s)$ is fixed, by definition of ω_{β_0} in (7.19), the monotonicity of ω_u and the Hölder estimate (7.3), the radius $\bar{r}$ depends on $\mathbf{data}, \max\{\nu, 1\}, \omega_x(\cdot), \omega_u(\cdot), \varrho_{x_0}, \|I_{1,\chi}^f(\cdot, 1)\|_{L^\infty(\Omega)}$ and on an upper bound on $E(u; B_{\varrho_{x_0}}(x_0))$ and the tail term $\mathbf{Tail}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0))$.

Next, by using (7.24), equation (7.23) becomes

$$(7.25) \quad E(Du; B_{i+1}) \leq \frac{1}{2} E(Du; B_i) + c \left[\omega_{\beta_0}(\varrho_i) k_i + \varrho_i \left(\int_{B_i} |f|^{2^*} dx \right)^{1/2^*} + \varrho_i^{1-2s} \mathbf{Tail}(u - (u)_{B_{\varrho_i/2}(\bar{x})}; B_{\varrho_i/2}(\bar{x})) \right] \\ \leq \frac{1}{2} E(Du; B_i) + c \left[\omega_{\beta_0}(\varrho_i) k_i + \varrho_i \left(\int_{B_i} |f|^{2^*} dx \right)^{1/2^*} + \varrho_i \left[r_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)) \right] \right. \\ \left. + \frac{1}{|\beta_0 - 2s|} \varrho_i^{\beta_0-2s} [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} + \left(1 + \frac{1}{|\beta_0 - 2s|} \right) r_{x_0}^{\beta_0-2s} [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} \right]$$

for all $i \in \mathbb{N}_0$, with $c = c(\mathbf{data}, \max\{\nu, 1\})$. In the last line we used Proposition 7.3 with $r = \varrho_i/2$ and $R = r_{x_0}$; moreover, since $\beta_0 \neq 2s$, we computed

$$\int_{\varrho_i/2}^{r_{x_0}} \mu^{\beta_0-2s} \frac{d\mu}{\mu} \leq \frac{r_{x_0}^{\beta_0-2s} + \varrho_i^{\beta_0-2s}}{|\beta_0 - 2s|}.$$

By summing (7.25) for $i \in \{0, 1, \dots, m-1\}$ for $m \in \mathbb{N}$ and performing simple algebraic manipulations, similar to those from Proof of Theorem 2.3, we get

$$\begin{aligned} \sum_{i=0}^m E(Du; B_i) &\leq E(Du; B_0) + \frac{1}{2} \sum_{i=0}^m E(Du; B_i) + c \sum_{i=0}^{m-1} \omega_{\beta_0}(\varrho_i) k_i + c \sum_{i=0}^{\infty} \varrho_i \left(\int_{B_i} |f|^{2^*} dx \right)^{1/2^*} \\ &+ c [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} \sum_{i=0}^{\infty} \varrho_i^{1+\beta_0-2s} + c \left[r_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)) + r_{x_0}^{\beta_0-2s} [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} \right] \sum_{i=0}^{\infty} \varrho_i, \end{aligned}$$

with $c = c(\mathbf{data}, \max\{\nu, 1\})$. Reabsorbing the first term of the right-hand side, and using that

$$\sum_{i=0}^{\infty} \varrho_i = \sum_{i=0}^{\infty} \tau^i r = c(\mathbf{data}, \max\{\nu, 1\}) r, \quad \text{and similarly} \quad \sum_{i=0}^{\infty} \varrho_i^{1+\beta_0-2s} = c(\mathbf{data}, \max\{\nu, 1\}) r^{1+\beta_0-2s}$$

due to $\beta_0 = \beta_0(s)$ and (7.1), we find

$$\begin{aligned} (7.26) \quad \sum_{i=0}^m E(Du; B_i) &\leq 2E(Du; B_0) + c \sum_{i=0}^{m-1} \omega_{\beta_0}(\varrho_i) k_i + c \sum_{i=0}^{\infty} \varrho_i \left(\int_{B_i} |f|^{2^*} dx \right)^{1/2^*} \\ &+ c [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} r [r^{\beta_0-2s} + r_{x_0}^{\beta_0-2s}] + c r r_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)), \end{aligned}$$

with constant $c = c(\mathbf{data}, \max\{\nu, 1\})$. Combining (7.26) and (7.22) with (3.9) for $\bar{j} = 0$, $\alpha = 1$, we get

$$\begin{aligned} k_{m+1} &\leq k_0 + c E(Du; B_r(\bar{x})) + c \sum_{i=0}^{m-1} \omega_{\beta_0}(\varrho_i) k_i + c \|\mathbf{I}_{1,\chi}^f(\cdot, r_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))} \\ &+ c [u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} r [\bar{r}^{\beta_0-2s} + r_{x_0}^{\beta_0-2s}] + c r r_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)) \\ (7.27) \quad &=: M_0 + c \sum_{i=0}^{m-1} \omega_{\beta_0}(\varrho_i) k_i, \end{aligned}$$

with the obvious definition for M_0 , and with $c = c(\mathbf{data}, \max\{\nu, 1\})$. We now claim that

$$(7.28) \quad k_m \leq 2M_0 \quad \text{for all } m \in \mathbb{N}_0,$$

provided that we choose the radius $\bar{r} > 0$ sufficiently small.

We proceed by induction: the case $m = 0$ is trivial by the definition of M_0 . Let us suppose the claim is true for $m = 0, 1, \dots, m_0$ for some $m_0 \in \mathbb{N}_0$, and prove it for $m_0 + 1$. By (7.27) and the inductive hypothesis,

$$(7.29) \quad k_{m_0+1} \leq M_0 + c \sum_{i=0}^{m_0} \omega_{\beta_0}(\varrho_i) k_i \leq M_0 \left[1 + 2\bar{c} \sum_{i=0}^{\infty} \omega_{\beta_0}(\varrho_i) \right],$$

with $\bar{c} = \bar{c}(\mathbf{data}, \max\{\nu, 1\})$. Using the standard estimate (3.4), recalling that

$$\beta_0 = \beta_0(s), \quad \tau = \tau(\mathbf{data}, \max\{\nu, 1\}) \leq 1/16,$$

by the expression of (7.19), using (3.4) and the Dini continuity of $\omega_x(\cdot)$, $\omega_u(\cdot)$, we may reduce the size of $\bar{r}$ in such a way that

$$(7.30) \quad \sum_{i=0}^{\infty} \omega_{\beta_0}(\varrho_i) \leq \frac{1}{\log(1/\tau)} \int_0^{2\bar{r}} \omega_x(\varrho) \frac{d\varrho}{\varrho} + \frac{1}{\beta_0 \log(1/\tau)} \int_0^{2[u]_{C^{0,\beta_0}(B_{r_{x_0}}(x_0))} \bar{r}^{\beta_0}} \omega_u(\varrho) \frac{d\varrho}{\varrho} + \frac{\bar{r}^{2(1-s)}}{1 - \tau^{2(1-s)}} \leq \frac{1}{2\bar{c}}.$$

Combining (7.30) with (7.29), we deduce $k_{m_0+1} \leq 2M_0$, which proves the claim (7.28).

Now, for all Lebesgue points $\bar{x} \in B_r(x_0)$ of Du , by (7.28) we have

$$(7.31) \quad |Du(\bar{x})| = \lim_{m \rightarrow \infty} k_m \leq 2M_0 \quad \implies \quad \|Du\|_{L^\infty(B_r(x_0))} \leq 2M_0.$$

Recalling (7.21) and (3.2), for all $\bar{x} \in B_r(x_0)$, we estimate

$$k_0 + E(Du; B_r(\bar{x})) \leq c(n) \left(\int_{B_{2r}(x_0)} |Du|^2 dx \right)^{1/2}$$

since $B_r(\bar{x}) \subset B_{2r}(x_0)$; we then insert the estimate (7.3) into (7.31), expand the expression (7.27) of M_0 , and use that, since (5.3) holds, $1 + \beta_0 - 2s > 0$ and $r \leq r_{x_0} \leq \varrho_{x_0} \leq 1$,

$$r[r^{\beta_0-2s} + r_{x_0}^{\beta_0-2s}] \leq r^{1+\beta_0-2s} + r_{x_0}^{1+\beta_0-2s} \leq 2$$

and similarly

$$\varrho_{x_0}^{\gamma-\beta_0-2s} r[r^{\beta_0-2s} + r_{x_0}^{\beta_0-2s}] \leq 2\varrho_{x_0}^{\gamma-4s} \leq 2\varrho_{x_0}^{-2s}.$$

This yields

$$\begin{aligned} \|Du\|_{L^\infty(B_r(x_0))} &\leq c r^{-n/2} \|Du\|_{L^2(B_{2r}(x_0))} + c \varrho_{x_0}^{-n/2} \|Du\|_{L^2(B_{\varrho_{x_0}}(x_0))} \\ &\quad + c \|\mathbf{I}_{1,\chi}^f(\cdot, r_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))} + c r r_{x_0}^{-2s} \mathbf{Tal}(u - (u)_{B_{r_{x_0}}(x_0)}; B_{r_{x_0}}(x_0)) \\ &\quad + c r \varrho_{x_0}^{-2s} \mathbf{Tal}(u - (u)_{B_{\varrho_{x_0}}(x_0)}; B_{\varrho_{x_0}}(x_0)), \end{aligned}$$

with $c = c(\mathbf{data}, \max\{\nu, 1\})$, and with $r \leq \bar{r} \leq r_{x_0}/4$ fulfilling (7.24) and (7.30). This completes the proof of (2.21) (with $\bar{r}_{x_0} := \bar{r}$). $\square$

Remark 7.4. Taking advantage of the Lipschitz continuity of u , we can improve (7.20). As we want to make use of Theorem 2.5 in (7.17), we start from points $\bar{x} \in B_{\bar{r}_{x_0}/2}(x_0)$ and radii $r \leq \bar{r}_{x_0}/4$ and get

$$(7.32) \quad \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} \leq \Lambda \left[\omega_x(r) + \omega_u([u]_{C^{0,1}(B_{\bar{r}_{x_0}}(x_0))} r) \right],$$

with $\bar{A}$ as in (7.4). Additionally, we can apply Proposition 7.3 with $\beta = 1$, thus getting

$$(7.33) \quad r^{1-2s} \mathbf{Tal}(u - (u)_{B_r(\bar{x})}; B_r(\bar{x})) \leq c r \bar{r}_{x_0}^{-2s} \mathbf{Tal}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) \\ + c r \left[\bar{r}_{x_0}^{1-2s} + \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}.$$

Substituting this into (7.15), and taking into account that $|(Du)_{B_{2r}(\bar{x})}| \leq \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}$, gives

$$(7.34) \quad E(Du; B_\varrho(\bar{x})) \leq c \left[\frac{\varrho}{r} + \left(\frac{r}{\varrho} \right)^{n/2} \omega_1(r) \right] E(Du; B_{2r}(\bar{x})) + c \left(\frac{r}{\varrho} \right)^{n/2} \left\{ \omega_1(r) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right. \\ \left. + c r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c r \bar{r}_{x_0}^{-2s} \mathbf{Tal}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) \right. \\ \left. + c r \left[\bar{r}_{x_0}^{1-2s} + \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right\},$$

for $0 < \varrho \leq r \leq \bar{r}_{x_0}/4$, with $c = c(\mathbf{data}, \max\{\nu, 1\})$, and where $\omega_1(\cdot)$ is defined in (7.19) with $\beta_0 = 1$, that is

$$\omega_1(r) := \omega_x(r) + \omega_u([u]_{C^{0,1}(B_{\bar{r}_{x_0}}(x_0))} r) + r^{2(1-s)}.$$

We recall that u is the continuous representative of the Lebesgue class in Ω_u .

7.3. Gradient continuity of solutions. We prove here the continuity of the gradient under the additional assumption that the Riesz potential of the datum tends to zero locally uniformly as the radius vanishes. The proof is divided into two steps: first we show that the local excess of the gradient vanishes locally uniformly at small scales. Then, by summing the decay estimate (7.20) along a dyadic sequence of radii, we prove that the corresponding averages of the gradient form a Cauchy sequence; the conclusion then follows as in the proof of Theorem 2.3. Again, $x_0 \in \Omega_u$ and $0 < \bar{r}_{x_0} \leq r_{x_0}/4 \leq \varrho_{x_0}/16$ are such that $B_{r_{x_0}}(x_0) \Subset \Omega_u$, $B_{2\varrho_{x_0}}(x_0) \Subset \Omega$ and the local estimate (2.21) holds.

Proof of Theorem 2.6. Step 1: A VMO-type result. We first prove a local VMO-type estimate for Du , uniform in $B_{\bar{r}_{x_0}/4}(x_0)$. Specifically, we show that for every $\varepsilon \in (0, 1)$, there exists a radius $r_\varepsilon \leq \bar{r}_{x_0}/4$ small enough and independent of $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$, such that

$$(7.35) \quad E(Du; B_r(\bar{x})) = \left(\int_{B_r(\bar{x})} |Du - (Du)_{B_r(\bar{x})}|^2 dx \right)^{1/2} < \varepsilon$$

for all $r \in (0, r_\varepsilon]$ and for all $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$. To this end, let us fix $R_\varepsilon > 0$, to be determined in the course of the proof, and let $r > 0$ be arbitrary with $0 < 2r \leq R_\varepsilon \leq \bar{r}_{x_0}/2$. By this choice and (3.9), we have

$$r \left(\int_{B_r(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} \leq c \mathbf{I}_{1,\chi}^f(\bar{x}, 2r) \leq c \mathbf{I}_{1,\chi}^f(\bar{x}, R_\varepsilon)$$

and clearly $E(Du; B_{2r}(\bar{x})) \leq 2 \|Du\|_{L^\infty(B_{\bar{r}_{x_0}/4}(x_0))}$ for all $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$. Let us now fix $\tau_\varepsilon \in (0, 1)$ to be chosen later, and combine the last two estimates with (7.34), choosing $\varrho = \tau_\varepsilon r$. Recalling the monotonicity of $\omega_1(\cdot)$ and that $r \leq R_\varepsilon \leq \bar{r}_{x_0}/2$, we thus obtain

$$(7.36) \quad E(Du; B_{\tau_\varepsilon r}(\bar{x})) \leq \tilde{c} \left[\tau_\varepsilon + \tau_\varepsilon^{-n/2} \left[\omega_1(R_\varepsilon) + R_\varepsilon \bar{r}_{x_0}^{1-2s} + \sup_{0 < r \leq R_\varepsilon} r \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right. \\ \left. + \tilde{c} \tau_\varepsilon^{-n/2} \mathbf{I}_{1,\chi}^f(\bar{x}, R_\varepsilon) + \tilde{c} R_\varepsilon \bar{r}_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) \right],$$

with $\tilde{c}$ depending on **data** and $\max\{\nu, 1\}$. We can now fix τ_ε and R_ε : for $\varepsilon > 0$ fixed, we choose

$$(7.37) \quad \tau_\varepsilon = \frac{\varepsilon}{4\tilde{c} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} + 1},$$

and then we select $R_\varepsilon > 0$ independent of $\bar{x}$ such that

$$\tau_\varepsilon^{-n/2} \left[\omega_1(R_\varepsilon) + R_\varepsilon \bar{r}_{x_0}^{1-2s} + \sup_{0 < r \leq R_\varepsilon} r \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \leq \frac{\varepsilon}{4\tilde{c}}$$

(this is possible since $\sup_{0 < r \leq R_\varepsilon} r \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \searrow 0$ as $R_\varepsilon \rightarrow 0$) and

$$R_\varepsilon \bar{r}_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) \leq \frac{\varepsilon}{4\tilde{c}}.$$

Moreover, thanks to our assumption (2.22), we may further reduce R_ε , independently of $\bar{x}$, in such a way that

$$(7.38) \quad \tau_\varepsilon^{-n/2} \mathbf{I}_{1,\chi}^f(\bar{x}, R_\varepsilon) < \frac{\varepsilon}{4\tilde{c}};$$

merging the content of equations (7.37)–(7.38) with (7.36), we obtain

$$E(Du; B_{\tau_\varepsilon r}(\bar{x})) < \varepsilon \quad \text{for all } r \leq R_\varepsilon/2, \text{ for all } \bar{x} \in B_{\bar{r}_{x_0}/4}(x_0),$$

that is (7.35) with $r_\varepsilon = \tau_\varepsilon R_\varepsilon/2$.

Step 2: Du is continuous. We let $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$, and let $\tau \in (0, 1/2)$ a parameter and $R_1 \leq \bar{r}_{x_0}/2$ a radius both be determined later; we set, for $i \in \mathbb{N}$ $\varrho_i = \tau^i R_1$, $B_i(\bar{x}) = B_{\varrho_i}(\bar{x})$. In order to prove that Du is continuous, it suffices to show that

$$(7.39) \quad \bar{x} \mapsto \{(Du)_{B_i(\bar{x})}\}_{i \in \mathbb{N}_0} \text{ is a Cauchy sequence, uniformly in } B_{\bar{r}_{x_0}/4}(x_0)$$

(see indeed the argument at the end of the proof of Theorem 2.3, applied to Du).

Let us prove (7.39). First observe that, for $1 \leq k < m$, using the definition of ϱ_i and Hölder's inequality

$$(7.40) \quad |(Du)_{B_k(\bar{x})} - (Du)_{B_m(\bar{x})}| \leq \sum_{i=k}^{m-1} \int_{B_{i+1}} |Du - (Du)_{B_i(\bar{x})}| dx \leq \tau^{-n} \sum_{i=k}^{m-1} E(Du; B_i(\bar{x})),$$

see for instance (6.12). We then make use of (7.34) with $\varrho = \varrho_{i+1}$ and $r = \varrho_i/2$ for $i \in \mathbb{N}_0$, together with $|(Du)_{B_i(\bar{x})}| \leq \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}$ and the fact that ω_1 is monotone increasing, so that we find

$$(7.41) \quad E(Du; B_{i+1}(\bar{x})) \leq \tilde{c} \left\{ \tau + \tau^{-n/2} \omega_1(\varrho_i) \right\} E(Du; B_i(\bar{x})) + c \tau^{-n/2} \omega_1(\varrho_i) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \\ + c \tau^{-n/2} \varrho_i \left(\int_{B_i(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \tau^{-n/2} \varrho_i \bar{r}_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) \\ + c \tau^{-n/2} \varrho_i \bar{r}_{x_0}^{1-2s} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} + c \tau^{-n/2} \varrho_i \int_{\varrho_i/2}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))},$$

where $c, \tilde{c}$ depend on $\mathbf{data}$ and $\max\{\nu, 1\}$. We now fix $\tau = \tau(\mathbf{data}, \max\{\nu, 1\})$ and then $R_1 > 0$, both independent of $\bar{x}$, such that

$$\tilde{c} \left\{ \tau + \tau^{-n/2} \omega_1(R_1) \right\} \leq \frac{1}{2}.$$

Combining this information with (7.41), using $\omega_1(\varrho_i) \leq \omega_1(R_1)$ and recalling that now τ depends on $\mathbf{data}$ gives

$$\begin{aligned} E(Du; B_{i+1}(\bar{x})) &\leq \frac{1}{2} E(Du; B_i(\bar{x})) + c \omega_1(\varrho_i) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \\ &\quad + c \varrho_i \left(\int_{B_i(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \\ &\quad + c \varrho_i \bar{r}_{x_0}^{-2s} \left[\mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + \bar{r}_{x_0} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right]. \end{aligned}$$

We take the sum over $i = k, \dots, m$ and, after changing variables and enlarging summations on the right-hand side, we get

$$\begin{aligned} \sum_{i=k+1}^{m+1} E(Du; B_i(\bar{x})) &\leq \frac{1}{2} \sum_{i=k}^m E(Du; B_i(\bar{x})) + c \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \sum_{i=k}^{\infty} \omega_1(\varrho_i) \\ &\quad + c \sum_{i=k}^m \varrho_i \left(\int_{B_i(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} + c \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \sum_{i=k}^{\infty} \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \\ &\quad + c \bar{r}_{x_0}^{-2s} \left[\mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + \bar{r}_{x_0} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right] \sum_{i=k}^{\infty} \varrho_i; \end{aligned}$$

reabsorbing the first term on the right-hand side, letting $m \rightarrow +\infty$, using the definition of the radii ϱ_i , the choice of $\tau = \tau(\mathbf{data}, \max\{\nu, 1\})$ and the fact that $R_1 \leq \bar{r}_{x_0}$, the above inequality entails

$$\begin{aligned} (7.42) \quad \sum_{i=k}^{\infty} E(Du; B_i(\bar{x})) &\leq 2E(Du; B_k(\bar{x})) + c \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \sum_{i=k}^{\infty} \omega_1(\varrho_i) + c \sum_{i=k}^{\infty} \varrho_i \left(\int_{B_i(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} \\ &\quad + c \tau^k \left[\bar{r}_{x_0}^{1-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + \bar{r}_{x_0}^{2(1-s)} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right] \\ &\quad + c \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \sum_{i=k}^{\infty} \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu}. \end{aligned}$$

Now we recall that, by (3.4) ($\omega_1(\cdot)$ is indeed a modulus of continuity) and (3.9),

$$\sum_{i=k}^{\infty} \omega_1(\varrho_i) \leq c \int_0^{\varrho_{k-1}} \omega_1(t) \frac{dt}{t} \quad \text{and} \quad \sum_{i=k}^{\infty} \varrho_i \left(\int_{B_i(\bar{x})} |f|^{2^*} dx \right)^{1/2^*} \leq c \mathbf{I}_{1,\chi}^f(\bar{x}, \varrho_{k-1}),$$

and we notice that

$$\sum_{i=k}^{\infty} \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \leq c(s, \tau) \tau^{k \min\{1, 2(1-s)\}} \bar{r}_{x_0}^{2(1-s)} \left[\log \frac{\bar{r}_{x_0}}{R_1} + k \right].$$

Indeed, if $s \neq 1/2$, then

$$\int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} = \frac{\bar{r}_{x_0}^{1-2s} - \varrho_{i+1}^{1-2s}}{1-2s} \leq c(s, \tau) \max \{ \bar{r}_{x_0}^{1-2s}, \varrho_i^{1-2s} \}$$

so that, recalling that $\varrho_i \leq R_1 \leq \bar{r}_{x_0}/2$ and that $2(1-s) > 0$,

$$\sum_{i=k}^{\infty} \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \leq c(s) \sum_{i=k}^{\infty} \varrho_i \max \{ \bar{r}_{x_0}^{1-2s}, \varrho_i^{1-2s} \} \leq c(s, \tau) \bar{r}_{x_0}^{2(1-s)} \max \{ \tau^k, \tau^{2k(1-s)} \};$$

if $s = 1/2$, then

$$\int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{-1} d\mu = \log \frac{\bar{r}_{x_0}}{\tau^{i+1} R_1} = \log \frac{\bar{r}_{x_0}}{R_1} + (i+1) \log \frac{1}{\tau},$$

so that

$$\sum_{i=k}^{\infty} \varrho_i \int_{\varrho_{i+1}}^{\bar{r}_{x_0}} \mu^{-1} d\mu \leq c(\tau) R_1 \sum_{i=k}^{\infty} \tau^i \left(\log \frac{\bar{r}_{x_0}}{R_1} + i + 1 \right) \leq c \bar{r}_{x_0} \tau^k \left[\log \frac{\bar{r}_{x_0}}{R_1} + k + 1 \right].$$

Inserting this information into (7.42) yields

$$(7.43) \quad \sum_{i=k}^{\infty} E(Du; B_i(\bar{x})) \leq 2E(Du; B_k(\bar{x})) + \tilde{c} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \int_0^{\varrho_{k-1}} \omega_1(t) \frac{dt}{t} + \tilde{c} \mathbf{I}_{1,\chi}^f(\bar{x}, \varrho_{k-1}) \\ + c \bar{r}_{x_0}^{1-2s} \left[\tau^k \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + \tau^{k \min\{1, 2(1-s)\}} \bar{r}_{x_0} \left[\log \frac{\bar{r}_{x_0}}{R_1} + k \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right],$$

with $c, \tilde{c}$ depending on **data** and $\max\{\nu, 1\}$. By *Step 1*, see (7.35), we have that

$$\lim_{k \rightarrow \infty} E(Du; B_k(\bar{x})) = 0 \quad \text{uniformly in } \bar{x} \in B_{\bar{r}_{x_0}/4}(x_0).$$

Using this information, the Dini continuity of $\omega_1(\cdot)$, (2.22), and since $\tau = \tau(\mathbf{data}, \max\{\nu, 1\}) \in (0, 1/16)$, from (7.43) we deduce that for any $\varepsilon > 0$ we may find $k_\varepsilon > 0$, independent of $\bar{x}$, such that

$$\sum_{i=k}^{\infty} \sup_{\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)} E(Du; B_i(\bar{x})) < \tau^n \varepsilon \quad \text{for all } k \geq k_\varepsilon;$$

substituting this estimate into (7.40) gives

$$|(Du)_{B_k(\bar{x})} - (Du)_{B_m(\bar{x})}| < \varepsilon \quad \text{for all } k_\varepsilon \leq k < m \text{ and all } \bar{x} \in B_{\bar{r}_{x_0}/4}(x_0),$$

so that (7.39) is proved, thus completing the proof. $\square$

7.4. Gradient Hölder continuity. Carefully tracking the dependences in the previous proof gives the proof of Theorem 2.7, once the data are sufficiently regular.

Proof of Theorem 2.7. We fix an exponent χ such that $2_* < \chi < n$; since $n < n/(1-\beta)$, this choice is admissible in (3.5) with $m = n/(1-\beta)$. By (2.1) and (3.5), we have for $\bar{x} \in B_{r_{x_0}}(x_0)$ and $r \leq r_{x_0}$

$$(7.44) \quad r \left(\int_{B_r(\bar{x})} |f|^{2_*} dx \right)^{1/2_*} + r \left(\int_{B_r(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq c(n, \beta) r^\beta \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_r(\bar{x}))},$$

which in particular implies

$$\|\mathbf{I}_{1,\chi}^f(\cdot, r_{x_0})\|_{L^\infty(B_{r_{x_0}}(x_0))} \leq c(n, \beta) \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))};$$

combining this inequality with (2.21), the new Lipschitz continuity estimate (2.25) follows.

Let us now fix $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$, and let $r \leq \bar{r}_{x_0}/4$. Rewriting estimate (7.34) using (7.44) and also $E(Du; B_{2r}(\bar{x})) \leq 2\|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}$ due to the inclusion $B_{2r}(\bar{x}) \subset B_{\bar{r}_{x_0}}(x_0)$, yields

$$(7.45) \quad E(Du; B_\varrho(\bar{x})) \leq c \frac{\varrho}{r} E(Du; B_{2r}(\bar{x})) + c \left(\frac{r}{\varrho} \right)^{n/2} \left\{ \omega_1(r) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} + c r^\beta \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\bar{r}_{x_0}}(x_0))} \right. \\ \left. + c r \bar{r}_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + c r \left[\bar{r}_{x_0}^{1-2s} + \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right\},$$

where we can redefine

$$\omega_1(r) = \left[1 + \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}^\beta \right] r^{\min\{\beta, 2(1-s)\}}.$$

We now notice the elementary inequality

$$r \left[\bar{r}_{x_0}^{1-2s} + \int_r^{\bar{r}_{x_0}} \mu^{1-2s} \frac{d\mu}{\mu} \right] \leq c (1 + \bar{r}_{x_0}^{1-2s}) r^{\min\{\beta, 2(1-s)\}}.$$

valid for every $r \in (0, \bar{r}_{x_0}/4]$: the only non-trivial case in when $s = 1/2$, and in this case we can estimate

$$r \int_r^{\bar{r}_{x_0}} \mu^{-1} d\mu = r \log \frac{\bar{r}_{x_0}}{r} \leq c(\beta) \bar{r}_{x_0}^{1-\beta} r^\beta$$

as $\beta < 1$. If we insert this into (7.45), after multiplying both sides by $\varrho^{n/2}$ and factoring out the smallest power of r appearing in the error terms, namely $r^{\min\{\beta, 2(1-s)\}} \geq r$, after a few algebraic manipulations, we are lead to

$$(7.46) \quad \left(\int_{B_{\varrho}(\bar{x})} |Du - (Du)_{B_{\varrho}(\bar{x})}|^2 dx \right)^{1/2} \leq \widehat{C} r^{n/2 + \min\{\beta, 2(1-s)\}} + c \left(\frac{\varrho}{r} \right)^{n/2+1} \left(\int_{B_{2r}(\bar{x})} |Du - (Du)_{B_{2r}(\bar{x})}|^2 dx \right)^{1/2},$$

where we set

$$\widehat{C} = c(\mathbf{data}, \max\{\nu, 1\}, \beta) \left[\|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}^{1+\beta} + (1 + \bar{r}_{x_0}^{1-2s}) \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} + \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))} + \bar{r}_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{\bar{r}_{x_0}}(x_0)}; B_{\bar{r}_{x_0}}(x_0)) + 1 \right].$$

Therefore, by setting

$$\phi(\varrho) \equiv \phi_{\bar{x}}(\varrho) := \left(\int_{B_{\varrho}(\bar{x})} |Du - (Du)_{B_{\varrho}(\bar{x})}|^2 dx \right)^{1/2}$$

from (7.46) we have, absorbing the powers of 2 into the constants,

$$(7.47) \quad \phi(\varrho) \leq c \left(\frac{\varrho}{R} \right)^{n/2+1} \phi(R) + c \widehat{C} R^{n/2 + \min\{\beta, 2(1-s)\}}.$$

for all $0 < \varrho \leq R = 2r \leq \bar{r}_{x_0}/2$, as the estimate is trivial for $\varrho \in (R/2, R]$. Additionally, by means of (3.2) it is immediate to verify that $\varrho \mapsto \phi(\varrho)$ is a monotone function. This observation and (7.47) allow us to apply a standard iteration lemma [34, Lemma 2.1, Chapter III], which yields

$$\phi(r) \leq c \left[\frac{\phi(\bar{r}_{x_0}/4)}{(\bar{r}_{x_0})^{n/2 + \min\{\beta, 2(1-s)\}}} + \widehat{C} \right] r^{n/2 + \min\{\beta, 2(1-s)\}};$$

from this estimate and (3.2), since $B_{\bar{r}_{x_0}/4}(\bar{x}) \subseteq B_{\bar{r}_{x_0}}(x_0)$, we get

$$E(Du; B_r(\bar{x})) \leq c \left[\bar{r}_{x_0}^{-\min\{\beta, 2(1-s)\}} E(Du; B_{\bar{r}_{x_0}}(x_0)) + \widehat{C} \right] r^{\min\{\beta, 2(1-s)\}}$$

for all $0 < r \leq \bar{r}_{x_0}/4$ and all $\bar{x} \in B_{\bar{r}_{x_0}/4}(x_0)$, where $c = c(\mathbf{data}, \max\{\nu, 1\}, \beta)$. This estimate then yields our desired result (2.26) by Campanato's characterization of Hölder continuity [34, Chapter 3.1]. $\square$

We now move on to the proof of Theorem 2.8; to this end, we need to improve the exponent $2(1-s)$ appearing in the scale factor of the comparison estimate (7.7). This in turn boils down to improving the nonlocal estimates (7.11), (7.12) and (7.14), and this is where we will use the particular form of the nonlocal term Q_Φ^{nl} given by (2.28) and the newfound Hölder continuity of Du obtained in Theorem 2.7.

Before proving Theorem 2.8, we introduce the following notation and make a preliminary remark. Under assumption (2.28), we write $Q_\Phi^{\text{nl}} = Q_a^{\text{nl}}$.

Remark 7.5. For any affine function $\ell : \mathbb{R}^N \rightarrow \mathbb{R}^N$, following [41, Remark 3.4], we show that

$$(7.48) \quad Q_a^{\text{nl}} \ell = 0 \quad \text{in } \mathbb{R}^n$$

that holds in the P.V. sense if $s \geq 1/2$. In fact, write $\ell(x) = Mx + b$, for some matrix $M \in \mathbb{R}^{N \times n}$ and vector $b \in \mathbb{R}^N$. Then, for all $x \in \mathbb{R}^n$, we have

$$(7.49) \quad \int_{\mathbb{R}^n} a(x-y) (\ell(x) - \ell(y)) \frac{dy}{|y-x|^{n+2s}} = \int_{\mathbb{R}^n} a(x-y) M(x-y) \frac{dy}{|y-x|^{n+2s}} = \int_{\mathbb{R}^n} a(z) Mz \frac{dz}{|z|^{n+2s}} = 0,$$

where the last equality stems from

$$\int_{\mathbb{R}^n} a_{\alpha\beta}(z) M_{\beta j} z_j \frac{dz}{|z|^{n+2s}} = 0 \quad \text{for all } \alpha, \beta = 1, \dots, N, j = 1, \dots, n,$$

since the integrand is odd by (2.27); these integrals are to be meant in the P.V. sense. From (7.49), equation (7.48) follows.

Let us now introduce the following shorthand notation. For $x_0 \in \Omega$ and $\varrho > 0$, we define the affine function

$$(7.50) \quad \ell_{x_0, \varrho}(x) \equiv \ell_\varrho(x) := (u)_{B_\varrho(x_0)} + (Du)_{B_\varrho(x_0)}(x - x_0), \quad x \in \mathbb{R}^n.$$

In the following, we will use several times the following inequality: for all $y \in \mathbb{R}^n$ and $0 < \varrho \leq r$, we have

$$(7.51) \quad |\ell_{x_0, \varrho} - \ell_{x_0, r}|(x) \leq \left(\frac{r}{\varrho}\right)^n \int_{B_r(x_0)} |u - \ell_{x_0, r}| dx + \left(\frac{r}{\varrho}\right)^n |x - x_0| \int_{B_r(x_0)} |Du - (Du)_{B_r(x_0)}| dx.$$

To prove it, without loss of generality we may assume that $x_0 = 0$, and we set $(u)_r = (u)_{B_r}$ and $\ell_\varrho = \ell_{x_0, \varrho}$; then observe that $\int_{B_\varrho} (Du)_r x dx = 0$, so we have

$$\begin{aligned} |\ell_\varrho - \ell_r|(x) &= \left| \int_{B_\varrho} [u(z) + Du(z)x - (u)_r - (Du)_r x] dz \right| \\ &= \left| \int_{B_\varrho} [u(z) - (u)_r - (Du)_r z + Du(z)x - (Du)_r x] dz \right| \\ &\leq \int_{B_\varrho} |u - \ell_r|(z) dz + |[(Du)_\varrho - (Du)_r]x| \\ &\leq \left(\frac{r}{\varrho}\right)^n \int_{B_r} |u - \ell_r| dz + \left(\frac{r}{\varrho}\right)^n |x| \int_{B_r} |Du - (Du)_r| dz \end{aligned}$$

and (7.51) follows.

Lemma 7.6. *Let $0 < R_0 \leq 1$, and let $u \in W^{1,1}(B_{R_0}(x_0); \mathbb{R}^N) \cap L^1_{2s}$; suppose that $s > 1/2$. Then there exists a constant $c = c(n, N, s)$ such that*

$$(7.52) \quad \begin{aligned} \text{Tail}(u - \ell_{x_0, r}; B_r(x_0)) &\leq c \left(\frac{r}{R}\right)^{2s} \text{Tail}(u - \ell_{x_0, R}; B_R(x_0)) \\ &+ c \int_r^R \left(\frac{r}{\mu}\right)^{2s} \int_{B_\mu(x_0)} |u - \ell_{x_0, \mu}| dx \frac{d\mu}{\mu} + c r \int_r^R \left(\frac{r}{\mu}\right)^{2s-1} \int_{B_\mu(x_0)} |Du - (Du)_{B_\mu(x_0)}| dx \frac{d\mu}{\mu} \\ &+ c \left(\frac{r}{R}\right)^{2s} \int_{B_R(x_0)} |u - \ell_{x_0, R}| dx + c \left(\frac{r}{R}\right)^{2s} R \int_{B_R(x_0)} |Du - (Du)_{B_R(x_0)}| dx \end{aligned}$$

for all $0 < r \leq R \leq R_0$. In particular, by Poincaré's inequality, there holds

$$(7.53) \quad \begin{aligned} \text{Tail}(u - \ell_{x_0, r}; B_r(x_0)) &\leq c \left(\frac{r}{R}\right)^{2s} \text{Tail}(u - \ell_{x_0, R}; B_R(x_0)) \\ &+ c r \int_r^R \left(\frac{r}{\mu}\right)^{2s-1} \int_{B_\mu(x_0)} |Du - (Du)_{B_\mu(x_0)}| dx \frac{d\mu}{\mu} + c \left(\frac{r}{R}\right)^{2s} R \int_{B_R(x_0)} |Du - (Du)_{B_R(x_0)}| dx. \end{aligned}$$

Proof. By translation, we may assume that $x_0 = 0$ and set $\ell_\varrho = \ell_{0, \varrho}$. The proof follows the argument of [26, Lemma 3.2]; since the replacement of the constant averages $(u)_{B_\rho}$ by the affine maps ℓ_ρ produces additional terms involving the oscillation of Du , we include the details. We split

$$\begin{aligned} \text{Tail}(u - \ell_r; B_r) &\leq r^{2s} \int_{\mathbb{R}^n \setminus B_R} |u - \ell_r| \frac{dx}{|x|^{n+2s}} + r^{2s} \int_{B_R \setminus B_r} |u - \ell_r| \frac{dx}{|x|^{n+2s}} \\ &\leq \left(\frac{r}{R}\right)^{2s} \text{Tail}(u - \ell_R; B_R) + r^{2s} \int_{\mathbb{R}^n \setminus B_R} |\ell_R - \ell_r| \frac{dx}{|x|^{n+2s}} + r^{2s} \int_{B_R \setminus B_r} |u - \ell_r| \frac{dx}{|x|^{n+2s}}. \end{aligned}$$

We call the last two integrals T_1 and T_2 , respectively. We first study the case $r \leq R/4$. Hence, there exist $\gamma \in (1/4, 1/2]$ and an integer $k \geq 2$ such that $r = \gamma^k R$. By telescoping and exploiting that $\gamma \in (1/4, 1/2]$ and (7.51) multiple times, we get

$$\begin{aligned} T_1 &\leq c r^{2s} \sum_{i=0}^{k-1} \int_{\mathbb{R}^n \setminus B_R} |\ell_{\gamma^i R} - \ell_{\gamma^{i+1} R}| \frac{dx}{|x|^{n+2s}} \\ &\leq c \left(\frac{r}{R}\right)^{2s} \sum_{i=1}^{k-1} \int_{B_{\gamma^i R}} |u - \ell_{\gamma^i R}| dx + c \left(\frac{r}{R}\right)^{2s} R \sum_{i=1}^{k-1} \int_{B_{\gamma^i R}} |Du - (Du)_{\gamma^i R}| dx \end{aligned}$$

$$(7.54) \quad + c \left(\frac{r}{R} \right)^{2s} \int_{B_R} |u - \ell_R| dx + c R \left(\frac{r}{R} \right)^{2s} \int_{B_R} |Du - (Du)_R| dx;$$

we remark that above we used (3.16), also replacing $2s$ with $2s - 1$, which are admissible since $s > 1/2$. Now we shorten

$$(7.55) \quad G(\rho) := \int_{B_\rho} |u - \ell_\rho| dx, \quad H(\rho) := \int_{B_\rho} |Du - (Du)_\rho| dx$$

and we notice that for every $i \in \{1, \dots, k-1\}$ and every $\mu \in [\gamma^i R, \gamma^{i-1} R]$, by (7.51), (3.2) and the fact that $\gamma \in (1/4, 1/2]$, we have $G(\gamma^i R) \leq c G(\mu) + c \mu H(\mu)$ and $H(\gamma^i R) \leq c H(\mu)$. Consequently,

$$\left(\frac{r}{R} \right)^{2s} \sum_{i=1}^{k-1} G(\gamma^i R) \leq c \int_r^R \left(\frac{r}{\mu} \right)^{2s} G(\mu) \frac{d\mu}{\mu} + c r \int_r^R \left(\frac{r}{\mu} \right)^{2s-1} H(\mu) \frac{d\mu}{\mu}$$

and, recalling that $2s > 1$,

$$\left(\frac{r}{R} \right)^{2s} R \sum_{i=1}^{k-1} H(\gamma^i R) \leq c r \int_r^R \left(\frac{r}{\mu} \right)^{2s-1} H(\mu) \frac{d\mu}{\mu}.$$

Using the first estimate to bound the first term on the right-hand side in (7.54), we have

$$(7.56) \quad \left(\frac{r}{R} \right)^{2s} \sum_{i=1}^{k-1} \int_{B_{\gamma^i R}} |u - \ell_{\gamma^i R}| dx \leq c \int_r^R \left(\frac{r}{\mu} \right)^{2s} \int_{B_\mu} |u - \ell_\mu| dx \frac{d\mu}{\mu} + c r \int_r^R \left(\frac{r}{\mu} \right)^{2s-1} \int_{B_\mu} |Du - (Du)_\mu| dx \frac{d\mu}{\mu},$$

while using the second one yields

$$(7.57) \quad \left(\frac{r}{R} \right)^{2s} R \sum_{i=1}^{k-1} \int_{B_{\gamma^i R}} |Du - (Du)_{\gamma^i R}| dx \leq c r \int_r^R \left(\frac{r}{\mu} \right)^{2s-1} \int_{B_\mu} |Du - (Du)_\mu| dx \frac{d\mu}{\mu}$$

with $c = c(n, N, s)$. Combining (7.56)–(7.57) with (7.54), we get

$$T_1 \leq c \int_r^R \left(\frac{r}{\mu} \right)^{2s} \int_{B_\mu} |u - \ell_\mu| dx \frac{d\mu}{\mu} + c r \int_r^R \left(\frac{r}{\mu} \right)^{2s-1} \int_{B_\mu} |Du - (Du)_\mu| dx \frac{d\mu}{\mu} + c \left(\frac{r}{R} \right)^{2s} \int_{B_R} |u - \ell_R| dx + c \left(\frac{r}{R} \right)^{2s} R \int_{B_R} |Du - (Du)_R| dx,$$

with $c = c(n, N, s)$. Next we estimate T_2 : by telescoping and using again that $\gamma \in (1/4, 1/2]$, we find

$$(7.58) \quad T_2 \leq r^{2s} \sum_{i=0}^{k-1} \int_{B_{\gamma^{-i-1}r} \setminus B_{\gamma^{-i}r}} |u - \ell_r| \frac{dx}{|x|^{n+2s}} \leq c(n, s) \sum_{i=0}^k \gamma^{i2s} \int_{B_{\gamma^{-i}r}} |u - \ell_r| dx.$$

On the other hand, for $0 \leq i \leq k$, by telescoping and using (7.51) and that $\gamma \in (1/4, 1/2]$, we find

$$(7.59) \quad \int_{B_{\gamma^{-i}r}} |u - \ell_r| dx \leq c \sum_{m=0}^i \int_{B_{\gamma^{-m}r}} |u - \ell_{\gamma^{-m}r}| dx + c \gamma^{-i} r \sum_{m=0}^i \int_{B_{\gamma^{-m}r}} |Du - (Du)_{\gamma^{-m}r}| dx,$$

with $c = c(n, N)$. Merging the content of (7.58)–(7.59), and using Fubini's theorem on discrete sums, we obtain

$$(7.60) \quad \begin{aligned} T_2 &\leq c \sum_{i=0}^k \gamma^{i2s} \sum_{m=0}^i \int_{B_{\gamma^{-m}r}} |u - \ell_{\gamma^{-m}r}| dx + c \sum_{i=0}^k \gamma^{i(2s-1)} r \sum_{m=0}^i \int_{B_{\gamma^{-m}r}} |Du - (Du)_{\gamma^{-m}r}| dx \\ &\leq c \sum_{m=0}^k \int_{B_{\gamma^{-m}r}} |u - \ell_{\gamma^{-m}r}| dx \sum_{i=m}^\infty \gamma^{i2s} + c r \sum_{m=0}^k \int_{B_{\gamma^{-m}r}} |Du - (Du)_{\gamma^{-m}r}| dx \sum_{i=m}^\infty \gamma^{i(2s-1)} \\ &\leq c \sum_{m=0}^k \gamma^{m2s} \int_{B_{\gamma^{-m}r}} |u - \ell_{\gamma^{-m}r}| dx + c r \sum_{m=0}^k \gamma^{m(2s-1)} \int_{B_{\gamma^{-m}r}} |Du - (Du)_{\gamma^{-m}r}| dx \end{aligned}$$

with $c = c(n, N, s)$, since $\gamma \in (1/4, 1/2]$ and $2s > 1$, and thus we can majorize the geometric sums. Now we estimate as after (7.55): for $m = 0, \dots, k-1$, if $\mu \in [\gamma^{-m}r, \gamma^{-(m+1)}r]$ we again have $G(\gamma^{-m}r) \leq cG(\mu) + c\mu H(\mu)$ and $H(\gamma^{-m}r) \leq cH(\mu)$. Since $\gamma^{2sm} \approx (r/\mu)^{2s}$, $r\gamma^{(2s-1)m} \approx r(r/\mu)^{2s-1}$ for $\mu \in [\gamma^{-m}r, \gamma^{-(m+1)}r]$, we obtain

$$(7.61) \quad \sum_{m=0}^k \gamma^{m2s} G(\gamma^{-m}r) \leq c \int_r^R \left(\frac{r}{\mu}\right)^{2s} G(\mu) \frac{d\mu}{\mu} + cr \int_r^R \left(\frac{r}{\mu}\right)^{2s-1} H(\mu) \frac{d\mu}{\mu} + c \left(\frac{r}{R}\right)^{2s} G(R),$$

and

$$(7.62) \quad \sum_{m=0}^k \gamma^{m(2s-1)} H(\gamma^{-m}r) \leq c \int_r^R \left(\frac{r}{\mu}\right)^{2s-1} H(\mu) \frac{d\mu}{\mu} + c \left(\frac{r}{R}\right)^{2s-1} H(R).$$

Inserting (7.61)–(7.62) into (7.60), we obtain

$$\begin{aligned} T_2 \leq c \int_r^R \left(\frac{r}{\mu}\right)^{2s} \int_{B_\mu} |u - \ell_\mu| dx \frac{d\mu}{\mu} + c \left(\frac{r}{R}\right)^{2s} \int_{B_R} |u - \ell_R| dx \\ + cr \int_r^R \left(\frac{r}{\mu}\right)^{2s-1} \int_{B_\mu} |Du - (Du)_\mu| dx \frac{d\mu}{\mu} + cR \left(\frac{r}{R}\right)^{2s} \int_{B_R} |Du - (Du)_R| dx, \end{aligned}$$

where $c = c(n, N, s)$. Merging the estimates for T_1 and T_2 proves (7.52) when $r \leq R/4$.

For $R/4 \leq r \leq R$, using (7.51) we simply have

$$\begin{aligned} T_1 &\leq c \int_{B_R} |u - \ell_R| dx + cr^{2s} \int_{B_R} |Du - (Du)_R| dx \int_{\mathbb{R}^n \setminus B_R} \frac{dx}{|x|^{n+2s-1}} \\ &\leq c \left(\frac{r}{R}\right)^{2s} \int_{B_R} |u - \ell_R| dx + c \left(\frac{r}{R}\right)^{2s} R \int_{B_R} |Du - (Du)_R| dx, \end{aligned}$$

and

$$\begin{aligned} T_2 &\leq cr^{-n} \int_{B_R \setminus B_r} |u - \ell_r| dx \leq c \int_{B_R} |u - \ell_R| dx + c \int_{B_R} |\ell_R - \ell_r| dx \\ &\leq c \left(\frac{r}{R}\right)^{2s} \int_{B_R} |u - \ell_R| dx + c \left(\frac{r}{R}\right)^{2s} R \int_{B_R} |Du - (Du)_R| dx. \end{aligned}$$

□

Now we move on to the proof of Theorem 2.8.

Proof of Theorem 2.8. Theorem 2.7 already covers the case $s \leq 1/2$; we may therefore assume throughout the proof that $s > 1/2$. The argument is a bootstrap on the differentiability exponent: Theorem 2.7 indeed provides the initial regularity $Du \in C^{0,\beta}$ for every $\beta \leq 2(1-s)$. The special structure of the nonlocal term then allows, at each step, to use the already known Hölder regularity of Du to improve the decay of the nonlocal tail by a further amount $2(1-s)$. This is due to the cancellation of affine functions in the frozen nonlocal operator, which allows the tail to be estimated in terms of the oscillation of Du , rather than of u itself. All the estimates below are performed on the regular set Ω_u obtained in Theorem 2.7; the bootstrap does not change this set, but only improves the Hölder exponent of Du on compact subsets of Ω_u .

We prove the theorem by induction; more precisely, we show that for every $k \in \mathbb{N}_0$,

$$(7.63) \quad \text{Theorem 2.8 holds for every } \beta \in (0, \min\{2(k+1)(1-s), 1\}).$$

The statement for $k = 0$ follows from Theorem 2.7, since it covers the range $0 < \beta \leq 2(1-s)$. Assume that (7.63) holds for some $k_0 \in \mathbb{N}_0$ (such that $2(k_0+1)(1-s) < 1$); let us prove it for $k_0 + 1$. Let $0 < \beta < \min\{2(k_0+2)(1-s), 1\}$; if $\beta < 2(k_0+1)(1-s)$, then the conclusion is trivial by the inductive hypothesis; thus we may assume that

$$2(k_0+1)(1-s) \leq \beta < \min\{2(k_0+2)(1-s), 1\}.$$

Choose $\beta_0 < 2(k_0+1)(1-s)$ so close to $2(k_0+1)(1-s)$ that $\beta < \beta_0 + 2(1-s)$; we also choose $\beta_0 \neq 2s-1$. Since the assumption $f \in \mathcal{M}_{\text{loc}}^{n/(1-\beta)}$ implies $f \in \mathcal{M}_{\text{loc}}^{n/(1-\beta_0)}$, we obtain $Du \in C_{\text{loc}}^{0,\beta_0}(\Omega_u)$ and in particular, for every $x_0 \in \Omega_u$, after setting $R_{x_0} = \bar{r}_{x_0}/32$,

$$(7.64) \quad [Du]_{C^{0,\beta_0}(B_{4R_{x_0}}(x_0))} \leq C_{k_0}.$$

Let now $\bar{x} \in B_{R_{x_0}/4}(x_0)$ and $0 < r \leq R_{x_0}/32$ be fixed, and also let $\ell_{\bar{x},2r} = \ell_{2r}$ be the affine function defined in (7.50).

We go back to estimate (7.8) and we improve the estimate for the term $(III) = \langle Q_a^{\text{nl}}u, u - h \rangle_{\text{nl}}$; by using (7.48), the linearity of Q_a^{nl} , (2.27)–(2.28) and (7.6), we get

$$\begin{aligned}
 \langle Q_a^{\text{nl}}u, u - h \rangle_{\text{nl}} &= \langle Q_a^{\text{nl}}(u - \ell_{2r}), u - h \rangle_{\text{nl}} \\
 &= \int_{B_{2r}(\bar{x})} \int_{B_{2r}(\bar{x})} a(y-z) [(u - \ell_{2r})(y) - (u - \ell_{2r})(z)] \cdot [(u - h)(y) - (u - h)(z)] \frac{dy dz}{|y - z|^{n+2s}} \\
 &\quad - 2 \int_{B_r(\bar{x})} \int_{\mathbb{R}^n \setminus B_{2r}(\bar{x})} a(y-z) [(u - \ell_{2r})(y) - (u - \ell_{2r})(z)] \cdot (u - h)(z) \frac{dy dz}{|y - z|^{n+2s}} \\
 (7.65) \quad &=: (IV)_1 + (IV)_2.
 \end{aligned}$$

Using (2.27), (3.6) and (7.6), we find

$$\begin{aligned}
 |(IV)_1| &\leq c [u - \ell_{2r}]_{W^{s,2}(B_{2r}(\bar{x}))} [u - h]_{W^{s,2}(B_{2r}(\bar{x}))} \\
 &= c r^{2(1-s)} \left(\int_{B_{2r}(\bar{x})} |Du - (Du)_{B_{2r}(\bar{x})}|^2 dx \right)^{1/2} \left(\int_{B_{2r}(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \\
 &\leq c C_{k_0} r^{n/2+\beta_0+2(1-s)} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2},
 \end{aligned}$$

with $c = c(n, N, s, \max\{\nu, 1\})$, where in the last inequality we used (7.6) and (7.64). Then, by (2.27), (3.16), (7.6), (7.13), Hölder's and Poincaré's inequalities (which are applicable as $(\ell_{2r})_{B_{2r}(\bar{x})} = (u)_{B_{2r}(\bar{x})}$) we get

$$\begin{aligned}
 |(IV)_2| &\leq c \int_{B_r(\bar{x})} |(u - h)(z)| \left\{ \int_{\mathbb{R}^n \setminus B_{2r}(\bar{x})} (|u(y) - \ell_{2r}(y)| + |u(z) - \ell_{2r}(z)|) \frac{dy}{|y - \bar{x}|^{n+2s}} \right\} dz \\
 &= c r^{-2s} \text{Tail}(u - \ell_{2r}; B_{2r}(\bar{x})) \int_{B_r(\bar{x})} |(u - h)(z)| dx + c r^{-2s} \int_{B_r(\bar{x})} |u - h| |u - \ell_{2r}| dx \\
 &\leq c r^{n/2-2s} \text{Tail}(u - \ell_{2r}; B_{2r}(\bar{x})) \left(\int_{B_r(\bar{x})} |u - h|^2 dx \right)^{1/2} \\
 &\quad + c r^{-2s} \left(\int_{B_r(\bar{x})} |u - h|^2 dx \right)^{1/2} \left(\int_{B_{2r}(\bar{x})} |u - \ell_{2r}|^2 dx \right)^{1/2} \\
 &\leq c r^{n/2+1-2s} \text{Tail}(u - \ell_{2r}; B_{2r}(\bar{x})) \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \\
 &\quad + c r^{n/2+2(1-s)} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \left(\int_{B_{2r}(\bar{x})} |Du - (Du)_{B_{2r}(\bar{x})}|^2 dx \right)^{1/2} \\
 &\leq c r^{n/2+1-2s} \text{Tail}(u - \ell_{2r}; B_{2r}(\bar{x})) \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \\
 &\quad + c C_{k_0} r^{n/2+\beta_0+2(1-s)} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2},
 \end{aligned}$$

with $c = c(\text{data}, \max\{\nu, 1\})$, where in the last inequality we used (7.64). Then, by using the tail property (7.53) and (7.64), we get

$$\begin{aligned}
 &r^{-2s} \text{Tail}(u - \ell_{2r}; B_{2r}(\bar{x})) \\
 &\leq c R_{x_0}^{-2s} \text{Tail}(u - \ell_{R_{x_0}}; B_{R_{x_0}}(\bar{x})) + c \int_r^{R_{x_0}} \mu^{1-2s} \int_{B_\mu(\bar{x})} |Du - (Du)_{B_\mu(\bar{x})}| dx \frac{d\mu}{\mu} \\
 &\quad + c R_{x_0}^{1-2s} \int_{B_{R_{x_0}}(\bar{x})} |Du - (Du)_{B_{R_{x_0}}(\bar{x})}| dx \\
 &\leq c R_{x_0}^{-2s} \text{Tail}(u - \ell_{R_{x_0}}; B_{R_{x_0}}(\bar{x})) + c C_{k_0} \int_r^{R_{x_0}} \mu^{1-2s+\beta_0} \frac{d\mu}{\mu} + c C_{k_0} R_{x_0}^{1-2s+\beta_0}
 \end{aligned}$$

$$\leq c R_{x_0}^{-2s} \mathbf{Tail}(u - \ell_{R_{x_0}}; B_{R_{x_0}}(\bar{x})) + c C_{k_0} (R_{x_0}^{1-2s+\beta_0} + r^{1-2s+\beta_0}),$$

where $c = c(n, N, s)$. Moreover, by (3.19) and since $s > 1/2$, we estimate

$$\begin{aligned} & R_{x_0}^{-2s} \mathbf{Tail}(u - \ell_{R_{x_0}}; B_{R_{x_0}}(\bar{x})) \\ & \leq c R_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{R_{x_0}}(\bar{x})}; B_{R_{x_0}}(\bar{x})) + c |(Du)_{B_{R_{x_0}}(\bar{x})}| \int_{\mathbb{R}^n \setminus B_{R_{x_0}}(\bar{x})} \frac{dx}{|x - \bar{x}|^{n+2s-1}} \\ (7.66) \quad & \leq c R_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{4R_{x_0}}(x_0)}; B_{4R_{x_0}}(x_0)) + c R_{x_0}^{1-2s} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \end{aligned}$$

for all $\bar{x} \in B_{R_{x_0}}(x_0)$, where $c = c(n, N, s)$. We merge the content of (7.65)–(7.66), (7.33), and (7.64), obtaining

$$\begin{aligned} \|\langle Q_a^{\text{nl}} u, u - h \rangle\| & \leq c r^{n/2+1} \left\{ R_{x_0}^{-2s} \mathbf{Tail}(u - (u)_{B_{4R_{x_0}}(x_0)}; B_{4R_{x_0}}(x_0)) + R_{x_0}^{1-2s} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} \right. \\ & \quad \left. + c C_{k_0} (R_{x_0}^{1-2s+\beta_0} + r^{1-2s+\beta_0}) \right\} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \\ (7.67) \quad & \leq \tilde{C}_{k_0} r^{n/2+\min\{1, \beta_0+2(1-s)\}} \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \end{aligned}$$

for all $\bar{x} \in B_{R_{x_0}}(x_0)$, and $0 < r \leq R_{x_0}/32$, with the obvious definition for $\tilde{C}_{k_0}$. We now combine (7.8)–(7.10), (7.67), (7.44) and performing the same algebraic manipulations we obtain the improved comparison estimate

$$\begin{aligned} & \left(\int_{B_r(\bar{x})} |Du - Dh|^2 dx \right)^{1/2} \\ & \leq c r^{n/2} \left\{ \|A(\cdot, u(\cdot)) - \bar{A}\|_{L^\infty(B_r(\bar{x}))} + \tilde{C}_{k_0} r^{\min\{1, \beta_0+2(1-s)\}} \right\} \left(\int_{B_{2r}(\bar{x})} |Du|^2 dx \right)^{1/2} \\ & \quad + c r^{n/2+\beta} \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))} \\ & \leq c r^{n/2} \left\{ (\|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))}^\beta + 1) r^\beta + \tilde{C}_{k_0} r^{\min\{1, \beta_0+2(1-s)\}} \right\} \|Du\|_{L^\infty(B_{R_{x_0}}(x_0))} \\ & \quad + c r^{n/2+\beta} \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{\varrho_{x_0}}(x_0))} \\ (7.68) \quad & \leq \hat{C}_{k_0} r^{n/2+\beta} \end{aligned}$$

for all $0 < r \leq R_{x_0}/32$, where in the last line we used (7.32), (2.23), and the choice $\beta < \min\{1, \beta_0 + 2(1-s)\}$, with the obvious definition for $\hat{C}_{k_0}$. Equation (7.68) together with (7.16) gives

$$(7.69) \quad \left(\int_{B_\varrho(\bar{x})} |Du - (Du)_{B_\varrho(\bar{x})}|^2 dx \right)^{1/2} \leq c \left(\frac{\varrho}{r} \right)^{n/2+1} \left(\int_{B_r(\bar{x})} |Du - (Du)_{B_r(\bar{x})}|^2 dx \right)^{1/2} + \hat{C}_{k_0} r^{n/2+\beta};$$

this holds for all $\bar{x} \in B_{R_{x_0}}(x_0)$ and $0 < \varrho \leq r \leq R_{x_0}/32$. Equation (7.69) together with the standard iteration argument already used after (7.47) imply

$$E(Du; B_r(\bar{x})) \leq c \left[\frac{1}{R_{x_0}^\beta} E(Du; B_{R_{x_0}}(\bar{x})) + \hat{C}_{k_0} \right] r^\beta \leq c \left[\frac{1}{R_{x_0}^\beta} \|Du\|_{L^\infty(B_{\bar{r}_{x_0}}(x_0))} + \hat{C}_{k_0} \right] r^\beta$$

for all $r \leq R_{x_0}$ and $\bar{x} \in B_{R_{x_0}}(x_0)$, with $c = c(\text{data}, \max\{\nu, 1\}, \beta)$. This estimate and Campanato's characterization of Hölder continuity prove that $Du \in C_{\text{loc}}^{\beta, 0}(B_{R_{x_0}}(x_0))$ with quantitative estimate (2.30), and this completes the inductive step, and hence the proof. $\square$

7.5. Hierarchy of radii and admissible balls. For the reader's convenience, we summarize the hierarchy of radii used throughout the preceding proofs: several local scales have been introduced at different stages of the argument, each with a specific role. The following recap wants to collect their mutual relations and the corresponding ball inclusions, so that the localization arguments above can be read without ambiguity. For an exhaustion of compact sets $K_i \Subset \Omega$ as defined at the beginning of the Chapter, we define the threshold

$$\bar{r}_\beta(K_i) := \min \left\{ r_0, r_1, r_f(K_i), r_A(K_i), \frac{1}{4} \text{dist}(K_i, \partial\Omega) \right\} :$$

this is the maximal scale on the compact set K_i at which the smallness assumptions on the coefficients and on the datum are valid, and at which the structural restrictions involving r_0 and r_1 can be used. If $x_0 \in \Omega_{u,\beta}$, then, by definition, there are $i = i(x_0) \in \mathbb{N}$ and $0 < \varrho_{x_0} < \bar{r}_\beta(K_i)$ such that

$$x_0 \in K_i, \quad B_{\varrho_{x_0}}(x_0) \Subset B_{2\varrho_{x_0}}(x_0) \Subset \Omega, \quad \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x_0)) < \frac{\varepsilon_s}{2}.$$

The radius ϱ_{x_0} is the initial scale at which Proposition 5.6 is applied. Next, by the continuity of $x \mapsto \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(x))$, we choose a radius $0 < r_{x_0} \leq \varrho_{x_0}/8$ so small that

$$B_{r_{x_0}}(x_0) \Subset \text{int}(K_i), \quad B_{\varrho_{x_0}}(\bar{x}) \Subset \Omega \quad \text{and} \quad \mathcal{E}_\gamma(u; B_{\varrho_{x_0}}(\bar{x})) < \frac{\varepsilon_s}{2}$$

for every $\bar{x} \in B_{r_{x_0}}(x_0)$: this is the localization radius, as it allows us to apply Proposition 5.6 uniformly with center $\bar{x} \in B_{r_{x_0}}(x_0)$ and initial scale ϱ_{x_0} . In particular, for such centers, $B_{\varrho_{x_0}}(\bar{x}) \subset B_{2\varrho_{x_0}}(x_0) \Subset \Omega$.

In the gradient estimates we shall further reduce the working scale. We choose $0 < \bar{r}_{x_0} \leq r_{x_0}/4$, and this radius is the scale on which the Lipschitz estimate is obtained. It is chosen small enough to satisfy the smallness conditions required in the iteration/summation procedures, and it guarantees the inclusions $B_r(\bar{x}) \subset B_{2\bar{r}_{x_0}}(x_0) \subset B_{r_{x_0}}(x_0)$ if $\bar{x} \in B_{\bar{r}_{x_0}}(x_0)$ and $0 < r \leq \bar{r}_{x_0}$, and $B_{2r}(\bar{x}) \subset B_{\bar{r}_{x_0}}(x_0)$ if $0 < r \leq \bar{r}_{x_0}/4$ and $\bar{x} \in B_{\bar{r}_{x_0}/2}(x_0)$. Finally, in the bootstrap arguments for higher gradient regularity Theorem 2.8, we have to reduce the radius once more and choose $0 < R_{x_0} \leq \bar{r}_{x_0}$. This is the scale on which the already established C^{1,β_0} estimate is used as an input for the next step of the bootstrap. If $\bar{x} \in B_{R_{x_0}/4}(x_0)$ and $0 < r \leq R_{x_0}/8$, then $B_{2r}(\bar{x}) \subset B_{R_{x_0}/2}(x_0) \subset B_{R_{x_0}}(x_0) \subset B_{\bar{r}_{x_0}}(x_0)$. Thus the complete hierarchy of radii is

$$0 < R_{x_0} = \frac{\bar{r}_{x_0}}{32} \leq \bar{r}_{x_0} \leq \frac{r_{x_0}}{4} \leq \frac{\varrho_{x_0}}{32} < \varrho_{x_0} < \bar{r}_\beta(K_i),$$

and all balls appearing in the local and nonlocal estimates are contained in the region where the previously established estimates are available.

8. FULL REGULARITY FOR u -INDEPENDENT LOCAL OPERATORS

Suppose that $A(x, u) \equiv A(x)$, where $A(\cdot)$ satisfies the assumptions of the main theorems. In this case, the modulus of continuity in the u -variable (2.4) is trivial, that is, $\omega_u(\cdot) \equiv 0$. Returning to the approximate harmonicity result in Proposition 5.3, we note that the point x_0 and the radius ϱ no longer need to satisfy the smallness assumption (5.1): this is because (5.1) only enters through condition (5.40), which in the current setting is automatically satisfied for any choice of ε_0 . Consequently, for the remainder of the proofs, we may take $\Omega_u \equiv \Omega$.

Let us now track dependence of the radii. First, to verify the assumptions of Theorems 2.1 to 2.4 and 6.1, the radii have to satisfy the requirements of Paragraph 6.1, except for the smallness condition (6.5) and its uniform version (6.6). More precisely, if $x_0 \in \Omega$, we may set $d = \text{dist}(x_0, \partial\Omega)$ and choose

$$\varrho_{x_0} = \min \left\{ r_0, r_1, r_f(\overline{B_{d/2}(x_0)}), r_A(\overline{B_{d/2}(x_0)}), \frac{1}{4}d, \frac{1}{2} \right\};$$

we then set $r_{x_0} := \varrho_{x_0}/4$. With this choice, for every $\bar{x} \in B_{r_{x_0}}(x_0)$, we have $B_{\varrho_{x_0}}(\bar{x}) \subset B_{2\varrho_{x_0}}(x_0) \Subset \Omega$. Thus the balls $B_{\varrho_{x_0}}(x)$, with $x \in B_{r_{x_0}}(x_0)$, are all admissible, compare with Paragraph 7.5. In this case the choice of ϱ_{x_0} and r_{x_0} depends only on $\text{dist}(x_0, \partial\Omega)$, the structural radii r_0, r_1 , and the local radii $r_f(\cdot), r_A(\cdot)$, but not on u . Additionally, in the case of Theorems 2.4 and 6.1, we can actually quantify $r_f(\cdot)$, as by (6.15) and (3.5). Indeed, r_f has to satisfy (2.8), and in our case a sufficient condition can be obtained from the maximal quantity appearing in (6.15) in the case of Theorem 6.1: we have, since $B_{r_{x_0}}(x_0) \subset B_{d/2}(x_0)$ and $\varrho_{x_0} \leq r_f(\overline{B_{d/2}(x_0)}) = r_f$, if r_f is smaller than $d/4$,

$$\sup_{\substack{\bar{x} \in B_{r_{x_0}}(x_0) \\ 0 < t \leq \varrho_{x_0}}} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq \sup_{\substack{\bar{x} \in B_{d/2}(x_0) \\ 0 < t \leq r_f}} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq r_f^{\beta_0} \sup_{\substack{\bar{x} \in B_{d/2}(x_0) \\ 0 < t \leq d/4}} t^{2-\beta_0} \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi}$$

and the latter quantity is smaller than $\frac{1}{2}\varepsilon_s(\mathbf{data}, \beta_0)$ – and thus (6.7) is ensured – as soon as r_f is small enough, with the dependences stated in Theorem 2.9. Similarly for Theorem 2.4,

$$\sup_{\substack{\bar{x} \in B_{r_{x_0}}(x_0) \\ 0 < t \leq \varrho_{x_0}}} t^2 \left(\int_{B_t(\bar{x})} |f|^\chi dx \right)^{1/\chi} \leq c r_f^{\beta_0} \|f\|_{\mathcal{M}^{n/(2-\beta_0)}(B_{d/2}(x_0))}.$$

For Theorems 2.5–2.8, the coefficient radius r_A can also be quantified. Here the local *BMO*-smallness condition (5.2) is no longer an additional restriction: it follows directly from the global modulus of continuity (2.18). More precisely, if $0 < r_A \leq d/4$, then

$$\sup_{x \in B_{d/2}(x_0)} E_{r_A}(A; x) \leq \Lambda \omega_x(r_A) < \delta_0(\mathbf{data}, \beta_0(s)),$$

where $\delta_0(\mathbf{data}, \beta_0(s))$ comes from Proposition 5.6 and it corresponds to the choice of $\beta_0(s)$ in (7.1), provided that $r_A \leq d/4$ is sufficiently small; in particular, r_A depends only on $\text{dist}(x_0, \partial\Omega)$, $\mathbf{data}$ and $\omega_x(\cdot)$. Finally, we quantify r_f in the settings of Theorems 2.5 and 2.7; again taking $r_f \leq d/4$, all balls $B_\varrho(x)$ with $x \in B_{d/2}(x_0)$ and $0 < \varrho \leq r_f$ are contained in $B_{3d/4}(x_0) \Subset \Omega$; thus

$$\sup_{\substack{x \in B_{d/2}(x_0) \\ 0 < t \leq r_f}} t^2 \left(\int_{B_t(x)} |f|^\chi dx \right)^{1/\chi} \leq c r_f \|\mathbf{I}_{1,\chi}^f(\cdot, d/4)\|_{L^\infty(B_{d/2}(x_0))}$$

by (3.10), for Theorem 2.5. In order to have the latter quantity smaller than $\varepsilon_b(\mathbf{data}, \beta_0(s))$ and, in this way, have the datum smallness required in the preliminary construction (2.8) satisfied (this is employed after (7.2)), it is sufficient to choose r_f smaller than a constant depending on d , $\mathbf{data}$ and a bound on $\|\mathbf{I}_{1,\chi}^f(\cdot, d/4)\|_{L^\infty(B_{d/2}(x_0))}$. In the setting of Theorems 2.7–2.8, by (3.5), we have

$$\sup_{\substack{x \in B_{d/2}(x_0) \\ 0 < t \leq r_f}} t^2 \left(\int_{B_t(x)} |f|^\chi dx \right)^{1/\chi} \leq c(\beta) r_f^{1+\beta} \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{3d/4}(x_0))}.$$

Thus it is enough to choose $r_f \leq d/4$ so small, depending on d , $\mathbf{data}$, β and any bound on $\|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{3d/4}(x_0))}$, such that

$$c(\beta) r_f^{1+\beta} \|f\|_{\mathcal{M}^{n/(1-\beta)}(B_{3d/4}(x_0))} < \varepsilon_b(\mathbf{data}, \beta_0(s)),$$

for the same reason just explained. Finally, the radius $\bar{r} = \bar{r}_{x_0}$ has to satisfy the conditions (7.24) and (7.30); recalling that $\omega_u(\cdot) \equiv 0$, it is then easy to check that $\bar{r}_{x_0}$ will depend only on the quantities stated in Theorem 2.9; the very same reasoning applies to the radius $R_{x_0} = \bar{r}_{x_0}/8$ appearing in Theorem 2.8.

Acknowledgments. The authors thank L. Brasco for several helpful suggestions on a preliminary version of the paper, which helped improve and clarify the presentation.

Part of this work was carried out while C.A. Antonini held a postdoctoral fellowship of the National Institute for Advanced Mathematics (INdAM) at the University of Florence, and previously a postdoctoral position at the University of Parma. The latter position was funded by the project “Bando di Ateneo per la Ricerca 2024 – Azione D”, awarded to P. Baroni. C.A. Antonini is also supported by the Italian Ministry of University and Research (MUR) through the FIS 2 project “*SiGmA - Singularities in Geometric Analysis: Minimal Surfaces and Mean Curvature Flows*”, project code FIS-2023-02962, CUP G53C25000120001.

P. Baroni is supported by the European Research Council through the ERC Starting Grant project NEW, no. 101220121, and by the European Union – NextGenerationEU through the PRIN 2022 PNRR project 2022E9CF89 “Geometric Evolution Problems and Shape Optimization” (GEPSO), PNRR Italia Domani.

Data availability statement. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest. The authors declare that they have no conflict of interest.

REFERENCES

- [1] N. ABATANGELO, E. AFFILI, M. COZZI: Optimal boundary regularity for mixed local and nonlocal equations, *Math. Eng.* **8**, (2026) n. 1, 1–42.
- [2] E. ACERBI, G. MINGIONE: Gradient estimates for a class of parabolic systems. *Duke Math. J.* **136** (2007), n. 2, 285–320.
- [3] C.A. ANTONINI: Local and global $C^{1,\beta}$ -regularity for uniformly elliptic quasilinear operators of p -Laplace and Orlicz-Laplace type. <https://arxiv.org/abs/2601.07140> (2026).
- [4] C.A. ANTONINI, M. COZZI: Global gradient regularity and a Hopf lemma for quasilinear operators of mixed local-nonlocal type. *J. Differ. Equations* **425**, 342–382 (2025).
- [5] J.P.A. APAZA: Local Hölder regularity for quasilinear elliptic equations with mixed local-nonlocal operators, variable exponents, and weights. *J. Math. Anal. Appl.* **558** (2026), n. 2, Paper n. 130417.
- [6] A. KH. BALCI, L. BEHN, L. DIENING, J. STORN: Examples of p -harmonic maps. *SIAM Journal on Mathematical Analysis* **58** (2026), n. 1, 260–275.
- [7] P. BARONI: A new condition ensuring gradient continuity for minimizers of non-autonomous functionals with mild phase transition. *Calc. Var. Partial Differential Equations* **64** (2025), Paper n. 112.
- [8] P. BARONI, V. BÖGELEIN: Calderón-Zygmund estimates for parabolic $p(x,t)$ -Laplacian systems. *Rev. Mat. Iberoamericana* **30** (2014), n. 4, 1355–1386.
- [9] S. BIAGI, S. DIPIERRO, E. VALDINOCI, E. VECCHI: Mixed local and nonlocal elliptic operators: regularity and maximum principles. *Comm. Partial Differential Equations* **47** (2022), n. 3, 585–629.
- [10] S. BIAGI, S. DIPIERRO, E. VALDINOCI, E. VECCHI: A Faber-Krahn inequality for mixed local and nonlocal operators. *J. Anal. Math.*, **150** (2023), n. 2, 405–448.
- [11] A. BISWAS, M. MODASIYA: Mixed local-nonlocal operators: maximum principles, eigenvalue problems and their applications. *J. Anal. Math.* **156** (2025), n. 1, 47–81.
- [12] A. BISWAS, H. PRASAD: Lipschitz potential estimates for diffusion with jumps. *NoDEA, Nonlinear Differ. Equ. Appl.* **32** (2025), n. 5, Paper n. 88.
- [13] A. BISWAS, E. TOPP: Interior $C^{1,\alpha}$ regularity of mixed local-nonlocal (p, q) -energy minimizers for $p \leq sq$. <https://arxiv.org/abs/2512.07481> (2025).
- [14] L. BRASCO: *Handbook of Calculus of Variations for Absolute Beginners*. Unitext, Vol. 163, La Matematica per il 3+2, Springer, Cham, (2025).
- [15] S.-S. BYUN, K. DEEPAK, H.-S. LEE: Global gradient estimates for the mixed local and nonlocal problems with measurable nonlinearities. *Calc. Var. Partial Differential Equations* **63** (2024), n. 2, Paper n. 27.
- [16] Z.-Q. CHEN, P. KIM, R. SONG, Z. VONDRAČEK: Sharp Green function estimates for $\Delta + \Delta^{\alpha/2}$ in $C^{1,1}$ open sets and their applications. *Illinois J. Math.* **54** (2010), n. 3, 981–1024.
- [17] Z.-Q. CHEN, P. KIM, R. SONG, Z. VONDRAČEK: Boundary Harnack principle for $\Delta + \Delta^{\alpha/2}$. *Trans. Amer. Math. Soc.* **364** (2012), n. 8, 4169–4205.
- [18] I. CHLEBICKA, K. SONG, Y. YOUN, A. ZATORSKA-GOLDSTEIN: Riesz potential estimates for mixed local-nonlocal problems with measure data. *J. Lond. Math. Soc., II. Ser.* **112** (2025), n. 4, Paper n. e70310.
- [19] M. COLOMBO, G. MINGIONE: Calderón-Zygmund estimates and non-uniformly elliptic operators. *J. Funct. Anal.* **270** (2016), n. 4, 1416–1478.
- [20] D. CRUZ-URIBE, L. DIENING: Sharp $\mathcal{A}$ -harmonic approximation. *Appl. Anal.* **98** (2019), n. 1–2, 374–380.
- [21] K. V. DA SILVA, A. M. SALORT: A limiting problem for local/non-local p -Laplacians with concave-convex nonlinearities. *Z. Angew. Math. Phys.* **71** (2020), n. 6, Paper n. 191.
- [22] S. DAS: Gradient Hölder regularity in mixed local and nonlocal linear parabolic problem. *J. Math. Anal. Appl.* **535** (2024), n. 2, Paper n. 128140.
- [23] C. DE FILIPPIS, G. MINGIONE: Gradient regularity in mixed local and nonlocal problems. *Math. Ann.* **388** (2024), n. 1, pp. 261–328.
- [24] C. DE FILIPPIS, G. MINGIONE: Nonuniformly elliptic Schauder theory. *Invent. Math.* **234** (2023), no. 3, 1109–1196.
- [25] C. DE FILIPPIS, G. MINGIONE, S. NOWAK: Partial regularity in nonlocal systems I. *Adv. Math.*, to appear, <https://arxiv.org/abs/2501.08405> (2025).
- [26] C. DE FILIPPIS, G. MINGIONE, S. NOWAK: Partial regularity in nonlocal systems II. <https://arxiv.org/abs/2602.18848> (2026).
- [27] A. DI CASTRO, T. KUUSI, G. PALATUCCI: Local behavior of fractional p -minimizers. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **33** (2016), n. 5, 1279–1299.
- [28] E. DiBENEDETTO: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal., Theory Methods Appl.* **7** (1983), 827–850.
- [29] M. DING, MENGGAO, Y. FANG, YUZHOU, C. ZHANG: Local behavior of the mixed local and nonlocal problems with nonstandard growth. *J. Lond. Math. Soc., II. Ser.* **109** (2024), n. 6, Paper n. e12947.
- [30] F. DUZAAR, G. MINGIONE: Gradient continuity estimates. *Calc. Var. Partial Differ. Equ.* **39** (2010), n. 3–4, pp. 379–418.
- [31] P. GARAIN, J. KINNUNEN: On the regularity theory for mixed local and nonlocal quasilinear elliptic equations. *Trans. Am. Math. Soc.* **375** (2022), n. 8, 5393–5423.
- [32] P. GARAIN, E. LINDGREN: Higher Hölder regularity for mixed local and nonlocal degenerate elliptic equations. *Calc. Var. Partial Differential Equations* **62** (2023), n. 2, Paper n. 67.
- [33] M.G. GARRONI, J.L. MENALDI: *Second order elliptic integro-differential problems*. Chapman & Hall/CRC Research Notes in Mathematics, 430, 2002.

- [34] M. GIAQUINTA: *Multiple integrals in the calculus of variations and nonlinear elliptic systems*. Annals of Mathematics Studies 105, Princeton University Press, Princeton, NJ, 1983.
- [35] M. GIAQUINTA, E. GIUSTI: On the regularity of the minima of variational integrals. *Acta Math.* **148** (1982), 31–46.
- [36] D. GILBARG, N. S. TRUDINGER: *Elliptic partial differential equations of second order*. Classics in Mathematics, Springer-Verlag, Berlin, 2001.
- [37] E. GIUSTI: *Direct methods in the Calculus of Variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [38] J. KORVENPÄÄ, T. KUUSI, G. PALATUCCI: The obstacle problem for nonlinear integro-differential operators. *Calc. Var. Partial Differ. Equ.* **55** (2016), n. 3, Paper n. 63.
- [39] T. KUUSI, G. MINGIONE: Guide to nonlinear potential estimates. *Bull. Math. Sci.* **4** (2014), n. 1, 1–82.
- [40] T. KUUSI, G. MINGIONE, Y. SIRE: Nonlocal self-improving properties. *Anal. PDE* **8** (2015), n. 1, 57–114.
- [41] T. KUUSI, S. NOWAK, Y. SIRE: Gradient regularity and first-order potential estimates for a class of nonlocal equations. *Amer. J. Math.*, to appear.
- [42] G. LEONI: *A first course in fractional Sobolev spaces*. Graduate Studies in Mathematics, 229. American Mathematical Society, Providence, RI, 2023.
- [43] L. PICK, A. KUFNER, O. JOHN, S. FUČIK: *Function spaces Vol. 1*. De Gruyter Series in Nonlinear Analysis and Applications **14**, Walter de Gruyter & Co., Berlin, 2013.
- [44] E. M. STEIN: *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.
- [45] X. SU, E. VALDINOCI, Y. WEI, J. ZHANG: On some regularity properties of mixed local and nonlocal elliptic equations. *J. Differ. Equations* **416** (2025), Part 1, 576–613.

CARLO ALBERTO ANTONINI, DIPARTIMENTO DI MATEMATICA “FEDERICO ENRIQUES”, UNIVERSITÀ DEGLI STUDI DI MILANO, VIALE CESARE SALDINI 50, 20133, MILAN, ITALY, ORCID ID: 0000-0002-7663-1090
Email address: `carlo.antonini@unimi.it`, `antonini@altamatematica.it`

PAOLO BARONI, DIPARTIMENTO DI SCIENZE MATEMATICHE, FISICHE E INFORMATICHE, UNIVERSITÀ DI PARMA, PARCO AREA DELLE SCIENZE, 53/A, 43124 PARMA, ITALY, ORCID ID: 0000-0002-2683-4025
Email address: `paolo.baroni@unipr.it`