NONLOCAL SELF-IMPROVING PROPERTIES

TUOMO KUUSI, GIUSEPPE MINGIONE, AND YANNICK SIRE

Abstract. Solutions to nonlocal equations with measurable coefficients are higher differentiable.

Contents

1. Introduction

A basic and fundamental result in the theory of linear and nonlinear elliptic equations is given by the higher integrability of solutions. This falls in the realm of so called self-improving properties. The result was first pioneered by Meyers [24] and Elcrat & Meyers [13], and then extended in various directions and in several different contexts; see for instance [4, 15, 17, 20]. Modern proofs of this property in the nonlinear case rely on the so called Gehring lemma [16, 18]. In the simplest possible instance the result in question asserts that distributional $W^{1,2}(\Omega)$-solutions u to linear elliptic equations

$$-\mathrm{div}\,(A(x)Du) = f \in L^{\frac{2n}{n+2}+\delta_0}_{\mathrm{loc}}(\Omega)\,, \qquad \delta_0 > 0\,,$$

actually belong to a better Sobolev space

$$u \in W^{1,2+\delta}_{\mathrm{loc}}(\Omega)\,, \tag{1.1}$$

for some positive $\delta \leq \delta_0$. Here $\Omega \subset \mathbb{R}^n$ is an open subset and $n \geq 2$. The matrix $A(x)$ is supposed to be elliptic and with bounded and measurable entries, that is

$$\Lambda^{-1}|\xi|^2 \leq \langle A(x)\xi, \xi\rangle \qquad \text{and} \qquad |A(x)| \leq \Lambda \tag{1.2}$$

hold whenever $\xi \in \mathbb{R}^n$, $x \in \Omega$, where $\Lambda > 1$. The number $\delta > 0$ appearing in (1.1) is universal in the sense that, essentially, it does depend neither on the solution u nor the specific equation considered. It rather depends only on n, Λ, that is, on the ellipticity rate of the equation considered. The key point here is the measurability of the coefficients; when $A(\cdot)$ has more regular entries, higher regularity of solutions follows from the one available for equations with constant coefficients via perturbation. This is the reason why the result in (1.1) deeply lies at the core of regularity theory, and allows for a proof of several other regularity results; see for instance [17].

In this paper we are interested in studying self-improving properties of solutions to nonlocal problems. To outline the results in a special, yet meaningful model case, let us consider weak solutions $u \in W^{\alpha,2}(\mathbb{R}^n)$ of the following nonlocal equation:

$$\mathcal{E}_K(u,\eta) = \langle f, \eta\rangle \qquad \text{for every test function } \eta \in C_c^\infty(\mathbb{R}^n) \tag{1.3}$$

where $f \in L^{2+\delta_0}_{\rm loc}(\mathbb{R}^n)$ and

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

The measurable Kernel is assumed to satisfy the following uniform ellipticity assumptions:

$$\frac{1}{\Lambda|x-y|^{n+2\alpha}} \le K(x,y) \le \frac{\Lambda}{|x-y|^{n+2\alpha}} \tag{1.4}$$

for every $x,y \in \mathbb{R}^n$, where $\alpha \in (0,1)$ and $\Lambda \ge 1$. We recall that the fractional Sobolev space $W^{s,\gamma}$, for $\gamma \ge 1$ and $s \in (0,1)$, is given by the subspace of $L^\gamma(\mathbb{R}^n)$-functions u such that the following Gagliardo seminorm is finite (see for instance [11, 21])

$$[u]^\gamma_{s,\gamma} := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^\gamma}{|x-y|^{n+\gamma s}}\,dx\,dy\,. \tag{1.5}$$

In view of (1.1), a natural question to begin with is whether or not the inclusion

$$u \in W^{\alpha,2+\delta}_{\rm loc}(\mathbb{R}^n) \tag{1.6}$$

holds for some $\delta > 0$, possibly depending only on the ellipticity parameters of the equation and not on the solution itself. For the definition of local fractional Sobolev spaces, see Section 2. This has been answered in a very interesting and recent paper of Bass & Ren [2], who consider the function

$$\Gamma(x) := \left(\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\,dy\right)^{1/2}, \tag{1.7}$$

and prove that $\Gamma \in L^{2(1+\delta)}(\mathbb{R}^n)$ for some positive δ depending only on n, α, Λ and δ_0. Then (1.6) follows by characterisations of Bessel potential spaces [12, 26]. In this paper we provide a stronger and surprising result. Indeed, we see that *for nonlocal problems the self-improvement property extends to the differentiability scale*. This means that there exists some positive $\delta \in (0, 1-\alpha)$, depending only on n, α, Λ, such that

$$u \in W^{\alpha+\delta,2+\delta}_{\rm loc}(\mathbb{R}^n) \tag{1.8}$$

holds. This phenomenon is purely nonlocal, and has no parallel in the regularity theory of local equations, where, in order to get fractional Sobolev differentiability of Du, a similar fractional regularity must be assumed on the coefficients matrix $A(x)$, as for instance established in [22, 25].

In the classical local case, measurability is, in general, not sufficient to get any gradient differentiability. To see this already in the one dimensional case $n = 1$, it is sufficient to consider the following equation:

$$\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) = 0\,, \qquad \frac{1}{\Lambda} \le a(x) \le \Lambda\,, \tag{1.9}$$

and to note that

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

is a solution with $a(\cdot)$ being any measurable function satisfying nothing but the inequalities in (1.9). It is then easy to build similar multidimensional examples.

We remark that the differentiability gain is in fact the main information in (1.8), since a standard application of the fractional Sobolev embedding theorem gives that if $u \in W^{\alpha+\delta,2}$ for some $\delta > 0$, then (1.8) holds for some other number δ. Our results are actually covering a more general class of equations than the one in (1.3) and provide a full nonlocal analog of the classical higher integrability results valid in the local case. The precise statements are in the next section. Our results are a consequence of a new, fractional version of the Gehring lemma for fractional Sobolev functions that replaces the classical one valid in the local case.

We finally remark that, in recent times, there has been much attention to the regularity of solutions to nonlocal problems, especially in the basic case of kernels with measurable coefficients; see for instance [1, 3, 5, 7, 8, 14].

1.1. **Higher differentiability results.** A rather general statement concerning higher integrability for weak solutions to local problems involves non-homogeneous equations such as

$$-\operatorname{div}(A(x)Du) = -\operatorname{div} g + f \qquad \text{in } \Omega\,, \tag{1.10}$$

where the matrix $A(\cdot)$ has measurable coefficients and satisfies (1.2). Indeed, assuming that $g \in L^{2+\delta_0}_{\rm loc}(\Omega,\mathbb{R}^n)$ and $f \in L^{2n/(n+2)+\delta_0}_{\rm loc}(\Omega)$ hold for some $\delta_0 > 0$, it follows that there exists another positive number $\delta < \delta_0$, such that (1.1) holds. The exponent $2n/(n+2)$ is nothing but the conjugate of the Sobolev embedding exponent of $W^{1,2}$, that is $2n/(n-2)$.

A first nonlocal analog of (1.10) is given by

$$\mathcal{E}_K(u,\eta) = \mathcal{E}_K(g,\eta) + \int_{\mathbb{R}^n} f\eta\,dx \qquad \forall\, \eta \in C^\infty_{\rm c}(\mathbb{R}^n)\,, \tag{1.11}$$

considering weak solutions $u \in W^{\alpha,2}(\mathbb{R}^n)$. The assumptions are the natural counterpart of the local ones; we indeed take $g \in W^{\alpha+\delta_0,2}(\mathbb{R}^n)$ and

$$f \in L^{2_*+\delta_0}_{\rm loc}(\mathbb{R}^n) \tag{1.12}$$

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

$$2_* := \frac{2n}{n+2\alpha}\,, \qquad 2^* := \frac{2n}{n-2\alpha}\,, \qquad \frac{1}{2^*} + \frac{1}{2_*} = 1\,. \tag{1.13}$$

The terminology is motivated by the fractional version of the classical Sobolev embedding theorem, that is $W^{\alpha,2} \subset L^{2^*}$. On the other hand, we recall that the essence of the structure of equation (1.10) lies in the fact that the right hand side contains terms of all possible integer order. A full extension to the fractional case then leads us to consider right hand sides of *arbitrary fractional order*, not necessarily equal to the order of the considered nonlocal elliptic operator on the left hand side. Moreover, since higher integrability of solutions still holds when considering monotone quasilinear equations, we will also examine nonlinear integro-differential equations. Specifically, we will consider general equations of the type

$$\mathcal{E}^\varphi_K(u,\eta) = \mathcal{E}_H(g,\eta) + \int_{\mathbb{R}^n} f\eta\,dx \qquad \forall\, \eta \in C^\infty_{\rm c}(\mathbb{R}^n)\,. \tag{1.14}$$

The form $\mathcal{E}^\varphi_K(\cdot)$ is then defined by

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

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

$$|\varphi(t)| \le \Lambda|t|\,, \qquad \varphi(t)t \ge t^2, \qquad \forall\, t \in \mathbb{R} \tag{1.15}$$

making in fact $\mathcal{E}_K^{\varphi}$ a coercive form in $W^{\alpha,2}$, and thereby (1.14) an elliptic equation. While we assume (1.4) for $K(\cdot)$, the measurable kernel $H(\cdot)$ is now assumed to satisfy

$$|H(x,y)| \leq \frac{\Lambda}{|x-y|^{n+2\beta}} \tag{1.16}$$

for $\beta \in (0,1)$. In particular, β is also allowed to be larger than α. Here the function f is still assumed to satisfy (1.12) while the assumptions on g sharply match the structure in (1.14). We actually consider two different cases and the first one is when $2\beta \geq \alpha$. In this situation we assume the existence of a positive number $\delta_0 > 0$ such that

$$g \in W^{2\beta-\alpha+\delta_0,2}(\mathbb{R}^n)\,. \tag{1.17}$$

Needless to say, we also assume that $2\beta - \alpha + \delta_0 \in (0,1)$ to give (1.17) sense in terms of the seminorm (1.5); this in particular implies that $\beta < (1+\alpha)/2$. In the case $0 < 2\beta < \alpha$ we instead do not consider any differentiability on g, but only integrability:

$$g \in L^{p_0+\delta_0}(\mathbb{R}^n)\,, \qquad p_0 := \frac{2n}{n+2(\alpha-2\beta)}\,. \tag{1.18}$$

We then have the following *main result* of the paper:

Theorem 1.1. *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (1.14) under the assumptions (1.4) and (1.12)-(1.18). Then there exists a positive number $\delta \in (0, 1-\alpha)$, depending only on $n, \alpha, \Lambda, \beta, \delta_0$, but otherwise independent of the solution u and of the kernels $K(\cdot), H(\cdot)$, such that $u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$.*

Equation (1.11) is covered taking $\alpha = \beta$. The optimality of the assumptions on f and g can be checked by considering the model equation $(-\triangle)^{\alpha} u = (-\triangle)^{\beta} g + f$, and using Fourier analysis. They sharply relate to the fractional Sobolev embedding theorem. As in the case of the classical, local the Gehring lemma, explicit estimates on the exponent δ for Theorem 1.1 can be given by tracing back the dependence of the constants in the proof.

1.2. **Dual pairs (μ, U) and sketch of the proof.** In order to get (1.8) we here introduce a new approach and develop a method aimed at exploiting the *hidden cancellation properties* which are intrinsic in the definition of the nonlocal seminorm (1.5). To this aim, we introduce dual pairs of measures and functions (μ, U) in $\mathbb{R}^{2n}$, proving that a version of the Gehring lemma applies to them; see Section 1.3 below. A natural choice would be to consider the measure generated by the density $|x-y|^{-n}$, but this would not yield a finite measure. We therefore consider a perturbation of it, i.e. the measure defined by

$$\mu(A) := \int_A \frac{dx\,dy}{|x-y|^{n-2\varepsilon}}\,, \tag{1.19}$$

for suitably small $\varepsilon > 0$, whenever $A \subset \mathbb{R}^{2n}$ is a measurable subset. This is a locally finite, doubling Borel measure in $\mathbb{R}^{2n}$. Accordingly, for $x \neq y$, we introduce the function

$$U(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{\alpha+\varepsilon}}\,. \tag{1.20}$$

The main point here is that the measure μ and the function U are in duality when $u \in W^{\alpha,2}$ in the sense that for a function $u \in L^2(\mathbb{R}^n)$ we have that $U \in L^2(\mathbb{R}^{2n};\mu)$ holds iff $u \in W^{\alpha,2}(\mathbb{R}^n)$. This motivates in fact the following:

Definition 1. Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ and let $\varepsilon \in (0, \alpha/2)$. The couple (μ, U) defined in (1.19)-(1.20) is called a dual pair generated by the function u.

We then look at the higher μ-integrability for U proving that

$$U \in L^{2+\delta}_{\mathrm{loc}}(\mathbb{R}^{2n};\mu) \tag{1.21}$$

holds for some $\delta > 0$. Now, by the very definition of U, we have that (1.21) implies the higher differentiability of u, that is (1.8); see Section 6. This is the effect of the cancellations hidden in the definition of fractional norm in (1.5) we were mentioning above. In order to prove (1.21), we shall show decay estimates for the μ-measure of the level sets of U. The first step consists of deriving suitable energy estimates (i.e. Caccioppoli type inequalities) for U, see Theorem 3.1. We obtain a kind of reverse Hölder type inequality, that is

$$\left(\fint_{\mathcal{B}} U^2 d\mu\right)^{1/2} \lesssim \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q\, d\mu\right)^{1/q} + \text{“terms involving } g, f\text{”} \tag{1.22}$$

with $q < 2$, see Proposition 4.2. The estimate in (1.22) holds whenever $\mathcal{B} \equiv B \times B$ and $B \subset \mathbb{R}^n$ is a ball. Notice that if we discard from the sum above all the terms but the first one we formally obtain a reverse Hölder type inequality similar to those that hold for solutions to local problems.

Inequality (1.22) does not seem to be sufficient to proceed, since in order to prove estimates on level sets in $\mathbb{R}^{2n}$ we need information on every ball $\mathcal{B} \subset \mathbb{R}^{2n}$, not only those of diagonal type $B \times B$. To overcome such an apparently decisive lack of information, we have to introduce an extremely delicate localisation technique. Consider the level set $\{U > \lambda\}$; we use a Calderón-Zygmund type exit time argument in order to cover the level set with (almost disjoint) diagonal balls $B \times B$ and disjoint “off-diagonal” dyadic cubes $\mathcal{K}$

$$\{U > \lambda\} \subset \bigcup B \times B \cup \bigcup \mathcal{K}\,,$$

on which, for a suitably large number L, we have

$$\left(\fint_{B\times B} U^2\, d\mu\right)^{1/2} \approx \lambda \qquad \text{and} \qquad \left(\fint_{\mathcal{K}} U^2\, d\mu\right)^{1/2} \approx L\lambda\,,$$

see Sections 5.1 and 5.6. We call the cubes $\mathcal{K}$ off-diagonal, because they are “far” from the diagonal, in the sense that their distance from the diagonal is larger than their sidelength. The number L is introduced to make the decomposition along the diagonal predominant with respect to the decomposition outside the diagonal. Indeed, the exit time balls $B \times B$ will tend to be “larger” than the cubes $\mathcal{K}$, since they have been obtained via an exit time at a lower level λ, as shown by the first formula in the latest display.

Surprisingly enough, the fact that a cube $\mathcal{K}$ is off-diagonal allows us to prove that a reverse inequality of the type (1.22) also holds on $\mathcal{K}$ (see Lemma 5.3). This inequality, however, incorporates certain correction terms involving once again diagonal cubes. This introduces serious difficulties, since this time such cubes are not coming from any exit time argument, and there is no a priori control on them. Matching the resulting reverse inequalities with those in (1.22) is not an easy task and indeed requires an involved covering/combinatorial argument. See Sections 5.9 and 5.10, and in particular Lemma 5.6.

The final outcome of this lengthy procedure is an inequality on level sets of U, see Proposition 5.1, that implies the *higher integrability* of U, together with the new reverse Hölder type inequality reported in display (1.24) below. This holds for some $\delta > 0$ that does not depend on the solution u. See Theorem 6.1. We have therefore proved (1.21). We also remark that treating the complete problem of Theorem 1.1 up to the sharp interpolation range described by (1.17) requires additional ideas. As a matter of fact, the exit time arguments have to be adapted in order to realise a direct analog of the so called good-λ inequality principle: i.e.

no maximal operator is used here. In particular, we employ a simultaneous level set analysis via use of the composite quantity $\Psi(\cdot)$ in (5.1), where the number M (appearing in the definition of $\Psi(\cdot)$) is used to adapt the size of the levels at the exit time. This must eventually match with the specific form of the energy estimates available for solutions.

Finally, we would like to remark that, although we are here dealing with the case of scalar, linear growth nonlocal equations, our approach is only based on energy inequalities, and therefore can be extended to more general nonlinear operators of nonlocal type, see for example [9, 10]. This will be the object of future works.

1.3. **The fractional Gehring lemma for dual pairs.** The classical Gehring lemma does not simply deal with solutions to equations, but, more in general, with self-improving properties of reverse Hölder type inequalities. At the core of our approach lies in fact a new, fractional version of Gehring lemma valid for general fractional Sobolev functions, and not only for solutions to nonlocal equations. Here is a version of it.

Theorem 1.2 (Fractional Gehring lemma)**.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ for $\alpha \in (0,1)$. Let $\varepsilon \in (0, \alpha/2)$ and let (μ, U) be the dual pair generated by u in the sense of (1.19)-(1.20) and Definition 1. Assume that the following reverse Hölder type inequality with the tail holds for every $\sigma \in (0,1)$ and for every ball $B \subset \mathbb{R}^n$:*

$$\begin{aligned}\left(\fint_{\mathcal{B}} U^2 d\mu\right)^{1/2} &\leq \frac{c(\sigma)}{\varepsilon^{1/q-1/2}}\left(\fint_{2\mathcal{B}} U^q d\mu\right)^{1/q} \\ &\quad + \frac{\sigma}{\varepsilon^{1/q-1/2}}\sum_{k=2}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q\, d\mu\right)^{1/q}, \end{aligned} \tag{1.23}$$

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

$$\left(\fint_{\mathcal{B}} U^{2+\delta}\, d\mu\right)^{1/(2+\delta)} \leq c\sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^2\, d\mu\right)^{1/2}. \tag{1.24}$$

In the literature there are several extensions of Gehring lemma in general settings, for instance in metric spaces equipped with a doubling Borel measure, but Theorem 1.2 is completely different. Indeed, its central feature is actually that global higher integrability information is reconstructed from reverse inequalities that do not hold on every ball in $\mathbb{R}^{2n}$, but only on diagonal ones. This is a crucial loss of information that makes Theorem 1.2 hold not for any function $U \in L^2(\mathbb{R}^{2n};\mu)$, but rather only for dual pairs (μ, U). Moreover the presence of the infinite series on the right hand side of (1.23) gives to this inequality a delicate nonlocal character that adds relevant technical complications. Theorem 1.2 is a particular case of a more general result; we prefer to report this form again to make the basic ideas more transparent. A more comprehensive version including additional functions F and G on the right hand side of (1.23) can be proved as well; see Theorem 6.1 below.

The results of this paper have been announced in the preliminary research report [23].

Acknowledgments We wish to thank Vladimir Maz'ya for a useful discussion and Paolo Baroni for a careful reading of a preliminary version of the manuscript.

We also thank the referees for their extremely valuable work and they careful reading of the first draft of the paper: their comments led to an improved version.

2. Preliminaries and notation

In what follows we denote by c a general positive constant, possibly varying from line to line; special occurrences will be denoted by $c_1, c_2, \bar{c}_1, \bar{c}_2$ or the like. All such constants will always be *larger or equal than one*; moreover relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n, \Lambda, p, \alpha)$ means that c_1 depends only on n, Λ, p, α. We denote by

$$B(x_0, r) \equiv B_r(x_0) := \{x \in \mathbb{R}^n \, : \, |x - x_0| < r\}$$

the open ball with center x_0 and radius $r > 0$; when not important, or clear from the context, we shall omit denoting the center as follows: $B_r \equiv B(x_0, r)$; moreover, with B being a generic ball with radius r we will denote by σB the ball concentric to B having radius σr, $\sigma > 0$. Unless otherwise stated, different balls in the same context will have the same center. With $\mathcal{O} \subset \mathbb{R}^k$ being a measurable set with positive μ-measure and with h being a measurable map we shall denote by

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

its integral average. In the following we shall need to consider integrals and functions in $\mathbb{R}^n \times \mathbb{R}^n$. In this respect, instead of dealing with the usual balls in $\mathbb{R}^{2n}$, we prefer to deal with balls generated by a different metric, that is that relative to the norm (in $\mathbb{R}^{2n}$) defined by

$$\|(x_0, y_0)\| := \max\{|x_0|, |y_0|\}, \tag{2.1}$$

where $|\cdot|$ denotes the standard Euclidean norm in $\mathbb{R}^n$ and $x_0, y_0 \in \mathbb{R}^n$. These balls are denoted by $\mathcal{B}(x_0, y_0, \varrho)$, and are of course of the form

$$\mathcal{B}(x_0, y_0, \varrho) := B(x_0, \varrho) \times B(y_0, \varrho) \, .$$

In the case $x_0 = y_0$ we shall also use the shorter notation $\mathcal{B}(x_0, x_0, \varrho) \equiv \mathcal{B}(x_0, \varrho)$. With obvious meaning, these will be called diagonal balls. Moreover, with $\mathcal{B}(x_0, \varrho)$ being a fixed ball, we shall also denote $\mathcal{B} \equiv \mathcal{B}(x_0, x_0, \varrho)$ when no ambiguity shall arise, and $s\mathcal{B} := \mathcal{B}(x_0, s\varrho)$ for $s > 0$. Needless to say, since they are metric balls, and actually equivalent to the standard ones in $\mathbb{R}^{2n}$, we can apply to them several tools that are available for the usual balls. For instance, we shall later on apply the classical Vitali's covering lemma. Needless to say, it follows that $B_{\mathbb{R}^{2n}}((x_0, y_0), \varrho) = \{(x_0, y_0) \in \mathbb{R}^{2n} : |(x_0, y_0)| < \varrho\} \subset \mathcal{B}(x_0, y_0, \varrho)$. Accordingly, we shall denote

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

If A is a finite set, the symbol $\#A$ denotes the number of its elements. In the following we shall very often use the following elementary inequality

$$2^{2\beta k} \sum_{j=k-1}^{\infty} 2^{-2\beta j} \le \frac{8}{\beta} \, , \qquad \text{for } \beta \in (0, 1] \text{ and } k \ge 1 \, . \tag{2.2}$$

Finally, the local fractional Sobolev spaces are defined via the Gagliardo seminorm

$$[u]_{s,\gamma}(\Omega) := \left(\int_{\Omega} \int_{\Omega} \frac{|u(x) - u(y)|^\gamma}{|x - y|^{n+\gamma s}} \, dx \, dy \right)^{1/\gamma} \tag{2.3}$$

for $\gamma \ge 1$ and $s \in (0, 1)$. A function $u \in L^\gamma_{loc}(\mathbb{R}^n)$ belongs to $W^{s,\gamma}_{loc}(\mathbb{R}^n)$ if $[u]_{s,\gamma}(\Omega)$ is finite whenever Ω is an open bounded subset of $\mathbb{R}^n$.

The following two lemmas report some classical Poincaré-Sobolev type inequalities valid in the fractional setting; the proof of the first is exactly the one in [25], for the second we refer to [19]. See also [11, 21].

Lemma 2.1 (Fractional Poincaré inequality)**.** *Let $v \in L^p(B)$, with $B \subset \mathbb{R}^n$ being a ball of radius r, and let α be a real number such that such that $n + p\alpha \geq 0$; then the following inequality holds:*

$$\fint_B |v - (v)_B|^p \, dx \leq c r^{p\alpha} \int_B \fint_B \frac{|v(x) - v(y)|^p}{|x - y|^{n+p\alpha}} \, dx \, dy \, .$$

Note that the previous inequality in particular applies when $v \in W^{\alpha,p}(B)$, and in this case the quantity on the right hand side is finite.

Lemma 2.2 (Fractional Sobolev-Poincaré inequality)**.** *Let $v \in W^{\alpha,p}(B)$, for $\alpha \in (0,1)$, where $B \subset \mathbb{R}^n$ is a ball of radius r, or a cube of diameter r. If $p\alpha < n$, then the following inequality holds for a constant c depending only on n, α:*

$$\left(\fint_B |v - (v)_B|^{p^*} \, dx \right)^{1/p^*} \leq c r^{\alpha} \left(\int_B \fint_B \frac{|v(x) - v(y)|^p}{|x - y|^{n+p\alpha}} \, dx \, dy \right)^{1/p} ,$$

where $p^* := np/(n - p\alpha)$.

With 2_* being the exponent defined in (1.13), an immediate consequence of the previous lemma is the following inequality, that we report since it will be used several times:

$$\left(\fint_B |v - (v)_B|^2 \, dx \right)^{1/2} \leq c r^{\alpha} \left(\int_B \fint_B \frac{|v(x) - v(y)|^{2_*}}{|x - y|^{n+2_*\alpha}} \, dx \, dy \right)^{1/2_*} . \tag{2.4}$$

Moreover, if v is compactly supported in B, then $v - (v)_B$ above can be replaced by v.

3. The Caccioppoli inequality

3.1. **Preliminary reformulation of the assumptions.** We start by the assumptions made on g, that is (1.17)-(1.18). In order to give a unified proof for the two cases $2\beta \geq \alpha$ and $2\beta < \alpha$, and to simplify certain computations, we shall make a few preliminary reductions and will restate the assumptions in a more convenient way. First of all let us consider the case $2\beta \geq \alpha$, when (1.17) is in force. Let us notice that, eventually reducing the value of δ_0, and in particular taking $\delta_0 \leq \alpha/40$, (1.17) implies the existence of exponents p, γ and $\delta_1 > 0$, such that $g \in W^{\gamma(1+\delta_1),p(1+\delta_1)}(\mathbb{R}^n)$ and

$$2\beta > \gamma > 2\beta - \alpha \, , \quad 2 > p > \frac{2n}{n + 2(\gamma - 2\beta + \alpha)} \, , \qquad \delta_1 \leq \frac{\alpha}{4n} \, . \tag{3.1}$$

Indeed, let us set $\gamma = 2\beta - \alpha + \delta_0/2$ and recall that $W^{2\beta-\alpha+\delta_0,2}$ embeds in $W^{\gamma(1+\delta_1),p(1+\delta_1)}$ whenever $2\beta - \alpha + \delta_0 - n/2 = \gamma(1+\delta_1) - n/[p(1+\delta_1)]$. A lengthy computation then shows that any choice of p as above and $\delta_1 \leq 1$ satisfying the inequalities

$$\frac{(1+\delta_1)\delta_0}{[n + 2\gamma(1+\delta_1)]} < \delta_1 < \frac{(2+\delta_1)\delta_0}{[n + 2\gamma(1+\delta_1)]}$$

matches the conditions in (3.1). We now consider the case $2\beta < \alpha$, when (1.18) is in force. In this case we can instead assume the existence of numbers $p > 1$ and $\delta_1 > 0$ such that

$$g \in L^{p(1+\delta_1)}_{\mathrm{loc}}(\mathbb{R}^n) \, , \qquad p > \frac{2n}{n + 2(\alpha - 2\beta)} \, . \tag{3.2}$$

Let us now unify the previous conditions. In the case $2\beta \geq \alpha$ we clearly have that

$$\int_B \int_B \frac{|g(x) - g(y)|^{p(1+\delta_1)}}{|x - y|^{n+p(1+\delta_1)^2\gamma}} \, dx \, dy + \int_B \int_B \frac{|g(x) - g(y)|^p}{|x - y|^{n+p\gamma}} \, dx \, dy < \infty \tag{3.3}$$

for every ball $B \subset \mathbb{R}^n$. This comes by the definition of the space $W^{\gamma(1+\delta_1),p(1+\delta_1)}$. On the other hand, when $2\beta < \alpha$, then assumptions (1.18) do not involve any number γ. Thanks to the lower bound on p in (1.18), we can find a negative number γ, such that $|\gamma| \in (0, 1/10)$ is small enough to still verify (3.1). In this case we note that

$$
\begin{aligned}
\int_B \int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+p(1+\delta_1)^2\gamma}}\,dx\,dy &\leq \int_B \int_B \frac{(|g(x)|+|g(y)|)^{p(1+\delta_1)}}{|x-y|^{n+p(1+\delta_1)^2\gamma}}\,dx\,dy \\
&\leq \frac{cr^{-p(1+\delta_1)^2\gamma}}{-\gamma} \int_B |g|^{p(1+\delta_1)}\,dx < \infty \qquad (3.4)
\end{aligned}
$$

where r denotes the radius of B; a similar estimate follows for the second quantity in (3.3). Summarizing, in the rest of the paper we shall always assume that (3.1) and (3.3) hold. In the case $2\beta < \alpha$ the number γ is negative.

Remark 3.1. In the following we shall denote by c_b a constant that depends on $n, \alpha, \Lambda, p, \beta, \gamma$ and exhibits the following blow-up behaviour:

$$
\lim_{p \to 2n/[n+2(\gamma-2\beta+\alpha)]} c_b = \infty\,. \qquad (3.5)
$$

3.2. **The Caccioppoli estimate.** The Caccioppoli type inequality stated in the next theorem is an essential tool in the proof of Theorem 1.1.

Theorem 3.1. *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (1.14) under the assumptions of Theorem 1.1; in particular, (3.1) and (3.3) are in force. Let $B \equiv B(x_0, r) \subset \mathbb{R}^n$ be a ball, and let $\psi \in C_c^\infty(B(x_0, 3r/4))$ be a cut-off function such that $0 \leq \psi \leq 1$ and $|D\psi| \leq c(n)/r$. Then the Caccioppoli type inequality*

$$
\begin{aligned}
&\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy \\
&\quad \leq \frac{c}{r^{2\alpha}} \int_B |u(x)|^2\,dx + c\int_{\mathbb{R}^n\setminus B} \frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy \int_B |u(x)|\,dx \\
&\quad\quad + cr^{n+2\alpha}\left(⨍_B |f(x)|^{2_*}\,dx\right)^{2/2_*} \\
&\quad\quad + c_b r^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta)k}\left(\int_{2^kB} ⨍_{2^kB} \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2 \qquad (3.6)
\end{aligned}
$$

holds for a constant $c \equiv c(n, \Lambda, \alpha)$, which is in particular independent of p, and a constant $c_b \equiv c_b(n, \Lambda, \alpha, \beta, \gamma, p)$. The constant c_b exhibits the behaviour described in (3.5); moreover, all the terms appearing on the right hand side of (3.6) are finite.

Proof. In the weak formulation

$$
\begin{aligned}
&\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(u(x)-u(y))[\eta(x)-\eta(y)]K(x,y)\,dx\,dy \\
&\quad = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} [g(x)-g(y)][\eta(x)-\eta(y)]H(x,y)\,dx\,dy + \int_{\mathbb{R}^n} f\eta\,dx \qquad (3.7)
\end{aligned}
$$

we choose $\eta = u\psi^2$, where $\psi \in C_c^\infty(B)$ is the cut-off function coming from the statement. By a density argument η is an admissible test function. Then we have

$$
\begin{aligned}
&I_1 + I_2 + I_3 := \\
&\quad \int_B \int_B \varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)]K(x,y)\,dx\,dy \\
&\quad\quad + \int_{\mathbb{R}^n\setminus B} \int_B \varphi(u(x)-u(y))u(x)\psi^2(x)K(x,y)\,dx\,dy
\end{aligned}
$$

$$
\begin{aligned}
&- \int_B \int_{\mathbb{R}^n \setminus B} \varphi(u(x)-u(y))u(y)\psi^2(y)K(x,y)\,dx\,dy \\
&\quad = \int_B \int_B [g(x)-g(y)][u(x)\psi^2(x)-u(y)\psi^2(y)]H(x,y)\,dx\,dy \\
&\qquad + \int_{\mathbb{R}^n \setminus B} \int_B [g(x)-g(y)]u(x)\psi^2(x)H(x,y)\,dx\,dy \\
&\qquad + \int_B \int_{\mathbb{R}^n \setminus B} [g(y)-g(x)]u(y)\psi^2(y)H(x,y)\,dx\,dy \\
&\qquad + \int_B f(x)u(x)\psi^2(x)\,dx =: J_1+J_2+J_3+J_4\,.
\end{aligned}
\tag{3.8}
$$

We proceed in estimating the various pieces stemming from the previous identity.
Estimation of I_1. Let us first consider the case in which $\psi(x) \geq \psi(y)$. Then we write

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\quad = \varphi(u(x)-u(y))[u(x)-u(y)]\psi^2(x) + \varphi(u(x)-u(y))u(y)[\psi^2(x)-\psi^2(y)].
\end{aligned}
$$

Applying Young's inequality and recalling the first inequality in (1.15), we have

$$
\begin{aligned}
&\varphi(u(x)-u(y))u(y)[\psi^2(x)-\psi^2(y)] \\
&\quad = \varphi(u(x)-u(y))u(y)[\psi(x)-\psi(y)][\psi(x)+\psi(y)] \\
&\quad \geq -2|\varphi(u(x)-u(y))||u(y)||\psi(x)-\psi(y)|\psi(x) \\
&\quad \geq -\frac{1}{2}|u(x)-u(y)|^2\psi^2(x) - 2\Lambda^2 u^2(y)[\psi(x)-\psi(y)]^2\,.
\end{aligned}
$$

Connecting the content of the last two displays, and using this time the second inequality in (1.15), yields

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\qquad \geq \frac{1}{2}[u(x)-u(y)]^2\psi^2(x) - 2\Lambda^2 u^2(y)[\psi(x)-\psi(y)]^2\,.
\end{aligned}
\tag{3.9}
$$

Now, we consider the case in which $\psi(y) \geq \psi(x)$ and we similarly write

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\quad = \varphi(u(x)-u(y))[u(x)-u(y)]\psi^2(y) + \varphi(u(x)-u(y))u(x)[\psi^2(x)-\psi^2(y)]\,.
\end{aligned}
$$

Proceeding similarly to the case $\psi(x) \geq \psi(y)$, we arrive at

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\qquad \geq \frac{1}{2}[u(x)-u(y)]^2\psi^2(y) - 2\Lambda^2 u^2(x)[\psi(x)-\psi(y)]^2\,.
\end{aligned}
$$

In any case, using also (1.4), we conclude with

$$
\begin{aligned}
I_1 \quad \geq \quad & \frac{1}{c}\int_B \int_B \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}} \max\{\psi^2(x),\psi^2(y)\}\,dx\,dy \\
& -c\int_B \int_B |u(x)|^2 \frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\,,
\end{aligned}
$$

where c depends on Λ. Moreover, by noticing that

$$
[u(x)\psi(x)-u(y)\psi(y)]^2 \leq 2[u(x)(\psi(x)-\psi(y))]^2 + 2[\psi(y)(u(x)-u(y))]^2
$$

and integrating, we conclude with

$$
I_1 \quad \geq \quad \frac{1}{c}\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy
$$

$$-c\int_B\int_B |u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\,. \tag{3.10}$$

Estimation of I_2 and I_3. The estimation of the terms I_2 and I_3 is similar. Indeed, as for I_2, we start observing that a direct computation yields

$$[u(x)-u(y)]u(x)\psi^2(x)K(x,y)\geq -\Lambda\frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}$$

and therefore, by (1.15) we obtain (we can assume without loss of generality that $u(x)\neq u(y)$) that

$$\begin{aligned}\varphi(u(x)-u(y))u(x)\psi^2(x)K(x,y) &\geq -\Lambda\left|\frac{\varphi(u(x)-u(y))}{u(x)-u(y)}\right|\frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}\\ &\geq -\Lambda^2\frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}\,.\end{aligned}$$

Similarly, we obtain

$$-\varphi(u(x)-u(y))u(y)\psi^2(y)K(x,y)\geq -\Lambda^2\frac{|u(x)||u(y)|\psi^2(y)}{|x-y|^{n+2\alpha}}\,.$$

We then estimate

$$\begin{aligned}I_2+I_3 &\geq -c\int_{\mathbb{R}^n\setminus B}\int_B\frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}\,dx\,dy\\ &\geq -c\sup_{z\in\operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|z-y|^{n+2\alpha}}\,dy\int_B|u(x)|\psi^2(x)\,dx\\ &\geq -c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy\int_B|u(x)|\psi^2(x)\,dx\,. \end{aligned}\tag{3.11}$$

Here we have used the fact that since ψ is supported in $B(x_0,3r/4)$, we have

$$\frac{|x_0-y|}{|z-y|}\leq 1+\frac{|x_0-z|}{|z-y|}\leq 4 \tag{3.12}$$

whenever $z\in\operatorname{supp}\psi$ and $y\in\mathbb{R}^n\setminus B$.

Estimation of J_4. Fractional Sobolev's inequality yields

$$\begin{aligned}J_4 &\leq cr^n\left(\fint_B|u(x)\psi(x)|^{2^*}\,dx\right)^{1/2^*}\left(\fint_B|f(x)|^{2_*}\,dx\right)^{1/2_*}\\ &\leq cr^{n/2+\alpha}\left(\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{1/2}\\ &\qquad\cdot\left(\fint_B|f(x)|^{2_*}\,dx\right)^{1/2_*}\,,\end{aligned}$$

so that, applying Young's inequality with $\sigma\in(0,1)$, we have

$$\begin{aligned}J_4 &\leq \frac{c}{\sigma}r^{n+2\alpha}\left(\fint_B|f(x)|^{2_*}\,dx\right)^{2/2_*}\\ &\qquad+\sigma\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\,.\end{aligned}\tag{3.13}$$

The constant c depends only on n,α.

Estimation of J_1. We write

$$\begin{aligned}&u(x)\psi^2(x)-u(y)\psi^2(y)\\ &\quad=[u(x)\psi(x)-u(y)\psi(y)]\psi(y)+u(x)\psi(x)[\psi(x)-\psi(y)]\,.\end{aligned}$$

Therefore, using that $\psi \leq 1$ and (1.16), we have

$$\begin{aligned} J_1 &\leq \Lambda \int_B \int_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta}} |u(x)\psi(x)-u(y)\psi(y)| \frac{dx\,dy}{|x-y|^n} \\ &\quad + \Lambda \int_B \int_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta}} |u(x)\psi(x)||\psi(x)-\psi(y)| \frac{dx\,dy}{|x-y|^n} \\ &=: J_{1.1} + J_{1.2}\,. \end{aligned}$$

In turn, we estimate $J_{1.1}$ and $J_{1.2}$ separately. Recalling (3.1), we now set

$$t := 1 - \frac{2\beta-\gamma}{\alpha} \qquad \text{and} \qquad s := \frac{n}{\alpha}\left[\frac{1}{p} - \frac{1}{2}\right]. \tag{3.14}$$

Observe that $0 < t \leq 1 \Longleftrightarrow 2\beta - \alpha < \gamma \leq 2\beta$. Then we notice that

$$\begin{aligned} &2\beta \geq \gamma \qquad \text{and} \qquad 2 > p > \frac{2n}{n+2(\gamma-2\beta+\alpha)} \\ &\qquad \Longrightarrow \qquad 2 > p > \frac{2n}{n+2\alpha} = 2_* \Longrightarrow 0 < s < 1 \end{aligned} \tag{3.15}$$

and moreover

$$p > \frac{2n}{n+2(\gamma-2\beta+\alpha)} \Longrightarrow 0 < s < t\,. \tag{3.16}$$

We also record the identity $\alpha t = \gamma - (2\beta-\alpha)$. Let us now write

$$\begin{aligned} J_{1.1} &= cr^n \int_B \fint_B \left[r^{\alpha t} \frac{|g(x)-g(y)|}{|x-y|^{2\beta-\alpha+t\alpha}}\right] \left[\frac{|u(x)\psi(x)-u(y)\psi(y)|}{|x-y|^{\alpha}}\right]^{1-s} \\ &\qquad \cdot \left[r^{-\alpha t/s} \frac{|u(x)\psi(x)-u(y)\psi(y)|}{|x-y|^{\alpha(1-t/s)}}\right]^s \frac{dx\,dy}{|x-y|^n}\,. \end{aligned}$$

The definitions in (3.14) imply $(1-s)/2 + s/2^* + 1/p = 1$ and therefore, applying Hölder's inequality with the related choice of the exponents, we have

$$\begin{aligned} J_{1.1} &\leq cr^{n+\alpha t} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p(2\beta-\alpha+t\alpha)}}\,dx\,dy\right)^{1/p} \\ &\quad \cdot \left(\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{(1-s)/2} \\ &\quad \cdot \left(r^{-2^*\alpha t/s} \int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}}\,dx\,dy\right)^{s/2^*}. \end{aligned} \tag{3.17}$$

Before going on, let us estimate the last integral

$$\begin{aligned} \int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}}\,dx\,dy &\leq 2^{2^*-1} \int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}}\,dx\,dy \\ &\leq \frac{cr^{-2^*\alpha(1-t/s)}}{t-s} \fint_B |u(x)\psi(x)|^{2^*}\,dx \\ &\leq \frac{cr^{2^*\alpha t/s}}{t-s} \left(\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\,dy\right)^{2^*/2}. \end{aligned} \tag{3.18}$$

Plugging the inequality into (3.17) yields

$$\begin{aligned} J_{1.1} &\leq cr^{n/2+\alpha t} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p(2\beta-\alpha+t\alpha)}}\,dx\,dy\right)^{1/p} \\ &\quad \cdot \left(\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2}. \end{aligned}$$

Using Young's inequality, and keeping in mind that $\alpha t = \gamma - (2\beta - \alpha)$, leads to

$$\begin{aligned} J_{1.1} \quad &\leq \quad \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy \right)^{2/p} \\ &\quad + \sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\, dy \end{aligned}$$

whenever $\sigma \in (0,1)$. The constant c depends only on $n, \alpha, \Lambda, \beta, \gamma, p$. We then continue with the estimation of $J_{1.2}$. Upon setting $\eta := (1-\alpha)/2$ using Hölder's inequality with conjugate exponents $(2^*, 2_*)$ we have

$$\begin{aligned} J_{1.2} \quad &\leq \quad c\|D\psi\|_{L^\infty} r^n \int_B \fint_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta-1+\eta}} \frac{|u(x)\psi(x)|}{|x-y|^{-\eta}} \frac{dx\, dy}{|x-y|^n} \\ &\leq \quad c\|D\psi\|_{L^\infty} r^n \left(\int_B \fint_B \frac{|g(x)-g(y)|^{2_*}}{|x-y|^{2_*(2\beta-1+\eta)}} \frac{dx\, dy}{|x-y|^n} \right)^{1/2_*} \\ &\qquad \cdot \left(\int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{-2^*\eta}} \frac{dx\, dy}{|x-y|^n} \right)^{1/2^*} . \end{aligned}$$

In turn, by Lemma 2.2 (see also the remark below there) we have

$$\begin{aligned} &\int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{-2^*\eta}} \frac{dx\, dy}{|x-y|^n} \leq \frac{cr^{2^*\eta}}{1-\alpha} \fint_B |u(x)\psi(x)|^{2^*}\, dx \\ &\qquad \leq cr^{2^*(\eta+\alpha)} \left(\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\, dy \right)^{2^*/2} \end{aligned}$$

and, recalling that $p > 2_*$ by (3.15), we proceed with

$$\begin{aligned} &\int_B \fint_B \frac{|g(x)-g(y)|^{2_*}}{|x-y|^{2_*(2\beta-1+\eta)}} \frac{dx\, dy}{|x-y|^n} \\ &\quad = \int_B \fint_B \left(\frac{|g(x)-g(y)|}{|x-y|^{\gamma}} \right)^{2_*} \frac{1}{|x-y|^{2_*(2\beta-1+\eta-\gamma)}} \frac{dx\, dy}{|x-y|^n} \\ &\quad \leq \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy \right)^{2_*/p} \\ &\qquad \cdot \left(\int_B \fint_B \frac{1}{|x-y|^{\frac{2_*(2\beta-1+\eta-\gamma)p}{p-2_*}}} \frac{dx\, dy}{|x-y|^n} \right)^{1-2_*/p} \\ &\quad \leq \frac{cr^{-2_*(2\beta-1+\eta-\gamma)}}{\gamma-2\beta+\alpha} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy \right)^{2_*/p} \end{aligned}$$

where of course we used that $2\beta - 1 + \eta - \gamma = 2\beta - 1/2 - \alpha/2 - \gamma < 2\beta - \alpha - \gamma < 0$ due to $\eta := (1-\alpha)/2$ and (3.1). Connecting the estimates in the last three displays yields

$$\begin{aligned} J_{1.2} \quad &\leq \quad c\|D\psi\|_{L^\infty} r^{n/2+\gamma-2\beta+\alpha+1} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy \right)^{1/p} \\ &\qquad \cdot \left(\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\, dy \right)^{1/2} . \end{aligned}$$

Again using Young's inequality we conclude with

$$\begin{aligned} J_{1.2} \quad &\leq \quad \frac{c}{\sigma} r^2 \|D\psi\|^2_{L^\infty} r^{n+2(\gamma-2\beta+\alpha)} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy \right)^{2/p} \\ &\qquad + \sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\, dy \end{aligned}$$

which holds whenever $\sigma \in (0,1)$. Gathering together the estimates found for $J_{1.1}$ and $J_{1.2}$, and using that $r^2\|D\psi\|^2_{L^\infty} \le c(n)$, gives

$$\begin{aligned} J_1 &\le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{2/p} \\ &\quad + 2\sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\,. \end{aligned} \tag{3.19}$$

The constant c depends on $n, \alpha, \Lambda, \beta, \gamma, p$.

Estimation of J_2 and J_3. The estimation of the two terms is completely similar, and we therefore confine ourselves to estimate J_2. Using (1.16) we have

$$\begin{aligned} J_2 &\le \Lambda \int_{\mathbb{R}^n\setminus B} \int_B \frac{|g(x)-(g)_B|}{|x-y|^{n+2\beta}} |u(x)|\psi^2(x)\,dx\,dy \\ &\quad + \Lambda \int_{\mathbb{R}^n\setminus B} \int_B \frac{|g(y)-(g)_B|}{|x-y|^{n+2\beta}} |u(x)|\psi^2(x)\,dx\,dy \\ &=: J_{2.1} + J_{2.2}\,. \end{aligned}$$

In turn we estimate the two resulting terms. Using that $p \ge 2_*$ by (3.15), we have

$$\begin{aligned} J_{2.1} &\le c \sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{dy}{|z-y|^{n+2\beta}} \int_B |g(x)-(g)_B||u(x)|\psi(x)\,dx \\ &\le c r^n \sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{dy}{|z-y|^{n+2\beta}} \left(\fint_B |g(x)-(g)_B|^{2_*}\,dx \right)^{1/2_*} \\ &\qquad \cdot \left(\fint_B |u(x)\psi(x)|^{2^*}\,dx \right)^{1/2^*} \\ &\le c r^n \sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{dy}{|z-y|^{n+2\beta}} \left(\fint_B |g(x)-(g)_B|^{p}\,dx \right)^{1/p} \\ &\qquad \cdot \left(\fint_B |u(x)\psi(x)|^{2^*}\,dx \right)^{1/2^*} \\ &\le c r^{n/2+\gamma-2\beta+\alpha} \sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{r^{2\beta}dy}{|z-y|^{n+2\beta}} \\ &\quad \cdot \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{1/p} \left(\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy \right)^{1/2}. \end{aligned}$$

Therefore, using Young's inequality, we have

$$\begin{aligned} J_{2.1} &\le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)} \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{2/p} \\ &\quad + \sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy \end{aligned}$$

where we have also used that $\psi \equiv 0$ outside $B(x_0, 3r/4)$, and therefore (3.12), to estimate

$$\sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{r^{2\beta}dy}{|z-y|^{n+2\beta}} \le c(n,\beta)\,.$$

In order to estimate $J_{2.2}$ we need another splitting over annuli. Recalling again that $\psi \le 1$ and that $\psi \equiv 0$ outside $B(x_0, 3r/4)$, we have

$$J_{2.2} \le c \sum_{j=0}^{\infty} \int_{2^{j+1}B\setminus 2^j B} \int_B \frac{|g(y)-(g)_B|}{|x-y|^{n+2\beta}} |u(x)|\psi^2(x)\,dx\,dy$$

$$
\begin{aligned}
&\leq \quad cr^n \sum_{j=0}^{\infty} (2^j r)^{-2\beta} \fint_{2^{j+1}B} |g(y) - (g)_B| \, dy \fint_B |u(x)\psi(x)| \, dx \\
\text{(3.20)} \qquad &\leq \quad cr^n \sum_{j=0}^{\infty} (2^j r)^{-2\beta} \left(\fint_{2^{j+1}B} |g(y) - (g)_B|^p \, dy \right)^{1/p} \fint_B |u(x)\psi(x)| \, dx \,.
\end{aligned}
$$

The estimation of $J_{2.2}$ needs again a splitting; we start by telescoping summation

$$
\begin{aligned}
&\left(\fint_{2^{j+1}B} |g(y) - (g)_B|^p \, dy \right)^{1/p} \\
&\leq \left(\fint_{2^{j+1}B} |g(y) - (g)_{2^{j+1}B}|^p \, dy \right)^{1/p} + \sum_{k=0}^{j} |(g)_{2^{k+1}B} - (g)_{2^k B}| \\
&\leq \left(\fint_{2^{j+1}B} |g(y) - (g)_{2^{j+1}B}|^p \, dy \right)^{1/p} + \sum_{k=0}^{j} \left(\fint_{2^{k+1}B} |g(y) - (g)_{2^k B}|^p \, dy \right)^{1/p} \\
\text{(3.21)} \qquad &\leq 2 \sum_{k=0}^{j+1} \left(\fint_{2^k B} |g(y) - (g)_{2^k B}|^p \, dy \right)^{1/p} .
\end{aligned}
$$

Then an application of the fractional Poincaré inequality in Lemma 2.1 yields

$$
\left(\fint_{2^{j+1}B} |g(y) - (g)_B|^p \, dy \right)^{1/p} \leq c \sum_{k=0}^{j+1} (2^k r)^{\gamma} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} .
$$

Merging the content of the last display with the one of (3.20) gives

$$
\begin{aligned}
J_{2.2} \quad &\leq \quad cr^n \left[\sum_{j=0}^{\infty} \sum_{k=0}^{j+1} (2^j r)^{-2\beta} (2^k r)^{\gamma} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} \right] \\
&\qquad \cdot \fint_B |u(x)\psi(x)| \, dx \,.
\end{aligned}
$$

We now manipulate the content of the square brackets above, using discrete Fubini's theorem as follows:

$$
\begin{aligned}
&\sum_{j=0}^{\infty} \sum_{k=0}^{j+1} (2^j r)^{-2\beta} (2^k r)^{\gamma} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} \\
&\quad = r^{\gamma - 2\beta} \left(\int_B \fint_B \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} \sum_{j=0}^{\infty} 2^{-2\beta j} \\
&\qquad + r^{\gamma - 2\beta} \sum_{k=1}^{\infty} 2^{\gamma k} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} \sum_{j=k-1}^{\infty} 2^{-2\beta j} \\
&\quad \leq \frac{cr^{\gamma - 2\beta}}{\beta} \sum_{k=0}^{\infty} 2^{(\gamma - 2\beta)k} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} .
\end{aligned}
$$

We remark that in the previous display we have used the elementary inequality in (2.2). All in all we have, by using also Hölder's inequality and Lemma 2.1, that

$$
\begin{aligned}
J_{2.2} \quad &\leq \quad cr^{n+\gamma-2\beta} \sum_{k=0}^{\infty} 2^{(\gamma - 2\beta)k} \left(\int_{2^k B} \fint_{2^k B} \frac{|g(x) - g(y)|^p}{|x-y|^{n+p\gamma}} \, dx \, dy \right)^{1/p} \\
&\qquad \cdot \left(\fint_B |u(x)\psi(x)|^{2^*} \, dx \right)^{1/2^*}
\end{aligned}
$$

$$\leq \quad cr^{n/2+\gamma-2\beta+\alpha}\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}$$
$$\cdot\left(\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{1/2}.$$

Finally, using Young's inequality we conclude with

$$J_{2.2} \quad\leq\quad \frac{c}{\sigma}r^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2$$
$$+\sigma\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy$$

whenever $\sigma\in(0,1)$. Connecting the inequalities found for $J_{1.2}$ and $J_{2.2}$, and again recalling that J_3 can be estimates in a completely similar way, we have

$$J_2+J_3$$
$$\leq \frac{c}{\sigma}r^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2$$
$$+4\sigma\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\,. \tag{3.22}$$

The constant c depends on $n,\Lambda,\alpha,\beta,\gamma,p$.

Reabsorbing terms. Inserting the estimates for the terms I_i and J_i into (3.8), we conclude with

$$\frac{1}{c}\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy$$
$$\leq 7\sigma\int_B\int_B\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy$$
$$+c\int_B\int_B|u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy$$
$$+c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy\int_B|u(x)|\psi^2(x)\,dx+\frac{c}{\sigma}r^{n+2\alpha}\left(\fint_B|f(x)|^{2_*}\,dx\right)^{\frac{2}{2_*}}$$
$$+\frac{c_b}{\sigma}r^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2.$$

The constant c depends only on n,α,Λ and the constant c_b depends only on $n,\Lambda,\alpha,\beta,\gamma,p$. Now, taking $\sigma=1/(14c)$ and reabsorbing terms finishes the proof, together with the estimate

$$\int_B\int_B|u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy$$
$$\leq \|D\psi\|_\infty\int_B|u(x)|^2\int_{B_{2r}(x)}|x-y|^{-n+2-2\alpha}\,dy\,dx$$
$$=\frac{c(n)}{1-\alpha}\|D\psi\|_\infty r^{2-2\alpha}\int_B|u(x)|^2\,dx$$
$$\leq\frac{c(n)}{1-\alpha}\frac{1}{r^{2\alpha}}\int_B|u(x)|^2\,dx\,.$$

The finiteness of the terms appearing on the right in (3.6) follows directly from the fact that $u\in W^{\alpha,2}(\mathbb{R}^n)$ and from Section 4.3 below. □

Remark 3.2. In the above statement, one can replace u with $u-(u)_B$ by testing instead of $u\psi^2$ with $(u-(u)_B)\psi^2$.

Remark 3.3. All the constants denoted by c and appearing in Theorem 3.1 blow up as $\alpha \to 0$ or as $\alpha \to 1$. The blow-up of the constant c_b is more peculiar, and it is as in (3.5). This appears for instance in estimate (3.18), as in this case $s \to t$; see (3.16). In terms of assumption (1.12) the blow-up of c_b occurs for instance when $\delta_0 \to 0$. Moreover, the constant c_b blows up also when $\beta \to 0$ and $\gamma \to 2\beta-\alpha$.

4. The dual pair (μ, U) and reverse inequalities

4.1. **A doubling measure.** With ε initially satisfying the condition $0<\varepsilon<\alpha/2$, we consider the locally finite measure μ on $\mathbb{R}^n \times \mathbb{R}^n$ introduced in (1.19). We summarise its basic properties in the next

Proposition 4.1. *With μ being defined as in* (1.19)

- *Whenever $\mathcal{B} = B \times B$, and $B \subset \mathbb{R}^n$ is a ball with radius r, it holds that*
$$\mu(\mathcal{B}) = \frac{c_\varepsilon(n) r^{n+2\varepsilon}}{\varepsilon}\,, \tag{4.1}$$
where $c_\varepsilon(n)$ denotes a constant depending only on n, ε, and it satisfies $1/c(n) \le c_\varepsilon(n) \le c(n)$ for an other constant $c(n)$ depending only on n.
- *(doubling diagonal property) Whenever $A \ge 1$ we have*
$$\sup_{\tilde{x}\in\mathbb{R}^n,\varrho>0} \frac{\mu(\mathcal{B}(\tilde{x}, A\varrho))}{\mu(\mathcal{B}(\tilde{x},\varrho))} = A^{n+2\varepsilon}\,. \tag{4.2}$$
- *For every $A \ge 1$, there exists a constant $c_d \equiv c_d(n,A)$ such that*
$$\frac{\mu(\mathcal{B}(\tilde{x},\varrho))}{\mu(K_1 \times K_2)} \le \frac{c_d}{\varepsilon} \tag{4.3}$$
holds whenever $K_1, K_2 \subset B(\tilde{x},\varrho) \subset \mathbb{R}^n$ are cubes with side parallel to the coordinate axes and such that $|K_1| = |K_2| = \varrho^n/A^n$.
- *(standard doubling property) there exists a constant c, depending only on n, such that*
$$\sup_{\tilde{x},\tilde{y},\in\mathbb{R}^n,\varrho>0} \frac{\mu(\mathcal{B}(\tilde{x},\tilde{y},2\varrho))}{\mu(\mathcal{B}(\tilde{x},\tilde{y},\varrho))} \le \frac{c}{\varepsilon}\,. \tag{4.4}$$

Proof. The proof of (4.1) follows directly from the definition in (1.19) and a scaling argument, while (4.2) follows from (4.1). The proof of (4.3) is slightly less direct. First, observe that $K_1 \times K_2 \subset \mathcal{B}(\tilde{x},\varrho)$ and moreover that $|x-y|<2\varrho$ whenever $x \in K_1$ and $y \in K_2$. Therefore we can estimate

$$\begin{aligned} \mu(\mathcal{B}(\tilde{x},\varrho)) &= \frac{c(n)\varrho^{n+2\varepsilon}}{\varepsilon} \le \frac{c(n)A^{2n}}{\varepsilon}\frac{1}{\varrho^{n-2\varepsilon}}\int_{K_1}\int_{K_2} dx\,dy \\ &\le \frac{c(n)A^{2n}}{\varepsilon}\int_{K_1}\int_{K_2}\frac{dx\,dy}{|x-y|^{n-2\varepsilon}} = \frac{c(n)A^{2n}}{\varepsilon}\mu(K_1\times K_2) \end{aligned}$$

and the proof of (4.3) is complete. The proof of (4.4) is similar to the one of (4.3); this estimate will not be used in the rest of the paper. □

4.2. **Diagonal reverse Hölder type inequalities.** For $(x,y) \in \mathbb{R}^{2n}$, we define the functions

$$U(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{\alpha+\varepsilon}}, \quad G(x,y) := \frac{|g(x)-g(y)|}{|x-y|^{\gamma+2\varepsilon/p}}, \quad F(x,y) := |f(x)|\,, \tag{4.5}$$

the first two being defined when $x \neq y$. According to the Definition 1 the function u generates the dual pair (μ, U). From now on, we shall always assume the following restriction on the number ε:

$$0 < \varepsilon < \min\{\alpha/2, |\gamma|(1+\delta_1)/4, (2\beta-\gamma)p/4\}\,. \tag{4.6}$$

Lemma 4.1. *With the definitions in (4.5) it follows that*

$$U \in L^2(\mathbb{R}^{2n};\mu) \quad \textit{and} \quad F \in L^{2_*+\delta_f}_{\rm loc}(\mathbb{R}^{2n};\mu), \qquad \textit{with} \quad \delta_f \in [0,\delta_0]\,. \tag{4.7}$$

Moreover, assuming (4.6) it follows that

$$G \in L^{p+\delta_g}_{\rm loc}(\mathbb{R}^{2n};\mu) \qquad \textit{where} \quad \delta_g \in [0, p\delta_1] \tag{4.8}$$

Proof. The first inclusion in (4.7) is a direct consequence of the definition in (4.5). As for F, for a ball $\mathcal{B} = B \times B$, where $B \subset \mathbb{R}^n$ has radius $r>0$, we have

$$\int_{\mathcal{B}} F^{2_*+\delta_0}\,d\mu \;=\; \int_B\int_B \frac{|f(x)|^{2_*+\delta_0}}{|x-y|^{n-2\varepsilon}}\,dx\,dy \le \frac{cr^{2\varepsilon}}{\varepsilon}\int_B |f|^{2_*+\delta_0}\,dx\,.$$

This clearly implies that $F \in L^{2_*+\delta_f}_{\rm loc}(\mathbb{R}^{2n};\mu)$ as long as $\delta_f \le \delta_0$. To prove that $G \in L^{p+\delta_g}_{\rm loc}(\mathbb{R}^{2n};\mu)$, let us start with the case $2\beta \ge \alpha$, when $\gamma > 0$. By using (4.6) we have

$$\begin{aligned}\int_{\mathcal{B}} G^{p+p\delta_1}\,d\mu \;&=\; \int_B\int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)+2\varepsilon\delta_1}}\,dx\,dy\\ &\le\; cr^{\delta_1[\gamma p(1+\delta_1)-2\varepsilon]}\int_B\int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)^2}}\,dx\,dy\,.\end{aligned} \tag{4.9}$$

The last quantity is finite since we are assuming $g \in W^{\gamma(1+\delta_1),p(1+\delta_1)}$ so that $G \in L^{p+\delta_g}_{\rm loc}(\mathbb{R}^{2n};\mu)$ follows. We finally treat the case $2\beta < \alpha$. In this case, we have $2\varepsilon\delta_1 < 2\varepsilon < |\gamma|p(1+\delta_1) = -\gamma p(1+\delta_1)$ so that $\gamma p(1+\delta_1) + 2\varepsilon\delta_1 < 0$. We can therefore estimate as follows:

$$\begin{aligned}\int_{\mathcal{B}} G^{p+p\delta_1}\,d\mu \;&\le\; \int_B\int_B \frac{(|g(x)|+|g(y)|)^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)+2\varepsilon\delta_1}}\,dx\,dy\\ &\le\; \frac{cr^{-[\gamma p(1+\delta_1)+2\varepsilon\delta_1]}}{-[\gamma p(1+\delta_1)+2\varepsilon\delta_1]}\int_B |g|^{p(1+\delta_1)}\,dx < \infty\,,\end{aligned} \tag{4.10}$$

and (4.8) follows again since when $2\beta < \alpha$ we are precisely assuming that $g \in L^{p(1+\delta_1)}_{\rm loc}(\mathbb{R}^n)$; see (3.2). □

We are now going to state a few inequalities of later use. Let $v \in W^{\tilde\sigma,q}(B)$ for $\tilde\sigma \in (0,1)$ and $q \ge 1$; then the following fractional Sobolev inequality:

$$⨍_B |v-(v)_B|^2\,dx \le cr^{2\tilde\sigma}\left(\int_B ⨍_B \frac{|v(x)-v(y)|^q}{|x-y|^{n+\tilde\sigma q}}\,dx\,dy\right)^{2/q} \tag{4.11}$$

holds as a consequence of (2.4), provided $q \ge 2n/(n+2\tilde\sigma)$ and $\tilde\sigma > 0$. With $\varepsilon \in (0,\alpha/2)$ we study the compatibility of the following conditions:

$$\tilde\sigma := \alpha + \varepsilon - \frac{2\varepsilon}{q} \qquad \text{and} \qquad q \ge \frac{2n}{n+2\tilde\sigma} \tag{4.12}$$

in inequality (4.11); this gives $q \ge (2n+4\varepsilon)/(n+2\alpha+2\varepsilon)$. Recalling the definition of the function U in (4.5), and using (4.1), we gain

$$r^{2\tilde\sigma}\left(\int_B ⨍_B \frac{|u(x)-u(y)|^q}{|x-y|^{n+\tilde\sigma q}}\,dx\,dy\right)^{2/q} = \frac{c_\varepsilon(n)^{2/q} r^{2\alpha+2\varepsilon}}{\varepsilon^{2/q}}\left(⨍_{\mathcal{B}} U^q d\mu\right)^{2/q}\,,$$

with $c_\varepsilon(n)$ defined in (4.1). We therefore have the following:

Lemma 4.2. *Let $\varepsilon \in (0, \alpha/2)$, and let q be defined by*

$$q := \frac{2n + 4\varepsilon}{n + 2\alpha + 2\varepsilon} < 2\,. \tag{4.13}$$

Then the following inequality:

$$\fint_B |u - (u)_B|^2\, dx \leq \frac{cr^{2(\alpha+\varepsilon)}}{\varepsilon^{2/q}} \left(\fint_{\mathcal{B}} U^q d\mu \right)^{2/q} \tag{4.14}$$

holds for a constant c depending only on n and α, whenever B is a ball with radius r and $\mathcal{B} = B \times B$. The same inequality continues to hold when the ball B is replaced by a cube Q with side of length r, and consequently $\mathcal{B}$ is replaced by $Q \times Q$.

We are now ready for the main result of this section, that is the following:

Proposition 4.2 (Diagonal reverse Hölder type inequality)**.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (1.14) under the assumptions of Theorem 1.1; in particular, (3.1) and (3.3) are in force. Assume that ε satisfies (4.6). Then the following reverse Hölder type inequality with tail holds whenever $\mathcal{B} \subset \mathbb{R}^{2n}$ is a diagonal ball, and $\sigma \in (0,1)$:*

$$\begin{aligned}
\left(\fint_{\mathcal{B}} U^2 d\mu \right)^{1/2} \quad \leq \quad & \frac{c}{\sigma \varepsilon^{1/q-1/2}} \left(\fint_{2\mathcal{B}} U^q d\mu \right)^{1/q} \\
& + \frac{\sigma}{\varepsilon^{1/q-1/2}} \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}} U^q\, d\mu \right)^{1/q} \\
& + \frac{c[\mu(\mathcal{B})]^{\eta}}{\varepsilon^{1/2_*-1/2}} \left(\fint_{2\mathcal{B}} F^{2_*}\, d\mu \right)^{1/2_*} \\
& + \frac{c_b[\mu(\mathcal{B})]^{\theta}}{\varepsilon^{1/p-1/2}} \sum_{k=1}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{2^k\mathcal{B}} G^p\, d\mu \right)^{1/p}\,,
\end{aligned} \tag{4.15}$$

where θ and η denote the following positive exponents:

$$\theta := \frac{\gamma - 2\beta + \alpha + \varepsilon(2/p - 1)}{n + 2\varepsilon} \qquad \text{and} \qquad \eta := \frac{\alpha - \varepsilon}{n + 2\varepsilon}\,, \tag{4.16}$$

respectively. The constant c depends only on n, α, Λ, while the number $q \in (1,2)$ has been defined in (4.13). The constant c_b depends on $n, \alpha, \Lambda, \beta, \gamma, p$ and exhibits the behaviour described in (3.5). The infinite sums on the right side of (4.15) are finite.

Proof. In the rest of the proof all the constants depend *at least* on n, α, Λ. We write $\mathcal{B} \equiv B(x_0, r) \times B(x_0, r)$ and apply Theorem 3.1; we choose the cut-off function $\psi \in C_0^\infty((3/4)B)$ such that $0 \leq \psi \leq 1$, $|D\psi| \leq c(n)/r$ and $\psi \equiv 1$ on $(1/2)B$. Inequality (3.6) remains valid upon replacing u by $u - (u)_B$, see Remark 3.2. Indeed, notice that for such a function all the integrals on the right hand side of (3.6) are finite. For this see Section 4.3 and (4.19) below. All in all we have

$$\begin{aligned}
I_4 &:= \int_B \fint_B \frac{|[u(x) - (u)_B]\psi(x) - [u(y) - (u)_B]\psi(y)|^2}{|x-y|^{n+2\alpha}}\, dx\, dy \\
&\leq cr^{-2\alpha} \fint_B |u(x) - (u)_B|^2\, dx \\
&\quad + c\int_{\mathbb{R}^n \setminus B} \frac{|u(y) - (u)_B|}{|x_0 - y|^{n+2\alpha}}\, dy \fint_B |u(x) - (u)_B|\, dx \\
&\quad + cr^{2\alpha} \left(\fint_B |f(x)|^{2_*}\, dx \right)^{2/2_*}
\end{aligned}$$

$$
+c_b r^{2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2
$$

$$
=: J_5+J_6+J_7+J_8\,. \tag{4.17}
$$

We start rewriting I_4 as follows:

$$
I_4=\frac{1}{|B|}\int_{\mathcal{B}}\frac{|[u(x)-(u)_B]\psi(x)-[u(y)-(u)_B]\psi(y)|^2}{|x-y|^{2(\alpha+\varepsilon)}}\,d\mu(x,y)
$$

so that, with the current choice of ψ, we have

$$
\frac{r^{2\varepsilon}}{\varepsilon}\fint_{\mathcal{B}/2}U^2 d\mu \le \frac{c(n)}{|B|}\int_{\mathcal{B}/2}U^2\,d\mu\le cI_4\,.
$$

We estimate J_5 with the aid of (4.14) as follows:

$$
J_5\le \frac{cr^{2\varepsilon}}{\varepsilon^{2/q}}\left(\fint_{\mathcal{B}}U^q d\mu\right)^{2/q}\,.
$$

To estimate J_6 we split the term in annuli, and proceed somehow as in (3.21). As a matter of fact, we will prove that this term is finite; we indeed have

$$
\begin{aligned}
\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy &= \sum_{j=0}^{\infty}\int_{2^{j+1}B\setminus 2^jB}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy\\
&\le c\sum_{j=0}^{\infty}(2^jr)^{-2\alpha}\fint_{2^{j+1}B}|u(y)-(u)_B|\,dy\,.
\end{aligned} \tag{4.18}
$$

In turn, we again split every integral in the previous sum similarly to (3.21), and using Hölder's inequality we estimate as follows:

$$
\fint_{2^{j+1}B}|u(y)-(u)_B|\,dy\le 2\sum_{k=0}^{j+1}\left(\fint_{2^kB}|u(y)-(u)_{2^kB}|^q\,dy\right)^{1/q}\,.
$$

Each of the previous integrals can be then estimated with the aid of the fractional Poincaré inequality of Lemma 2.1 as follows:

$$
\begin{aligned}
\fint_{2^kB}|u(y)-(u)_{2^kB}|^q\,dy &\le c(2^kr)^{q(\alpha+\varepsilon)-2\varepsilon}\int_{2^kB}\fint_{2^kB}\frac{|u(x)-u(y)|^q}{|x-y|^{n+q\tilde\sigma}}\,dx\,dy\\
&= \frac{c(2^kr)^{q(\alpha+\varepsilon)}}{\varepsilon}\fint_{2^k\mathcal{B}}U^q d\mu\,,
\end{aligned}
$$

where $\tilde\sigma$ is as in (4.12) and c remains independent of ε. As a consequence, we obtain

$$
\fint_{2^{j+1}B}|u(y)-(u)_B|\,dy\le\frac{c}{\varepsilon^{1/q}}\sum_{k=0}^{j+1}(2^kr)^{\alpha+\varepsilon}\left(\fint_{2^k\mathcal{B}}U^q d\mu\right)^{1/q}\,.
$$

Connecting the content of the last display to the one of (4.18), yields

$$
\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy\le\frac{cr^{-\alpha+\varepsilon}}{\varepsilon^{1/q}}\sum_{j=0}^{\infty}\sum_{k=0}^{j+1}2^{-2\alpha j}2^{k(\alpha+\varepsilon)}\left(\fint_{2^k\mathcal{B}}U^q d\mu\right)^{1/q}\,.
$$

Reverting the order of summation gives

$$
\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{k=0}^{j+1}2^{-2\alpha j}2^{k(\alpha+\varepsilon)}\left(\fint_{2^k\mathcal{B}}U^q d\mu\right)^{1/q}\\
&\quad=\left(\fint_{\mathcal{B}}U^q d\mu\right)^{1/q}\sum_{j=0}^{\infty}2^{-2\alpha j}+\sum_{k=1}^{\infty}2^{k(\alpha+\varepsilon)}\left(\fint_{2^k\mathcal{B}}U^q d\mu\right)^{1/q}\sum_{j=k-1}^{\infty}2^{-2\alpha j}
\end{aligned}
$$

$$\leq \frac{c}{\alpha}\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}.$$

Observe that we have once again used the elementary inequality in (2.2) (with $\beta=\alpha$). All in all, combining the content of the last two displays yields

$$\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy \leq \frac{cr^{-\alpha+\varepsilon}}{\varepsilon^{1/q}}\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}, \tag{4.19}$$

so that, via another application of (4.14) we have

$$J_6 \leq \frac{cr^{2\varepsilon}}{\varepsilon^{2/q}}\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{1/q}.$$

With $\sigma\in(0,1)$, using Young's inequality we finally conclude with

$$J_6 \leq \frac{cr^{2\varepsilon}}{\sigma^2\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q} + \frac{\sigma^2 r^{2\varepsilon}}{\varepsilon^{2/q}}\left[\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}\right]^2.$$

For the estimation of J_7 we observe that

$$\begin{aligned}
\fint_B |f(x)|^{2_*}\,dx &= \fint_B\fint_B |f(x)|^{2_*}\,dx\,dy\\
&\leq \frac{c}{r^{n-2\varepsilon+2\varepsilon}}\int_B\fint_B |f(x)|^{2_*}\,dx\,dy\\
&\leq \frac{c}{r^{n+2\varepsilon}}\int_B\int_B \frac{|f(x)|^{2_*}}{|x-y|^{n-2\varepsilon}}\,dx\,dy\\
&\leq \frac{c}{\varepsilon}\fint_{\mathcal{B}} F^{2_*}\,d\mu\,.
\end{aligned}$$

Here we have used (4.1) to perform the last estimation and the very definition of the measure μ. By the definition of J_7 it then follows

$$J_7 \leq \frac{cr^{2\alpha}}{\varepsilon^{2/2_*}}\left(\fint_{\mathcal{B}} F^{2_*}\,d\mu\right)^{2/2_*}.$$

Next, the definitions of $G(\cdot)$ and μ imply

$$J_8 \leq \frac{cr^{2(\gamma-2\beta+\alpha+2\varepsilon/p)}}{\varepsilon^{2/p}}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta+2\varepsilon/p)k}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p}\right]^2.$$

Finally, connecting the estimates found for I_4 and $J_5,\ldots,J_8$ to (4.17) yields

$$\begin{aligned}
\frac{r^{2\varepsilon}}{\varepsilon}\fint_{\mathcal{B}/2} U^2 d\mu \;\leq\; & \frac{cr^{2\varepsilon}}{\sigma^2\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q}\\
&+\frac{\sigma^2 r^{2\varepsilon}}{\varepsilon^{2/q}}\left[\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}\right]^2\\
&+\frac{cr^{2\alpha}}{\varepsilon^{2/2_*}}\left(\fint_{\mathcal{B}} F^{2_*}\,d\mu\right)^{2/2_*}\\
&+\frac{cr^{2(\gamma-2\beta+\alpha+2\varepsilon/p)}}{\varepsilon^{2/p}}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta+2\varepsilon/p)k}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p}\right]^2
\end{aligned}$$

from which (4.15) follows immediately (since the ball $\mathcal{B}$ is arbitrary, and we can switch from $\mathcal{B}$ to $2\mathcal{B}$). The right hand side terms in (4.15) involving infinite sums are finite and this is checked in the next Remark. $\square$

Remark 4.1. A computation based on the definitions in (4.16) gives

$$\frac{2_*\eta}{1-2_*\eta} = \frac{2n(\alpha-\varepsilon)}{n^2+4\varepsilon n+4\alpha\varepsilon} \leq \frac{2}{n}$$

and

$$\frac{p\theta}{1-p\theta} = \frac{p(\gamma-2\beta+\alpha)+\varepsilon(2-p)}{n-p(\gamma-2\beta+\alpha)+\varepsilon p} \leq \frac{3}{n-p(\gamma-2\beta+\alpha)+\varepsilon p} := \Lambda_\theta\,.$$

4.3. **The tails are finite.** We here observe that all the terms on the right hand sides of (3.6) and (4.15) are finite, obviously confining ourselves to those involving infinite sums. We start by the terms involving u. The second term appearing on the right hand side of (3.6) is seen to be finite by estimating

$$\int_{\mathbb{R}^n\setminus B} \frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy \leq \int_{\mathbb{R}^n\setminus B} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy + \int_{\mathbb{R}^n\setminus B} \frac{|(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy\,.$$

The last integral in the above display is obviously finite, while the finiteness of the second one can be obtained as in (4.19). In fact, by (2.2) and $\varepsilon \in (0,\alpha/2)$, the right hand side of (4.19) can be further estimated as

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

This also proves the finiteness of the first infinite sum appearing on the right hand side of (4.15). We now come to the terms involving g, proving that the last series appearing in (4.15) is finite. The finiteness of the last series appearing in (3.6) is therefore implied by looking at the estimate for the term J_8 in the proof of Proposition 4.2. We start by the case $2\beta \geq \alpha$, where using (4.9) we have

$$2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p} \leq c2^{-k(2\beta-\gamma-\delta_1\gamma+n/[p(1+\delta_1)])}[g]_{\gamma(1+\delta_1),p(1+\delta_1)}$$

with $c \equiv c(n,\beta,\gamma,p,\delta_1,r)$. and since by (3.1) we have $\gamma < 2\beta$ and $\delta_1\gamma p(1+\delta_1) \leq n$, the convergence of the series follows. In the case $2\beta < \alpha$ we instead use (4.10) to have the following inequality, that again implies the convergence of the series in question:

$$2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p} \leq c2^{-k(2\beta+n/[p(1+\delta_1)])}\|g\|_{L^{p(1+\delta_1)}(\mathbb{R}^n)}\,.$$

5. Level sets estimates for dual pairs

In this section we prove a level set estimate which is at the core of the proof of our higher differentiability and integrability results. Let us first define a few functionals. With θ and η as in (4.16), for every $\mathcal{B} \equiv \mathcal{B}(x,\varrho) \subset \mathbb{R}^{2n}$ we define

$$\begin{aligned}\Psi_{H,M}(\mathcal{B}(x,\varrho)) &:= \left(\fint_{\mathcal{B}(x,\varrho)} U^2\,d\mu\right)^{1/2} + \frac{H[\mu(\mathcal{B}(x,\varrho))]^\eta}{\varepsilon^{1/2_*-1/2}}\left(\fint_{\mathcal{B}(x,\varrho)} F^{2_*}\,d\mu\right)^{1/2_*}\\ &\quad +\frac{M[\mu(\mathcal{B}(x,\varrho))]^\theta}{\varepsilon^{1/p-1/2}}\left(\fint_{\mathcal{B}(x,\varrho)} G^p\,d\mu\right)^{1/p},\end{aligned} \tag{5.1}$$

where $H, M \geq 1$ and $\mathcal{B}(x,\varrho) \subset \mathbb{R}^{2n}$. We also define the functionals

$$\Upsilon_0(\mathcal{B}(x,\varrho)) := \left(\fint_{\mathcal{B}(x,\varrho)} F^{2_*+\delta_f}\,d\mu\right)^{1/(2_*+\delta_f)} + \left(\fint_{\mathcal{B}(x,\varrho)} G^{p+\delta_g}\,d\mu\right)^{1/(p+\delta_g)}, \tag{5.2}$$

$$\Upsilon_1(\mathcal{B}(x,\varrho)) := \sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{\mathcal{B}(x,2^k\varrho)} U^q \, d\mu \right)^{1/q} \tag{5.3}$$

and

$$\Upsilon_{2,M}(\mathcal{B}(x,\varrho)) := \frac{M[\mu(\mathcal{B}(x,\varrho))]^\theta}{\varepsilon^{1/p-1/2}} \sum_{k=0}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{\mathcal{B}(x,2^k\varrho)} G^p \, d\mu \right)^{1/p} . \tag{5.4}$$

We shall denote

$$\Psi(\mathcal{B}(x,\varrho)) := \Psi_{1,1}(\mathcal{B}(x,\varrho)).$$

and shall often use the abbreviations

$$\Psi_{H,M}(\mathcal{B}(x,\varrho)) \equiv \Psi_{H,M}(x,\varrho)\,, \qquad \Upsilon_0(\mathcal{B}(x,\varrho)) \equiv \Upsilon_0(x,\varrho)$$

and so forth. Finally, we can define

$$\mathrm{ADD}(\mathcal{B}(x,\varrho)) \equiv \mathrm{ADD}(x,\varrho) := \Psi(x,\varrho) + \Upsilon_0(x,\varrho) + \Upsilon_1(x,\varrho) + \Upsilon_{2,1}(x,\varrho)\,. \tag{5.5}$$

The aim of this section is to prove the following:

Proposition 5.1. *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (1.14) under the assumptions of Theorem 1.1; in particular, (3.1) and (3.3) are in force. Let μ be the measure defined in (1.19), with ε satisfying (4.6). Consider a ball $\mathcal{B}(x_0,2\varrho_0) \subset \mathbb{R}^{2n}$ such that $\varrho_0 \le 1$, and related concentric balls*

$$\mathcal{B}(x_0,\varrho_0) \subset \mathcal{B}(x_0,t) \subset \mathcal{B}(x_0,s) \subset \mathcal{B}(x_0,3\varrho_0/2) \tag{5.6}$$

for $\varrho_0 < t < s < 3\varrho_0/2$. There exists a constant $c_s \equiv c_s(n,\alpha,\Lambda)$ independent of ε and p, and constants $c_f \equiv c_f(n,\alpha,\Lambda,\varepsilon) > 1$, $c_g \equiv c_g(n,\alpha,\Lambda,\beta,\gamma,p,\varepsilon) \ge 1$, $\kappa_f \equiv \kappa_f(n,\alpha,\Lambda,\varepsilon) \in (0,1)$, $\kappa_g \equiv \kappa_g(n,\alpha,\Lambda,\beta,p,\varepsilon) \in (0,1)$, such that the inequality

$$\begin{aligned} \frac{1}{\lambda^2} \int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2 \, d\mu \quad &\le \quad \frac{c_s}{\varepsilon^{3(2-q)/q}\lambda^q} \int_{\mathcal{B}(x_0,s)\cap\{U>\lambda\}} U^q d\mu \\ &\quad + \frac{c_f \lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)}} \int_{\mathcal{B}(x_0,s)\cap\{F>\kappa_f\lambda\}} F^{2_*} \, d\mu \\ &\quad + \frac{c_g \lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}(x_0,s)\cap\{G>\kappa_g\lambda\}} G^p \, d\mu \end{aligned} \tag{5.7}$$

holds whenever $\lambda \ge \lambda_0$, where

$$\lambda_0 := \frac{c_a}{\varepsilon} \left(\frac{\varrho_0}{s-t} \right)^{2n} \mathrm{ADD}(x_0, 2\varrho_0)\,. \tag{5.8}$$

The constant c_a introduced in the last display depends on $n,\alpha,\Lambda,\beta,\gamma$, but is still independent of ε.

Remark 5.1. Unlike κ_f, c_f, the constants κ_g, c_g exhibit the following behaviour:

$$\lim_{p\to 2n/[n+2(\gamma-2\beta+\alpha)]} \frac{1}{\kappa_g} = \lim_{\gamma\to 2\beta} \frac{1}{\kappa_g} = \infty = \lim_{p\to 2n/[n+2(\gamma-2\beta+\alpha)]} c_g = \lim_{\gamma\to 2\beta} c_g\,. \tag{5.9}$$

The proof of Proposition 5.1 is rather delicate and falls into twelve steps. It will take the rest of this section.

5.1. **Diagonal balls and Vitali's covering.** The proof starts with an exit time argument for the functional $\Psi_{H,M}(\cdot)$, aimed at covering the "diagonal" level set of U. The constants $H, M \geq 1$ shall be fixed in due course of the proof, and the whole argument is independent of their particular values until the moment these are fixed. They will be used to give a different weight to the integrals of F^{2_*} and G^p: at the exit time, the averages of F^{2_*}, G^p will be smaller than the one of U^2 provided H, M are chosen to be large enough, respectively. Let us consider concentric diagonal balls as in (5.6). Let $\kappa \in (0,1]$ be a free parameter to be again chosen in due course of the proof and define

$$\tilde{\lambda}_0 := \kappa^{-1} \sup_{\frac{s-t}{40^n} \leq \varrho \leq \frac{\varrho_0}{2}} \ \sup_{x \in B(x_0,t)} \left\{ \Psi_{H,M}(x,\varrho) + \Upsilon_0(x,\varrho) + \Upsilon_1(x,\varrho) + \Upsilon_{2,M}(x,\varrho) \right\}. \tag{5.10}$$

All the foregoing steps of proofs are independent of the specific choice of κ until this will be in fact made in (5.55) below. For the same κ (to be defined later) and for $\lambda \geq \tilde{\lambda}_0$, define further the "diagonal level set"

$$D_{\kappa\lambda} := \left\{ (x,x) \in \mathcal{B}(x_0,t) \ : \ \sup_{0<\varrho<\frac{s-t}{40^n}} \Psi_{H,M}(x,\varrho) > \kappa\lambda \right\}. \tag{5.11}$$

Since, by the definition in (5.10), for every $(x,x) \in \mathcal{B}(x_0,t)$ it follows

$$\Psi_{H,M}(x,\varrho) \leq \kappa\tilde{\lambda}_0 \leq \kappa\lambda \qquad \text{whenever } \varrho \in [(s-t)/40^n, \varrho_0/2], \tag{5.12}$$

then we find for all $(x,x) \in D_{\kappa\lambda}$ the exit radius $\varrho(x) \in (0,(s-t)/40^n)$ such that

$$\Psi_{H,M}(x,\varrho(x)) \geq \kappa\lambda, \qquad \text{while} \quad \sup_{\varrho(x)<\varrho<\frac{s-t}{40^n}} \Psi_{H,M}(x,\varrho) \leq \kappa\lambda. \tag{5.13}$$

Collect enlarged balls into the covering $\{\mathcal{B}(x,2\varrho(x)) \ : \ (x,x) \in D_{\kappa\lambda}\}$. Balls of the type $\mathcal{B}(x,t,\varrho)$ are, as explained in Section 2, metric balls with respect to the metric (2.1). We therefore apply Vitali's covering theorem to find a countable set J_D, and related diagonal points $\{(x_j,x_j)\}_{j\in J_D}$, such that

$$\bigcup_{(x,x)\in D_{\kappa\lambda}} \mathcal{B}(x,2\varrho(x)) \subset \bigcup_{j\in J_D} \mathcal{B}(x_j,10\varrho(x_j)) \subset \mathcal{B}(x_0,s) \tag{5.14}$$

and

$$\{\mathcal{B}(x_j,2\varrho(x_j))\}_{j\in J_D} \quad \text{is a family of mutually disjoint balls}. \tag{5.15}$$

Notice that, implicit in (5.14), is the fact that since $\varrho(x_j) \leq (s-t)/40^n$ and $x_j \in B(x_0,t)$ for every $x_j \in J_D$, then $\mathcal{B}(x_j,10\varrho(x_j)) \subset \mathcal{B}(x_0,s)$. By (5.12)-(5.13) and the doubling property in (4.2), it follows that

$$\begin{aligned}
\sum_{j\in J_D} \int_{\mathcal{B}(x_j,10\varrho(x_j))} U^2 \, d\mu \quad &\leq \quad \sum_{j\in J_D} \mu(\mathcal{B}(x_j,10\varrho(x_j)))[\Psi_{H,M}(\mathcal{B}(x_j,10\varrho(x_j)))]^2 \\
&\leq \quad 10^{n+2\varepsilon}\kappa^2\lambda^2 \sum_{j\in J_D} \mu(\mathcal{B}(x_j,\varrho(x_j))) \,.
\end{aligned} \tag{5.16}$$

We shall denote in short

$$\mathcal{B}_j := \mathcal{B}(x_j,\varrho(x_j)), \qquad \sigma\mathcal{B}_j := \mathcal{B}(x_j,\sigma\varrho(x_j)), \qquad \sigma > 0. \tag{5.17}$$

Finally, since we are assuming that $\varrho_0 \leq 1$, by (4.1) we observe that

$$\mu(\mathcal{B}(x_0,2\varrho_0)) \leq \frac{c2^{n+2\varepsilon}}{\varepsilon} =: L \equiv L(n,\varepsilon). \tag{5.18}$$

5.2. **Dyadic cubes, and two constants.** This section has a very technical nature, and reports a few facts that are true independently of the specify context we are moving in. In order to cover the off-diagonal level sets of U, we need a more elaborate argument based on classical Calderón-Zygmund coverings. To this aim, we start recalling basic properties of dyadic cubes in $\mathbb{R}^{2n}$. They differ from the usual ones since they are "centred" at x_0 and the is size is adapted to the size of the starting ball $\mathcal{B}(x_0, s)$. Define

$$k_0 := \left[-\log_2 \left(\frac{s-t}{n10^{10n}} \right) \right] + 1 \, , \tag{5.19}$$

where $[\cdot]$ denotes the integer part of a given number, with the (unnecessary large) constant 10^{10n} having also a symbolic meaning. Let Δ_k, $k \geq k_0$, be the disjoint collection - centered at x_0 - of half-open cubes of sidelength 2^{-k} whose closures are touching $\overline{B}(x_0, (s+t)/2)$, i.e.

$$\begin{aligned} \Delta_k := \{ & x_0 + 2^{-k}v + [0, 2^{-k})^n \ : \\ & v \in \mathbb{Z}^n \, , \ (x_0 + 2^{-k}v + [0, 2^{-k}]^n) \cap \overline{B}(x_0, (s+t)/2) \neq \emptyset \} \, . \end{aligned}$$

Notice that, with such a definition, by using (5.19) it follows that $k \geq k_0$ implies

$$B(x_0, t) \subset \bigcup_{K \in \triangle_k} K \subset B(x_0, s) \, . \tag{5.20}$$

The cubes defined above are, up to a translation aimed at centring everything at x_0, the standard dyadic cubes in $\mathbb{R}^n$. Let us recall a few basic properties. Let Δ the family of all cubes from the families Δ_k, that is $\Delta := \{K \in \Delta_k \ : \ k \geq k_0\}$. Defined this way, every cube K in Δ_{k+1}, $k \geq k_0$, has only one predecessor $\widetilde{K} \in \Delta_k$ such that $K \subset \widetilde{K}$. Moreover, if $K_1 \in \Delta_{k_1}$ and $K_2 \in \Delta_{k_2}$ with $k_0 \leq k_1 < k_2$ and also $K_1 \cap K_2 \neq \emptyset$, then $K_2 \subset K_1$. Starting from the previous cubes, we fix the notation for the corresponding ones in $\mathbb{R}^{2n}$. We set, again for $k \geq k_0$

$$\Xi_k := \{ \mathcal{K} \equiv K_1 \times K_2 \ : \ K_1, K_2 \in \Delta_k \} \, , \qquad \Xi := \bigcup_{k \geq k_0} \Xi_k \, ,$$

while the diagonal cubes build up the family

$$\tilde{\Xi}_k := \{ \mathcal{K} \equiv K \times K \ : \ K \in \Delta_k \} \, . \tag{5.21}$$

With the above definition it follows from (5.20) that

$$\mathcal{B}(x_0, t) \subset \bigcup_{\mathcal{K} \in \Xi_k} \mathcal{K} \subset \mathcal{B}(x_0, s) \tag{5.22}$$

holds whenever $k \geq k_0$. Notice that, by defining the product cubes as above, we are actually once again considering dyadic cubes in $\mathbb{R}^{2n}$, with the same properties of the cubes from $\triangle_k$. We also notice that if $\Xi \ni \mathcal{K} = K_1 \times K_2$ then $\widetilde{\mathcal{K}} = \tilde{K}_1 \times \tilde{K}_2$ is its unique predecessor. Finally, let $\mathcal{K} \in \Xi$; then there exist $K_1, K_2 \in \triangle_k$ such that $\mathcal{K} = K_1 \times K_2$; in this case we let

$$k(\mathcal{K}) = k \, . \tag{5.23}$$

Next, again with $\mathcal{K} = K_1 \times K_2$, we define the *cube projections* as

$$P_1(\mathcal{K}) \equiv P_1\mathcal{K} := K_1 \times K_1 \qquad \text{and} \qquad P_2(\mathcal{K}) \equiv P_2\mathcal{K} := K_2 \times K_2 \, ,$$

whenever $K_1, K_2 \in \Delta_k$. In order to shorten the notation, in the following we shall also write $P_h(\mathcal{K}) = P_h\mathcal{K}$ for $h = 1, 2$. It hence follows

$$P_1(K_1 \times K_2) = P_2(K_2 \times K_1) \, . \tag{5.24}$$

For a given cube $\mathcal{K} \equiv K_1 \times K_2$ we define

$$\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) := \operatorname{dist}(K_1, K_2) \tag{5.25}$$

and its symmetric (or mirror reflected) cube with respect to the diagonal Diag, is defined by

$$\operatorname{Symm}(\mathcal{K}) = \operatorname{Symm}(K_1 \times K_2) := K_2 \times K_1 \,. \tag{5.26}$$

For future convenience we collect a few basic facts that are a direct consequence of the definitions above, and in particular of (5.23)-(5.26).

Proposition 5.2. *Let $\mathcal{K} = K_1 \times K_2 \in \Xi$. The following facts are true:*

- $P_1\mathcal{K}, P_2\mathcal{K} \in \Xi$.
- $\mu(P_1\mathcal{K}) = \mu(P_2\mathcal{K})$ *and* $k(\mathcal{K}) = k(P_1\mathcal{K}) = k(P_2\mathcal{K})$.
- *If* $\Xi \ni \mathcal{H} \subset \mathcal{K}$, *then* $k(\mathcal{K}) \leq k(\mathcal{H})$.
- *If $\widetilde{\mathcal{K}}$ is the predecessor of $\mathcal{K}$, then*

$$\tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \leq \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \,. \tag{5.27}$$

- *The following relations hold:*

$$\operatorname{dist}(P_1\mathcal{K}, P_2\mathcal{K}) = \sqrt{2}\,\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \tag{5.28}$$

$$\operatorname{dist}(\mathcal{K}, \operatorname{Diag}) = \frac{\operatorname{dist}(P_1\mathcal{K}, P_2\mathcal{K})}{2} = \frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{\sqrt{2}} \tag{5.29}$$

$$\begin{aligned} \operatorname{dist}(\mathcal{K}, P_1\mathcal{K}) = \operatorname{dist}(\mathcal{K}, P_2\mathcal{K}) &= \operatorname{dist}(K_1, K_2) = \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \\ \tilde{\operatorname{dist}}(P_1\operatorname{Symm}(\mathcal{K}), P_2\operatorname{Symm}(\mathcal{K})) &= \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \,. \end{aligned} \tag{5.30}$$

- *Let $F\colon (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a locally μ-integrable function which is symmetric; i.e. $F(x,y) = F(y,x)$ holds for every $x,y \in \mathbb{R}^n$, then*

$$\int_{\mathcal{K}} F\,d\mu = \int_{\operatorname{Symm}(\mathcal{K})} F\,d\mu$$

holds whenever $\mathcal{K} \in \Xi$. In particular, $\mu(\mathcal{K}) = \mu(\operatorname{Symm}(\mathcal{K}))$ and, moreover, it holds that $k(\mathcal{K}) = k(\operatorname{Symm}(\mathcal{K}))$.

In the next two lemmas we introduce the ε-independent constants c_{dd} and $\tilde{c}_d$, and these will be very often used in the following.

Lemma 5.1. *There exists a constant c_{dd}, depending only on n, and in particular independent of ε, such that the following inequality holds true for $h \in \{1,2\}$:*

$$\begin{aligned} c_{dd} \quad \geq \quad & \sup_{\mathcal{K}\in\Xi} \left\{ \frac{1}{\varepsilon} \left(\frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}} \right)^{n-2\varepsilon} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \right\} \\ & + \sup_{\mathcal{K}\in\Xi\,, \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{K})}} \varepsilon \left\{ \left(\frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}} \right)^{2\varepsilon-n} \frac{\mu(P_h\mathcal{K})}{\mu(\mathcal{K})} \right\} + 1 \,. \end{aligned} \tag{5.31}$$

Proof. Indeed, observe that that using the definition of the measure μ together with (5.25) (and assuming without loss of generality that $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) > 0$) we have

$$\mu(P_h\mathcal{K}) = \frac{c(n)}{\varepsilon} 2^{-k(\mathcal{K})(n+2\varepsilon)} \quad \text{and} \quad \mu(\mathcal{K}) \leq \frac{2^{-2k(\mathcal{K})n}}{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})^{n-2\varepsilon}} \,.$$

This allows to bound the first quantity in (5.31) in a universal way

$$\frac{1}{\varepsilon} \left(\frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}} \right)^{n-2\varepsilon} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \leq c(n) \,.$$

On the other hand, again by the definition (5.25), notice that if $x \in K_1$ and $y \in K_2$ then $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \leq |x-y| \leq 2\sqrt{n}[2^{-k(\mathcal{K})} + \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})]$ so that, the very definition of the measure μ yields

$$\mu(\mathcal{K}) \geq \frac{2^{-2k(\mathcal{K})n}}{(2\sqrt{n})^{n-2\varepsilon}[2^{-k(\mathcal{K})} + \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})]^{n-2\varepsilon}}\,.$$

Then we have

$$\begin{aligned} &\varepsilon \left(\frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{2\varepsilon-n} \frac{\mu(P_h\mathcal{K})}{\mu(\mathcal{K})} \\ &\quad \leq c(n)\left(\frac{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{2\varepsilon-n} \frac{[2^{-k(\mathcal{K})} + \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})]^{n-2\varepsilon}}{2^{-2k(\mathcal{K})n+k(\mathcal{K})(n+2\varepsilon)}} \leq c(n) \end{aligned}$$

where we have used that $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{K})}$. We have therefore proved that (5.31) holds for a constant c_{dd} depending only on n. □

The second constant is presented in the next

Lemma 5.2. *There exists a constant $\tilde{c}_d$, depending only on n, in particular independent of ε, such that the following inequality holds:*

$$\sup_{\substack{\widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K} \\ \tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \geq 2^{-k(\mathcal{K})}}} \frac{\mu(\widetilde{\mathcal{K}})}{\mu(\mathcal{K})} \leq \tilde{c}_d\,. \tag{5.32}$$

Proof. Let us consider a dyadic cube $\mathcal{K} = K_1 \times K_2 \subset \mathbb{R}^{2n}$, with $\widetilde{\mathcal{K}}$ being its predecessor, and such that $\tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \geq 2^{-k(\mathcal{K})}$. Triangle inequality gives

$$|x-y| \leq 2\sqrt{n}2^{-k(\mathcal{K})+1} + \operatorname{dist}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \leq 8\sqrt{n}\operatorname{dist}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})$$

whenever $(x,y) \in K_1 \times K_2$. By the very definition of μ and (5.25), and finally using the inequality in the previous line when performing the final estimation, we get

$$\begin{aligned} \mu(\widetilde{\mathcal{K}}) &\leq \tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})^{-(n-2\varepsilon)}|\widetilde{K_1} \times \widetilde{K_2}| \\ &= 4^n\tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})^{-(n-2\varepsilon)}|K_1 \times K_2| \leq c(n)\mu(\mathcal{K})\,, \end{aligned}$$

and the proof of the lemma is complete. □

5.3. **Off-diagonal cubes and Calderón-Zygmund coverings.** We start reporting an adaptation of the classical Calderón-Zygmund decomposition lemma. The argument is completely similar to the classical one and for a proof we refer for instance to [27], taking into account that the measure μ is doubling and absolutely continuous with respect to the Lebesgue measure.

Theorem 5.1. *Let Q_0 being a cube in $\mathbb{R}^{2n}$ and let $\tilde{U}$ be a non-negative function in $L^1(Q_0)$. Let $\tilde{\lambda}$ be a real number such that*

$$\fint_{Q_0} \tilde{U}\,d\mu \leq \tilde{\lambda}\,.$$

There exists a countable, but possibly finite, family of pairwise disjoint dyadic cubes $\{Q_i\}$, with sides parallel to those of Q_0, such that

$$\tilde{\lambda} < \fint_{Q_i} \tilde{U}\,d\mu \qquad \text{and} \qquad \fint_{\widetilde{Q_i}} \tilde{U}\,d\mu \leq \tilde{\lambda} \qquad \text{holds for every } Q_i$$

where $\widetilde{Q_i}$ denotes the predecessor of Q_i, and

$$\tilde{U} \leq \tilde{\lambda} \qquad a.e. \text{ in } \quad Q_0 \setminus \bigcup_i Q_i\,.$$

We now start to cover the off-diagonal part of the level set of U. To this end, let us consider the cubes from the family Ξ_{k_0} and, accordingly, the quantity

$$\lambda_1 := \max\left\{\tilde{\lambda}_0, \sup_{\mathcal{K}\in\Xi_{k_0}}\left(\fint_{\mathcal{K}} U^2\,d\mu\right)^{1/2}\right\}. \tag{5.33}$$

We recall that the numbers $\tilde{\lambda}_0$ and k_0 have been determined in (5.10) and (5.19), respectively. Let us observe that (5.22) implies that the family $\{\mathcal{K}\}_{\mathcal{K}\in\Xi_{k_0}}$ forms a disjoint covering of $\mathcal{B}(x_0,t)$. With $\lambda \geq \lambda_1$ we now apply Theorem 5.1 with the choice $Q_0 \equiv \mathcal{K}_0$, for every single cube $\mathcal{K}_0 \in \Xi_{k_0}$; we therefore obtain a family of disjoint dyadic cubes $Q_i(\mathcal{K}_0)$ such that

$$\lambda^2 < \fint_{Q_i(\mathcal{K}_0)} U^2 d\mu \qquad \text{and} \qquad \fint_{\widetilde{Q_i(\mathcal{K}_0)}} U^2 d\mu \leq \lambda^2 \qquad \text{holds for every } Q_i\,,$$

where, as usual, $\widetilde{Q_i(\mathcal{K}_0)}$ denotes the predecessor of $Q_i(\mathcal{K}_0)$, and

$$U \leq \lambda \qquad \text{holds a.e. in} \quad \mathcal{K}_0 \setminus \bigcup_i Q_i(\mathcal{K}_0)\,.$$

Putting all such families of cubes together we get a countable family

$$\mathcal{U}_\lambda := \bigcup_{\mathcal{K}_0\in\Xi_{k_0}} \{Q_i(\mathcal{K}_0)\} \equiv \{\mathcal{K}\}$$

of disjoint dyadic cubes $\mathcal{K}$ which are such that

$$\lambda^2 < \fint_{\mathcal{K}} U^2 d\mu \qquad \text{and} \qquad \fint_{\widetilde{\mathcal{K}}} U^2 d\mu \leq \lambda^2 \qquad \text{holds for every } \mathcal{K}\in\mathcal{U}_\lambda \tag{5.34}$$

where $\widetilde{\mathcal{K}}$ denotes the predecessor of $\mathcal{K}$, and such that

$$U \leq \lambda \qquad \text{holds a.e. in} \quad \mathcal{B}(x_0,t) \setminus \bigcup_{\mathcal{K}\in\mathcal{U}_\lambda} \mathcal{K}\,. \tag{5.35}$$

Remark 5.2. The symmetry of the function U and Proposition 5.2 imply that

$$\int_{\mathcal{K}} U^2\,d\mu = \int_{\mathrm{Symm}(\mathcal{K})} U^2\,d\mu$$

holds whenever $\mathcal{K} \in \Xi$. It then follows that $\mathcal{K} \in \mathcal{U}_\lambda$, iff $\mathrm{Symm}(\mathcal{K}) \in \mathcal{U}_\lambda$.

5.4. **First removal of nearly diagonal cubes.** In this step we are going to show that, in order to cover the level sets of U^2, it is sufficient to restrict our attention to those dyadic cubes that are "far" from the diagonal in a suitably quantified sense. Specifically, the word far refers to the fact that for such cubes it happens that their distance to the diagonal is larger than their size. These are really the relevant cubes to analyse, since we shall see that the remaining ones can be covered by the balls considered in (5.14)-(5.15). We therefore start considering the following family of *nearly diagonal* cubes:

$$\mathcal{U}_\lambda^d := \left\{\mathcal{K}\in\mathcal{U}_\lambda \;:\; \tilde{\mathrm{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) < 2^{-k(\mathcal{K})}\,,\ \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K}\right\}.$$

With $\mathcal{K} \in \mathcal{U}_\lambda^d$, consider now a point $(\tilde{x},\tilde{x}) \in \mathrm{Diag}$ such that $\mathrm{dist}((\tilde{x},\tilde{x}),\widetilde{\mathcal{K}}) = \mathrm{dist}(\mathrm{Diag},\widetilde{\mathcal{K}})$ and a diagonal ball $\mathcal{B}(\tilde{x},\varrho) \subset \mathbb{R}^{2n}$ with radius ϱ larger or equal than

$$\frac{5\sqrt{n}\,\tilde{\mathrm{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})}{2} + 5\sqrt{n}2^{-k(\mathcal{K})+1}\,.$$

Keeping (5.29) in mind and applying it to $\widetilde{\mathcal{K}}$, it follows that $\widetilde{\mathcal{K}} \subset \mathcal{B}(\tilde{x},\varrho)$. Ultimately, we can find a diagonal ball $\mathcal{B} \equiv \mathcal{B}(\tilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$, such that $\mathcal{K} \subset \mathcal{B}$. Notice that in this case, by using (4.3) from Proposition 4.1 and recalling that $(\tilde{x}, \tilde{x}) \in \mathrm{Diag}$, we conclude there exists a constant c_d, which is only depending on n, such that

$$1 \le \frac{\mu(\mathcal{B})}{\mu(\mathcal{K})} \le \frac{c_d}{\varepsilon} \equiv \frac{c_d(n)}{\varepsilon}\,. \tag{5.36}$$

Therefore, if $\mathcal{K} \in \mathcal{U}_\lambda^d$, then the lower bound in (5.34) yields

$$\lambda^2 < \fint_{\mathcal{K}} U^2\,d\mu \le \frac{\mu(\mathcal{B})}{\mu(\mathcal{K})} \fint_{\mathcal{B}} U^2\,d\mu \le \frac{c_d}{\varepsilon} \fint_{\mathcal{B}} U^2\,d\mu\,.$$

Assuming that the number $\kappa \in (0,1]$ introduced in (5.10) satisfies

$$\kappa \in (0, \kappa_0]\,, \qquad \kappa_0 := \frac{\varepsilon^{1/2}}{\sqrt{2c_d}}\,, \tag{5.37}$$

all in all we have proved that

$$\forall\ \mathcal{K} \in \mathcal{U}_\lambda^d \quad \exists\ \mathcal{B}^{\mathcal{K}} \equiv B^{\mathcal{K}} \times B^{\mathcal{K}} \quad \text{s.t.} \quad \fint_{\mathcal{B}^{\mathcal{K}}} U^2\,d\mu > \kappa^2\lambda^2 \quad \text{and} \quad \mathcal{K} \subset \mathcal{B}^{\mathcal{K}}\,.$$

This means that, being $\tilde{x}$ the centre of $\mathcal{B}^{\mathcal{K}}$, by the exit time condition (5.13) it follows that $(\tilde{x}, \tilde{x}) \in D_{\kappa\lambda}$ and then $\mathcal{B}^{\mathcal{K}} \subset B(\tilde{x}, \varrho(\tilde{x}))$. By (5.14) it hence follows that

$$\bigcup_{\mathcal{K} \in \mathcal{U}_\lambda^d} \mathcal{K} \subset \bigcup_{j \in J_D} 10\mathcal{B}_j\,. \tag{5.38}$$

Notice that here, in order to find the ball $\mathcal{B}^{\mathcal{K}}$ and apply the exit time condition in (5.13), we have used that the radius of the diagonal ball $\mathcal{B} \equiv \mathcal{B}(\tilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$ is smaller that $(s-t)/40^n$. In turn, this is a consequence of the fact that $k(\mathcal{K}) \ge k_0$ and of the fact that k_0 is large enough as prescribed in (5.19).

5.5. **Off-diagonal reverse Hölder inequalities.** As we saw in the previous section, $\mathcal{U}_\lambda^d$ has already been covered by the diagonal cover. Thus, we shall now only consider so-called off-diagonal cubes:

$$\mathcal{U}_\lambda^{nd} := \left\{ \mathcal{K} \in \mathcal{U}_\lambda \ :\ \tilde{\mathrm{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \ge 2^{-k(\mathcal{K})}\,,\ \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K} \right\}. \tag{5.39}$$

We notice that (5.27) implies

$$\mathcal{K} \in \mathcal{U}_\lambda^{nd} \qquad \Longrightarrow \qquad \tilde{\mathrm{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \ge 2^{-k(\mathcal{K})}\,.$$

The goal is thus to sort and estimate suitable off-diagonal sums of the measures of cubes belonging to $\mathcal{U}_\lambda^{nd}$. The following lemma is our basic tool. It roughly tells that for non-diagonal cubes reverse Hölder inequalities hold automatically, and independently of the fact that the function solves an equation. The prize to pay is the appearance of certain correction diagonal terms, and this is eventually treated by some combinatorial lemmas.

Lemma 5.3 (Off-diagonal reverse inequality)**.** *Let $k \ge k_0$ and suppose that $\mathcal{K} \in \Xi_k$. There exists a constant $c_{nd} \equiv c_{nd}(n, \alpha)$, independent of ε, such that whenever $\tilde{\mathrm{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \ge 2^{-k}$ the inequality*

$$\begin{aligned} \left(\fint_{\mathcal{K}} U^2\,d\mu\right)^{1/2} \quad \le \quad & c_{nd}\left(\fint_{\mathcal{K}} U^q\,d\mu\right)^{1/q} \\ & + \frac{c_{nd}}{\varepsilon^{1/q}}\left(\frac{2^{-k}}{\tilde{\mathrm{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{P_1\mathcal{K}} U^q\,d\mu\right)^{1/q} \\ & + \frac{c_{nd}}{\varepsilon^{1/q}}\left(\frac{2^{-k}}{\tilde{\mathrm{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{P_2\mathcal{K}} U^q\,d\mu\right)^{1/q} \end{aligned}$$

holds with the number q being defined in (4.13). In particular, the above inequality holds whenever $\mathcal{K} \in \mathcal{U}^{nd}_{\lambda}$.

Proof. Let $\mathcal{K} \equiv K_1 \times K_2 \in \Xi_k$ and find points $x_1 \in \overline{K}_1$ and $y_1 \in \overline{K}_2$ such that $\operatorname{dist}(K_1, K_2) = |x_1 - y_1|$. By the triangle inequality we obtain, whenever $x, y \in \mathcal{K}$

$$\begin{aligned} |x-y| &\leq \operatorname{dist}(K_1,K_2) + |x_1 - x| + |y_1 - y| \\ &\leq \operatorname{dist}(K_1,K_2) + 2\sqrt{n}2^{-k} \\ &\leq 3\sqrt{n}\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) = 3\sqrt{n}\operatorname{dist}(K_1,K_2). \end{aligned}$$

Therefore we have

$$1 \leq \frac{|x-y|}{\operatorname{dist}(K_1,K_2)} \leq 3\sqrt{n} \qquad \forall\ (x,y) \in \mathcal{K}\,, \tag{5.40}$$

with the first inequality in the above display which is a trivial consequence of the definition of $\operatorname{dist}(K_1, K_2)$. Next, thanks to (5.40), the very definition of μ yields

$$\mu(\mathcal{K}) \approx \frac{4^{-nk}}{\operatorname{dist}(K_1,K_2)^{n-2\varepsilon}}\,, \tag{5.41}$$

with the constant involved being independent of ε, but just depending on n. By using (5.40) and (5.41) we then have

$$\begin{aligned} &\left(\fint_{\mathcal{K}} U^2\,d\mu\right)^{1/2} \\ &\quad = \left(\frac{1}{\mu(\mathcal{K})}\int_{K_1}\int_{K_2} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2} \\ &\quad \leq c\left(\frac{\operatorname{dist}(K_1,K_2)^{n-2\varepsilon-(n+2\alpha)}}{4^{-nk(\mathcal{K})}}\int_{K_1}\int_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2} \\ &\quad \leq c\operatorname{dist}(K_1,K_2)^{-(\alpha+\varepsilon)}\left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2}\,, \end{aligned} \tag{5.42}$$

where c depends only on n. We further estimate the integral on the right using Minkowski's inequality:

$$\begin{aligned} \left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2} &\leq \left(\fint_{K_1}|u(x)-(u)_{K_1}|^2\,dx\right)^{1/2} \\ &\quad + \left(\fint_{K_2}|u(x)-(u)_{K_2}|^2\,dx\right)^{1/2} \\ &\quad + |(u)_{K_1}-(u)_{K_2}|\,. \end{aligned}$$

By using the fractional Poincaré inequality of Lemma 4.2 applied on cubes, and recalling that $P_h\mathcal{K} = K_h \times K_h$ for $h \in \{1, 2\}$, we deduce that

$$\left(\fint_{K_h}|u(x)-(u)_{K_h}|^2\,dx\right)^{1/2} \leq \frac{c2^{-k(\alpha+\varepsilon)}}{\varepsilon^{1/q}}\left(\fint_{P_h\mathcal{K}} U^q d\mu\right)^{1/q}\,, \qquad h \in \{1,2\}$$

with the implied constant c depending only on n and α. Finally, by Hölder's inequality, and using (5.40) and (5.41) repeatedly, we get

$$\begin{aligned} |(u)_{K_1}-(u)_{K_2}| &\leq \fint_{K_1}\fint_{K_2}|u(x)-u(y)|\,dx\,dy \\ &\leq \left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^q\,dx\,dy\right)^{1/q} \end{aligned}$$

$$
\begin{aligned}
&\leq c\left(\frac{1}{\operatorname{dist}(K_1,K_2)^{n-2\varepsilon}\mu(\mathcal{K})}\int_{K_1}\int_{K_2}|u(x)-u(y)|^q\,dx\,dy\right)^{1/q}\\
&\leq c\left(\fint_{\mathcal{K}}|u(x)-u(y)|^q\,d\mu\right)^{1/q}\\
&\leq c\operatorname{dist}(K_1,K_2)^{\alpha+\varepsilon}\left(\fint_{\mathcal{K}}U^q\,d\mu\right)^{1/q}
\end{aligned}
$$

with $c \equiv c(n)$. Combining the content of the last four displays and recalling the definition in (5.25) finishes the proof. □

We remark that the previous lemma works for any function $u \in W^{\alpha,2}$ and does not require that u solves any equation; moreover, the lemma works for every positive integer k. Applying it in the present situation we instead get the following:

Corollary 5.1. *Let $k \geq k_0$ be an integer, and suppose that $\mathcal{K} \in \Xi_k$ is such that $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k}$ holds. Assume that*

$$
\left(\fint_{\mathcal{K}}U^2\,d\mu\right)^{1/2} \geq \lambda
$$

and that the number κ introduced in (5.10) satisfies

$$
\kappa \in (0,\kappa_1]\,, \qquad \kappa_1 := \frac{\varepsilon^{1/q}}{2^{1/q}3c_{nd}}\,, \tag{5.43}
$$

where $c_{nd} \equiv c_{nd}(n,\alpha)$ has been defined in Lemma 5.3. Then it holds that

$$
\begin{aligned}
\mu(\mathcal{K}) \leq\ & \frac{3^q c_{nd}^q}{\lambda^q}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&+\frac{3^q c_{nd}^q}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&+\frac{3^q c_{nd}^q}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.
\end{aligned} \tag{5.44}
$$

In particular, the inequality (5.44) holds whenever $\mathcal{K} \in \mathcal{U}_\lambda^{nd}$.

Proof. Appealing to Lemma 5.3, and using the elementary inequality $(a+b+c)^q \leq 3^{q-1}(a^q+b^q+c^q)$ valid for all nonnegative numbers $a,b,c \in \mathbb{R}$, we get

$$
\frac{\lambda^q}{3^{q-1}c_{nd}^q} \leq \fint_{\mathcal{K}}U^q\,d\mu + \frac{1}{\varepsilon}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\left(\fint_{P_1\mathcal{K}}U^q\,d\mu+\fint_{P_2\mathcal{K}}U^q\,d\mu\right).
$$

To estimate the integrals appearing on the right hand side we note that by (5.43) we have

$$
\fint_E U^q\,d\mu \leq \kappa_1^q\lambda^q + \frac{1}{\mu(E)}\int_{E\cap\{U>\kappa\lambda\}}U^q\,d\mu
$$

with $E \in \{\mathcal{K}, P_1\mathcal{K}, P_2\mathcal{K}\}$ so that, recalling that $\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K}) \geq 2^{-k}$, we gain

$$
\begin{aligned}
\frac{\lambda^q}{3^{q-1}c_{nd}^q} \leq\ & \frac{3\kappa_1^q\lambda^q}{\varepsilon}+\frac{1}{\mu(\mathcal{K})}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&+\frac{1}{\varepsilon\mu(P_1\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&+\frac{1}{\varepsilon\mu(P_2\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.
\end{aligned}
$$

Now (5.44) follows inserting (5.43) in the last estimate and reabsorbing terms. □

5.6. **Families of off-diagonal cubes.** With $\mathcal{U}^{nd}_\lambda$ that has been defined in (5.39), consider now the families

$$\mathcal{M}^h_\lambda := \left\{\mathcal{K} \in \mathcal{U}^{nd}_\lambda \ : \ \fint_{P_h\mathcal{K}} U^q \, d\mu \le (10n)^{n+2}\kappa^q\lambda^q\right\} \tag{5.45}$$

and

$$\mathcal{N}^h_\lambda := \left\{\mathcal{K} \in \mathcal{U}^{nd}_\lambda \ : \ \fint_{P_h\mathcal{K}} U^q \, d\mu > (10n)^{n+2}\kappa^q\lambda^q\right\} \tag{5.46}$$

for $h \in \{1,2\}$, where the number κ has been introduced in (5.10) and q is defined in (4.13). Furthermore define

$$\mathcal{M}_\lambda := \mathcal{M}^1_\lambda \cap \mathcal{M}^2_\lambda \qquad \text{and} \qquad \mathcal{N}_\lambda := \mathcal{N}^1_\lambda \cup \mathcal{N}^2_\lambda \tag{5.47}$$

so that the following decomposition in disjoint families holds:

$$\mathcal{U}^{nd}_\lambda = \mathcal{M}_\lambda \cup \mathcal{N}_\lambda \, . \tag{5.48}$$

We then have the following:

Lemma 5.4 (Soft off-diagonal summation)**.** *The inequality*

$$\sum_{\mathcal{K}\in\mathcal{M}_\lambda} \mu(\mathcal{K}) \le \frac{6^q c^q_{nd}}{\lambda^q} \int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q \, d\mu \tag{5.49}$$

holds whenever the number κ in (5.10) satisfies

$$\kappa \in (0, \kappa_2] \, , \qquad \kappa_2 := \frac{\varepsilon^{1/q}}{8^{1/q}3c_{nd}(10n)^{(n+2)/q}} \, . \tag{5.50}$$

The constant $c_{nd} \equiv c_{nd}(n,\alpha)$ has been defined in Lemma 5.3 and appears in Corollary 5.1; it is independent of ε.

Proof. It is sufficient to prove that if $\mathcal{K} \in \mathcal{M}_\lambda$, then

$$\mu(\mathcal{K}) \le \frac{6^q c^q_{nd}}{\lambda^q} \int_{\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \, . \tag{5.51}$$

After this (5.49) follows since the initial family $\mathcal{U}_\lambda$ is disjoint and (5.22) holds. For the proof of (5.51), notice that if $\mathcal{K} \in \mathcal{M}^h_\lambda$, then we have, for $h \in \{1,2\}$, that

$$\begin{aligned} &\frac{3^q c^q_{nd}}{\varepsilon\lambda^q} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ &\quad \le \mu(\mathcal{K}) \frac{3^q c^q_{nd}}{\varepsilon\lambda^q} \fint_{P_h\mathcal{K}} U^q \, d\mu \le \mu(\mathcal{K}) \frac{3^q c^q_{nd}}{\varepsilon\lambda^q} (10n)^{n+2}\kappa^q\lambda^q \le \frac{\mu(\mathcal{K})}{8} \, . \end{aligned} \tag{5.52}$$

Using this last estimate for $h \in \{1,2\}$ in combination with (5.44), and reabsorbing terms, gives (5.51); the proof is therefore complete. □

It remains to study the family $\mathcal{N}_\lambda$ defined in (5.47). To this aim, we introduce the family of diagonal cubes defined by

$$P_h\mathcal{N}_\lambda := \{P_h\mathcal{K} \ : \ \mathcal{K} \in \mathcal{N}^h_\lambda\} \, , \qquad h \in \{1,2\} \, .$$

Keeping (5.24) and Remark 5.2 in mind, we have that

$$\mathcal{K} \in \mathcal{N}^1_\lambda \Longleftrightarrow \operatorname{Symm}(\mathcal{K}) \in \mathcal{N}^2_\lambda \tag{5.53}$$

whenever $\mathcal{K} \in \Xi$. Now, let us make a remark; consider $T \in P_1\mathcal{N}_\lambda$, then $T = P_1(\mathcal{K})$ for some $\mathcal{K} \in \mathcal{N}^1_\lambda$. Therefore $T = P_2(\operatorname{Symm}(\mathcal{K}))$ by (5.24) and by (5.53) we have $\operatorname{Symm}(\mathcal{K}) \in \mathcal{N}^2_\lambda$. We conclude that $T \in P_2\mathcal{N}_\lambda$ and eventually that $P_1\mathcal{N}_\lambda \subset P_2\mathcal{N}_\lambda$.

In a similar way it follows $P_2\mathcal{N}_\lambda \subset P_1\mathcal{N}_\lambda$. We therefore conclude that $P_1\mathcal{N}_\lambda = P_2\mathcal{N}_\lambda = P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$. Let $P\mathcal{N}_\lambda$ be a disjoint subfamily of $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$ such that

$$\bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \mathcal{H} = \bigcup_{\mathcal{K}\in P_1\mathcal{N}_\lambda\cup P_2\mathcal{N}_\lambda} \mathcal{K}\,. \tag{5.54}$$

Note that, since all the cubes of the family $P\mathcal{N}_\lambda$ are themselves dyadic cubes, such an extracted disjoint covering always exists. We remark that a straightforward consequence of the definitions is that all cubes from $P\mathcal{N}_\lambda$ obviously belong to $P_1\mathcal{N}_\lambda\cup P_2\mathcal{N}_\lambda$ and are therefore diagonal cubes.

5.7. **Determining κ.** We here determine the parameter κ in (5.10). By choosing

$$\kappa := \min\{\kappa_0, \kappa_1, \kappa_2\} \equiv \min\left\{\frac{\varepsilon^{1/2}}{\sqrt{2c_d}}, \frac{\varepsilon^{1/q}}{2^{1/q}3c_{nd}}, \frac{\varepsilon^{1/q}}{8^{1/q}3c_{nd}(10n)^{(n+2)/q}}\right\}, \tag{5.55}$$

conditions (5.37), (5.43) and (5.50) are all satisfied. Therefore the content and the results of Sections 5.4-5.6 are at our disposal. Recalling that c_d in (5.36) (coming from Proposition 4.1) depends only on n, and that c_{nd} from Lemma 5.3 depends only on n, α, we conclude there exists a new constant c_κ, such that

$$\kappa \geq \varepsilon^{1/q}/c_\kappa\,, \qquad c_\kappa \equiv c_\kappa(n,\alpha)\,. \tag{5.56}$$

5.8. **Further removal of nearly-diagonal cubes.** We recall that our final goal is to estimate the measure of the level sets of U. Since the nearly diagonal part has already been covered, we proceed in excluding from the subsequent analysis those cubes covered by the balls in (5.14)-(5.15). Therefore we introduce

$$\mathcal{N}_{\lambda,d} := \left\{\mathcal{K}\in\mathcal{N}_\lambda \;:\; \mathcal{K}\subset \bigcup_{j\in J_D} 10\mathcal{B}_j\right\} \tag{5.57}$$

and, accordingly

$$\mathcal{N}_{\lambda,nd} := \mathcal{N}_\lambda\setminus\mathcal{N}_{\lambda,d} \qquad \text{and} \qquad \mathcal{N}^h_{\lambda,nd} := \mathcal{N}_{\lambda,nd}\cap\mathcal{N}^h_\lambda\,, \quad \text{for} \quad h\in\{1,2\}\,. \tag{5.58}$$

We observe that the main difficulty in handling the cubes from the family $P\mathcal{N}_\lambda$ stems from the fact that they do not belong to the family $\mathcal{U}_\lambda$, i.e. they do not come from an exit time argument and therefore no control is available on the values taken by U^2 on such cubes. This will be bypassed via a very delicate combinatorial argument. The next lemma is instrumental to that.

Lemma 5.5. *Let $\mathcal{K}\in\mathcal{N}_{\lambda,nd}$ be such that $P_h\mathcal{K}\subset\mathcal{H}$ for some $\mathcal{H}\in P\mathcal{N}_\lambda$ and some $h\in\{1,2\}$. Then $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})\geq 2^{-k(\mathcal{H})}$ holds.*

Proof. First, let us consider a cube $\mathcal{H}\in P\mathcal{N}_\lambda$; take the diagonal ball $\mathcal{B}(\mathcal{H})\equiv \mathcal{B}(x_\mathcal{H}, 2^{-(k(\mathcal{H})+1)})$, $(x_\mathcal{H}, x_\mathcal{H})$ being the center of $\mathcal{H}$. It follows that

$$\mathcal{B}(\mathcal{H})\subset\mathcal{H}\subset\sqrt{n}\mathcal{B}(\mathcal{H})\,. \tag{5.59}$$

Therefore we have by Hölder's inequality and the definition of $P\mathcal{N}_\lambda$ that

$$\begin{aligned}(10n)^{(n+2)/q}\kappa\lambda &< \left(\fint_{\mathcal{H}} U^q\,d\mu\right)^{1/q}\\ &\leq \left(\frac{\mu(10n\mathcal{B}(\mathcal{H}))}{\mu(\mathcal{B}(\mathcal{H}))}\fint_{10n\mathcal{B}(\mathcal{H})} U^q\,d\mu\right)^{1/q}\\ &\leq (10n)^{(n+2)/q}\left(\fint_{10n\mathcal{B}(\mathcal{H})} U^2\,d\mu\right)^{1/2}.\end{aligned} \tag{5.60}$$

By the definition of $D_{\kappa\lambda}$ in (5.11) it follows that $(x_{\mathcal{H}}, x_{\mathcal{H}}) \in D_{\kappa\lambda}$ and then the exit time condition (5.13) gives $\mathcal{B}(\mathcal{H}) \subset \mathcal{B}(x_{\mathcal{H}}, \varrho(x_{\mathcal{H}}))$. We are using that the radius of the ball $10n\mathcal{B}(\mathcal{H})$ is smaller that $(s-t)/40^n$. In turn, this is a consequence of the fact that $k(\mathcal{H})+1 \geq k_0$ and of (5.19). Then (5.14) implies

$$10n\mathcal{B}(\mathcal{H}) \subset \bigcup_{j\in J_D} 10\mathcal{B}_j\,. \tag{5.61}$$

Now, in order to prove the lemma, assume by contradiction that $\tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) < 2^{-k(\mathcal{H})}$ and let $\mathcal{B}(\mathcal{H})$ be the ball determined in (5.59), and for which (5.61) holds. We are going to show that

$$\mathcal{K} \subset 10n\mathcal{B}(\mathcal{H}) \tag{5.62}$$

holds, and this then contradicts the assumption $\mathcal{K} \in \mathcal{N}_{\lambda,nd}$ by (5.61). In order to show (5.62) we observe that Proposition 5.2 and the fact that $P_h\mathcal{K} \subset \mathcal{H}$ give

$$\operatorname{dist}(\mathcal{K},\mathcal{H}) \leq \operatorname{dist}(\mathcal{K}, P_h\mathcal{K}) = \tilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \leq 2^{-k(\mathcal{H})}\,.$$

Again by Proposition 5.2 we have $k(P_h\mathcal{K}) = k(\mathcal{K})$ and $k(\mathcal{K}) \geq k(\mathcal{H})$. Therefore, since $\mathcal{H} \subset \sqrt{n}\mathcal{B}(\mathcal{H})$ and the radius of $\mathcal{B}(\mathcal{H})$ is $2^{-(k(\mathcal{H})+1)}$, then (5.62) must hold. The proof of the lemma is complete. □

5.9. **Summation in $\mathcal{N}_{\lambda,nd}$.** The aim of this section is to prove the following:

Lemma 5.6 (Hard off-diagonal summation)**.** *There exists a constant c, depending only on n, α, such that the estimate*

$$\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}} \mu(\mathcal{K}) \leq \frac{c}{\lambda^q} \int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q\, d\mu \tag{5.63}$$

holds, where κ has been determined in (5.55).

Proof. Step 1: Classifying cubes. Here we classify the cubes from $\mathcal{N}_{\lambda,nd}$ according to their projections, thereby partitioning $\mathcal{N}_{\lambda,nd}$ in suitable disjoint subfamilies. For every $\mathcal{H} \in P\mathcal{N}_\lambda$ set

$$\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) := \{\mathcal{K} \in \mathcal{N}_{\lambda,nd} \ : \ P_h\mathcal{K} \subset \mathcal{H}\}\,, \qquad h \in \{1,2\}\,.$$

Since $P\mathcal{N}_\lambda$ is a disjoint covering of $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda = P_1\mathcal{N}_\lambda = P_2\mathcal{N}_\lambda$, we have the following decomposition in mutually disjoint families:

$$\mathcal{N}^h_{\lambda,nd} = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \mathcal{N}^h_{\lambda,nd}(\mathcal{H})\,. \tag{5.64}$$

This means that for $\mathcal{H}_1, \mathcal{H}_2 \in P\mathcal{N}_\lambda$ it follows that $\mathcal{N}^h_{\lambda,nd}(\mathcal{H}_1) \cap \mathcal{N}^h_{\lambda,nd}(\mathcal{H}_2) \neq \emptyset$ implies $\mathcal{H}_1 = \mathcal{H}_2$. In fact, assume that a cube $\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H}_1) \cap \mathcal{N}^h_{\lambda,nd}(\mathcal{H}_2)$ and $\mathcal{H}_1 \neq \mathcal{H}_2$, then we would have that $P_h\mathcal{K} \subset \mathcal{H}_1 \cap \mathcal{H}_2$ against the fact that $\mathcal{H}_1$ and $\mathcal{H}_2$ have a non-empty intersection, being elements of the disjoint covering $P\mathcal{N}_\lambda$. Next, let us recall that for every $\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H})$ it is $k(\mathcal{K}) = k(P_h\mathcal{K}) \geq k(\mathcal{H})$, and this leads us to define the following classes:

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i := \left\{\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H}) \ : \ k(\mathcal{K}) = i + k(\mathcal{H})\right\}$$

for $h \in \{1,2\}$ and for every integer $i \geq 0$. Therefore, the decomposition in mutually disjoint families

$$\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) = \bigcup_{i\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i$$

holds, in the sense that $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i \cap [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_j \neq \emptyset$ implies that $i = j$. Next, take $\mathcal{H} \in P\mathcal{N}_\lambda$; by Lemma 5.5 we have that if $\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H})$, that is if $P_h\mathcal{K} \subset \mathcal{H}$, then it follows that $\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{H})}$ and this finally leads us to classify elements of $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i$ in the following way:

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j} := \left\{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i \ : \ 2^{j-k(\mathcal{H})} \leq \tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) < 2^{j+1-k(\mathcal{H})}\right\} ,$$

for $h \in \{1,2\}$ and $i, j \geq 0$ being integers. Again we have the decomposition

$$\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) = \bigcup_{i,j\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j} \tag{5.65}$$

and these are disjoint classes in the sense that, if $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i_1,j_1} \cap [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i_2,j_2} \neq \emptyset$, then it is $(i_1, j_1) = (i_2, j_2)$. All in all, keeping (5.64) and (5.65) in mind, we have that the following decomposition in mutually disjoint classes holds:

$$\mathcal{N}^h_{\lambda,nd} = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \bigcup_{i,j\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j} \, . \tag{5.66}$$

Step 2: Sums and further partitions. Let us fix $\mathcal{H} \in P\mathcal{N}_\lambda$; our aim here is to prove that the following inequality holds for $h \in \{1,2\}$:

$$\begin{aligned} &\frac{1}{\varepsilon} \sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ &\qquad \leq \frac{c(n)}{\alpha^2} \int_{\mathcal{H}\cap\{U>\kappa\lambda\}} U^q \, d\mu \, . \end{aligned} \tag{5.67}$$

We start by recalling that, by the very definitions in (5.46) and (5.47), and again (5.27), we have that $\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{K})}$ as soon as $\mathcal{K} \in \mathcal{N}_{\lambda,nd}$; (5.31) yields

$$\frac{1}{\varepsilon} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \leq c_{dd} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{n-2\varepsilon} ,$$

for $h \in \{1,2\}$, and moreover, if $\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$, we also have that

$$\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})} = \frac{1}{2^i} \frac{2^{-k(\mathcal{H})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})} \leq \frac{1}{2^{i+j}} \, .$$

Using the inequalities in the last two displays we can estimate as follows:

$$\begin{aligned} &\frac{1}{\varepsilon} \sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ &\leq c_{dd} \sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ &= c_{dd} \sum_{i,j=0}^{\infty} \sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}} \left(\frac{2^{-k(\mathcal{K})}}{\tilde{\text{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ &\leq c(n) \sum_{i,j=0}^{\infty} \left(\frac{1}{2^{i+j}}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon} \sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \, . \end{aligned} \tag{5.68}$$

In order to evaluate the last sum we have to further decompose $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$. For each integer $i \geq 0$, $\mathcal{H}$ contains precisely $4^{ni} = 2^{2ni}$ disjoint cubes from $\Xi_{i+k(\mathcal{H})}$ and exactly 2^{ni} disjoint cubes from $\tilde{\Xi}_{i+k(\mathcal{H})}$; see the definition in (5.21) and in the preceding display. As a consequence, it contains at most 2^{ni} disjoint (diagonal) cubes from the class $\tilde{\Xi}_{i+k(\mathcal{H})} \cap (P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda)$. We anyway consider all the diagonal cubes $\tilde{\Xi}_{i+k(\mathcal{H})}$ from $\mathcal{H}$ and relabel them as follows:

$$\{\widetilde{\mathcal{H}} \in \tilde{\Xi}_{i+k(\mathcal{H})} \,:\, \widetilde{\mathcal{H}} \subset \mathcal{H}\} = \{\mathcal{H}_i^m \,:\, 1 \leq m \leq 2^{ni}\}\,, \tag{5.69}$$

so that, in particular

$$\sum_{m=1}^{2^{ni}} \int_{\mathcal{H}_i^m \cap \{U > \kappa\lambda\}} U^q \, d\mu \leq \int_{\mathcal{H} \cap \{U > \kappa\lambda\}} U^q \, d\mu\,. \tag{5.70}$$

Now, let us concentrate one moment on the elements of $[\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j}$, a similar argument then apply to $[\mathcal{N}^2_{\lambda,nd}(\mathcal{H})]_{i,j}$. For any $\mathcal{K} \in [\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j}$, there is the unique cube from the diagonal class (5.21), that we denote by $\mathcal{H}_i^m(\mathcal{K})$, such that $P_1\mathcal{K} = \mathcal{H}_i^m(\mathcal{K})$. Now note that for $h \in \{1,2\}$ one can split $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$ as

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m} := \left\{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j} \,:\, P_h\mathcal{K} = \mathcal{H}_i^m\right\}\,, \qquad m \in \{1, \ldots, 2^{ni}\}\,.$$

Since $\mathcal{N}^1_{\lambda,nd}$ is a family of dyadic cubes, we must have that if we have that $\mathcal{K}_1, \mathcal{K}_2 \in [\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j,m}$ and $\mathcal{K}_1 \neq \mathcal{K}_2$, then $P_2\mathcal{K}_1 \cap P_2\mathcal{K}_2 = \emptyset$, i.e., the second components are disjoint (otherwise the two cubes would coincide). A similar argument holds when looking at $\mathcal{N}^2_{\lambda,nd}$. It then follows that

$$\#[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m} \leq c(n)2^{n(i+j)}\,, \qquad h \in \{1,2\}\,, \tag{5.71}$$

for every choice of $i,j \geq 0$ and $m \in \{1, \ldots, 2^{ni}\}$. We use now use (5.70)-(5.71) to estimate as follows:

$$\begin{aligned}
\sum_{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}} \int_{P_h\mathcal{K} \cap \{U > \kappa\lambda\}} U^q \, d\mu &= \sum_{m=1}^{2^{ni}} \sum_{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m}} \int_{\mathcal{H}_i^m \cap \{U > \kappa\lambda\}} U^q \, d\mu \\
&\leq c(n)2^{n(i+j)} \sum_{m=1}^{2^{ni}} \int_{\mathcal{H}_i^m \cap \{U > \kappa\lambda\}} U^q \, d\mu \\
&\leq c(n)2^{n(i+j)} \int_{\mathcal{H} \cap \{U > \kappa\lambda\}} U^q \, d\mu\,.
\end{aligned}$$

Using also (2.2) it then follows:

$$\begin{aligned}
&\sum_{i,j=0}^{\infty} \left(\frac{1}{2^{i+j}}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon} \sum_{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}} \int_{P_h\mathcal{K} \cap \{U > \kappa\lambda\}} U^q \, d\mu \\
&\qquad \leq c(n) \sum_{i,j=0}^{\infty} \left(\frac{1}{2^{i+j}}\right)^{q(\alpha+\varepsilon)-2\varepsilon} \int_{\mathcal{H} \cap \{U > \kappa\lambda\}} U^q \, d\mu \\
&\qquad \leq \frac{c(n)}{[q(\alpha+\varepsilon)-2\varepsilon]^2} \int_{\mathcal{H} \cap \{U > \kappa\lambda\}} U^q \, d\mu \\
&\qquad \leq \frac{c(n)}{\alpha^2} \int_{\mathcal{H} \cap \{U > \kappa\lambda\}} U^q \, d\mu\,.
\end{aligned}$$

Notice that we have used that, since $q > 1$ and $\varepsilon < \alpha/2$, it is $q(\alpha+\varepsilon) - 2\varepsilon > \alpha/2$. Combining the inequality in the last display with (5.68) yields (5.67).

Step 3: Summation. Let now $\mathcal{K} \in \mathcal{N}^1_{\lambda,nd}$. There are then two cases: either $\mathcal{K} \in \mathcal{M}^2_\lambda$ or $\mathcal{K} \in \mathcal{N}^2_\lambda$ (the relevant definitions are in (5.45), (5.46) and (5.58)). Now, if $\mathcal{K} \in \mathcal{M}^2_\lambda$, then using (5.44) and (5.52), and reabsorbing terms, we obtain that

$$\mu(\mathcal{K}) \leq \frac{6^q c^q_{nd}}{\lambda^q} \int_{\mathcal{K} \cap \{U > \kappa\lambda\}} U^q \, d\mu$$

$$+\frac{6^q c_{nd}^q}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.$$

If, on the other hand, $\mathcal{K}\in\mathcal{N}^2_\lambda$, then using (5.44) we get

$$\begin{aligned}\mu(\mathcal{K}) \quad\leq\quad & \frac{3^q c_{nd}^q}{\lambda^q}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\ &+\frac{3^q c_{nd}^q}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\ &+\frac{3^q c_{nd}^q}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})}\left(\frac{2^{-k}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.\end{aligned}$$

A similar reasoning holds if $\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}$. Summing up over the cubes $\mathcal{K}\in\mathcal{N}_{\lambda,nd}=\mathcal{N}^1_{\lambda,nd}\cup\mathcal{N}^2_{\lambda,nd}$ then yields

$$\begin{aligned}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\leq\frac{6^q c_{nd}^q}{\lambda^q}&\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\ +\frac{6^q c_{nd}^q}{\varepsilon\lambda^q}&\sum_{\mathcal{K}\in\mathcal{N}^1_{\lambda,nd}}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\left(\frac{2^{-k(\mathcal{K})}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\ +\frac{6^q c_{nd}^q}{\varepsilon\lambda^q}&\sum_{\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}}\frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})}\left(\frac{2^{-k(\mathcal{K})}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.\end{aligned}\tag{5.72}$$

Observe that a key point in the previous inequality, due to the argument at the beginning of Step 3, is that terms involving integrals over $P_h\mathcal{K}$ appear on the right hand side if and only if $\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}$, for $h\in\{1,2\}$. By the symmetry of U and μ, by (5.53) and subsequent remarks, and yet using Proposition 5.2, we have that if $\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}$, then $\operatorname{Symm}(\mathcal{K})\in\mathcal{N}^1_{\lambda,nd}$ and vice-versa; moreover, again by Proposition 5.2 the following identity holds:

$$\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu=\int_{P_1\operatorname{Symm}(\mathcal{K})\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.$$

We therefore deduce that the last two terms in (5.72) coincide. Therefore, also recalling (5.64), inequality (5.72) can be rewritten for instance as

$$\begin{aligned}&\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\leq\frac{c}{\lambda^q}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\ &+\frac{c}{\varepsilon\lambda^q}\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\sum_{\mathcal{K}\in\mathcal{N}^1_{\lambda,nd}(\mathcal{H})}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\left(\frac{2^{-k(\mathcal{K})}}{\tilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\end{aligned}$$

for a constant c depending on n,α. To estimate the last term we make use of (5.67), and this yields

$$\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\leq\frac{c}{\lambda^q}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\frac{c}{\lambda^q}\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.$$

At this stage (5.63) follows observing that

$$\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu\leq 2\int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}}U^q\,d\mu\,.$$

This is turn true since the families $P\mathcal{N}_\lambda$ and $\mathcal{N}_{\lambda,nd}$ are made of mutually disjoint cubes and all their members are contained in $\mathcal{B}(x_0,s)$ (since these families are contained in Ξ and (5.22) holds). The proof of Lemma 5.6 is complete. $\square$

5.10. **Conclusion of the off-diagonal analysis.** We are now ready to prove the following lemma, which summarises the decomposition results in the off-diagonal case:

Lemma 5.7 (Off-diagonal level set inequality)**.** *The inequality*

$$\begin{aligned}\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \quad &\leq \quad 10^{n+2}\kappa^2\lambda^2\sum_{j\in J_D}\mu(\mathcal{B}_j)\\ &\qquad +c\lambda^{2-q}\int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q\,d\mu \end{aligned}\tag{5.73}$$

holds for a constant c depending only on n,α, while the number κ has been defined in (5.55) and exhibits the dependence displayed in (5.56).

Proof. We have that the decompositions in disjoint classes $\mathcal{U}_\lambda = \mathcal{U}_\lambda^d\cup\mathcal{U}_\lambda^{nd}$ and $\mathcal{U}_\lambda^{nd} = \mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,d}\cup\mathcal{N}_{\lambda,nd}$ and we recall that all the cubes from $\mathcal{U}_\lambda^{nd}$ are mutually disjoint. Moreover, by (5.38) and (5.57) it follows

$$\left(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^d}\mathcal{K}\right)\cup\left(\bigcup_{\mathcal{K}\in\mathcal{N}_{\lambda,d}}\mathcal{K}\right)\subset\bigcup_{j\in J_D}10\mathcal{B}_j\,.$$

Therefore

$$\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda}\mathcal{K}\subset\left(\bigcup_{j\in J_D}10\mathcal{B}_j\right)\cup\left(\bigcup_{\mathcal{K}\in\mathcal{M}_\lambda}\mathcal{K}\right)\cup\left(\bigcup_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mathcal{K}\right)$$

follows. Keeping this in mind and recalling (5.35), we start estimating

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu\leq\sum_j\int_{10\mathcal{B}_j\cap\{U>\lambda\}}U^2\,d\mu+\sum_{\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\lambda\}}U^2\,d\mu\,.$$

By (5.34) it follows that if $\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}\subset\mathcal{U}_\lambda^{nd}$, then

$$⨍_{\mathcal{K}} U^2\,d\mu\leq\frac{\mu(\widetilde{\mathcal{K}})}{\mu(\mathcal{K})}⨍_{\widetilde{\mathcal{K}}}U^2\,d\mu\leq\tilde{c}_d\lambda^2\,.$$

Note that we have used (5.32) since $\mathcal{K}\in\mathcal{U}_\lambda^{nd}$ implies by the definition in (5.39) that $\tilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}},P_2\widetilde{\mathcal{K}})\geq 2^{-k(\mathcal{K})}$. Therefore we conclude with

$$\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}\qquad\Longrightarrow\qquad\int_{\mathcal{K}\cap\{U>\lambda\}}U^2\,d\mu\leq\tilde{c}_d\lambda^2\mu(\mathcal{K})\,.$$

Using this last inequality together with (5.16) yields

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}}U^2\,d\mu\leq 10^{n+2\varepsilon}\kappa^2\lambda^2\sum_{j\in J_D}\mu(\mathcal{B}_j)+\tilde{c}_d\lambda^2\sum_{\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\,,$$

and (5.73) follows by just using Lemmas 5.4 and 5.6. $\square$

Remark 5.3. An interesting point of Lemma 5.7 is that it does not make use of the fact that u is a solution. All the estimates just rely on the fact that u belongs to the Sobolev space $W^{\alpha,2}$. This is ultimately linked to the fact that the analysis made in Sections 2-10 is made in a zone where the kernel of the operator, that is $|x-y|^{-(n+2\alpha)}$, is not very singular. The ultimate outcome is that the whole issue reduces now to estimate $\sum_j\mu(\mathcal{B}_j)$. Therefore, it remains to perform the analysis close to the diagonal and this will be done in the next section.

5.11. **Diagonal estimates.** Whenever $\mathcal{B}_j$ is a ball from the covering determined in (5.14)-(5.15), from (5.13) it follows that $\Psi_{H,M}(\mathcal{B}_j) \geq \kappa\lambda$. By the very definition of $\Psi_{H,M}(\cdot)$ in (5.1) it then follows that at least one of the following three inequalities must hold:

$$\left(\fint_{\mathcal{B}_j} U^2 d\mu \right)^{1/2} \geq \frac{\kappa\lambda}{3}, \tag{5.74}$$

$$\frac{H[\mu(\mathcal{B}_j)]^\eta}{\varepsilon^{1/2_*-1/2}} \left(\fint_{\mathcal{B}_j} F^{2_*} \, d\mu \right)^{1/2_*} \geq \frac{\kappa\lambda}{3}, \tag{5.75}$$

$$\frac{M[\mu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p-1/2}} \left(\fint_{\mathcal{B}_j} G^p \, d\mu \right)^{1/p} \geq \frac{\kappa\lambda}{3}. \tag{5.76}$$

Here κ has been defined in (5.55). We now examine the occurrence of each of three cases separately.

Occurrence of (5.74) (and estimate of the tail at the exit time). In the case (5.74) holds then using (4.15) we have

$$\begin{aligned} \kappa\lambda \ \leq\ & \frac{c}{\sigma\varepsilon^{1/q-1/2}} \left(\fint_{2\mathcal{B}_j} U^q d\mu \right)^{1/q} \\ & + \frac{\sigma}{\varepsilon^{1/q-1/2}} \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}_j} U^q \, d\mu \right)^{1/q} \\ & + \frac{c_1[\mu(\mathcal{B}_j)]^\eta}{\varepsilon^{1/2_*-1/2}} \left(\fint_{2\mathcal{B}_j} F^{2_*} \, d\mu \right)^{1/2_*} \\ & + \frac{c_2[\mu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p-1/2}} \sum_{k=1}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{2^k\mathcal{B}_j} G^p \, d\mu \right)^{1/p} \end{aligned} \tag{5.77}$$

for all $\sigma \in (0,1]$. The constants c_1, c depend only on n, α, Λ, while $c_2 := 3c_b$ and therefore it depends on $n, \alpha, \Lambda, \beta, \gamma, p$ and exhibits the behaviour described in (3.5). With $\mathcal{B}_j \equiv \mathcal{B}(x_j, \varrho(x_j))$ we determine the integer $m \geq 0$ such that

$$2^{-m}\varrho_0/2 \leq \varrho(x_j) < 2^{-m+1}\varrho_0/2. \tag{5.78}$$

Notice that since $\varrho(x_j) < (s-t)/40^n$, we have $m \geq 3$ and moreover $(s-t)/40^n \leq \varrho_0/40^n \leq 2^{m-1}\varrho(x_j)$, so that the definition (5.10) implies

$$\Upsilon_0(2^{m-1}\mathcal{B}_j) + \Upsilon_1(2^{m-1}\mathcal{B}_j) + \Upsilon_{2,M}(2^{m-1}\mathcal{B}_j) \leq \kappa\tilde{\lambda}_0. \tag{5.79}$$

On the other hand the terms indexed before m can be estimated using Hölder's inequality and the exit time condition in (5.13) as follows:

$$\left(\fint_{2^k\mathcal{B}_j} U^q \, d\mu \right)^{1/q} \leq \Psi_{H,M}(2^k\mathcal{B}_j) \leq \kappa\lambda \qquad \text{if } 1 \leq k \leq m-1.$$

By using the inequalities in the last two displays we then have

$$\begin{aligned} & \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}_j} U^q \, d\mu \right)^{1/q} \\ & \quad = \sum_{k=1}^{m-2} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}_j} U^q \, d\mu \right)^{1/q} \end{aligned}$$

$$
\begin{aligned}
&\quad +2^{-(m-1)(\alpha-\varepsilon)}\sum_{k=0}^{\infty}2^{-k(\alpha-\varepsilon)}\left(\fint_{2^{k+m-1}\mathcal{B}_j}U^q\,d\mu\right)^{1/q}\\
&\le \kappa\lambda\sum_{k=1}^{m-2}2^{-k(\alpha-\varepsilon)}+2^{-(m-1)(\alpha-\varepsilon)}\Upsilon_1(2^{m-1}\mathcal{B}_j)\\
&\le \kappa\lambda\sum_{k=1}^{m-2}2^{-k(\alpha-\varepsilon)}+2^{-(m-1)(\alpha-\varepsilon)}\kappa\tilde{\lambda}_0\\
&\le \kappa\lambda\sum_{k=1}^{\infty}2^{-k(\alpha-\varepsilon)}\le\frac{4\kappa\lambda}{\alpha-\varepsilon}\le\frac{8\kappa\lambda}{\alpha}\,,
\end{aligned}
$$

where we have used (2.2) and that $\varepsilon<\alpha/2$. In a completely similar way, again using (5.79), we have

$$
\begin{aligned}
\frac{c_2[\mu(\mathcal{B}_j)]^{\theta}}{\varepsilon^{1/p-1/2}}\sum_{k=1}^{\infty}2^{-k(2\beta-\gamma-2\varepsilon/p)}\left(\fint_{2^k\mathcal{B}_j}G^p\,d\mu\right)^{1/p} &\le \frac{4c_2\kappa\lambda}{(2\beta-\gamma-2\varepsilon/p)M}\\
&\le \frac{8c_2\kappa\lambda}{(2\beta-\gamma)M}\,,
\end{aligned}
$$

where we also used the upper bound on ε in (4.6). By (5.78) and the fact that $m\ge 3$ we gain that $2\varrho(x_j)\le\varrho_0/2$ so that (5.13) and Hölder's inequality yield

$$
\frac{c_1[\mu(\mathcal{B}_j)]^{\eta}}{\varepsilon^{1/2_*-1/2}}\left(\fint_{2\mathcal{B}_j}F^{2_*}\,d\mu\right)^{1/2_*}\le\frac{c_1\Psi_{H,M}(2\mathcal{B}_j)}{H}\le\frac{c_1\kappa\lambda}{H}\,.
$$

By merging the inequalities in the last three displays with (5.77) we obtain

$$
\kappa\lambda\le\frac{c}{\sigma\varepsilon^{1/q-1/2}}\left(\fint_{2\mathcal{B}_j}U^q d\mu\right)^{1/q}+\frac{\sigma}{\varepsilon^{1/q-1/2}}\frac{8\kappa\lambda}{\alpha}+\frac{c_1\kappa\lambda}{H}+\frac{8c_2\kappa\lambda}{(2\beta-\gamma)M}\,. \tag{5.80}
$$

We recall that up to now the parameters $H,M\ge 1$ in the definition in (5.1) have not yet been chosen as well as $\sigma\in(0,1)$. Taking hence

$$
\sigma:=\frac{\varepsilon^{1/q-1/2}\alpha}{56}\,,\qquad H:=6c_1\,,\qquad M:=\frac{56c_2}{2\beta-\gamma} \tag{5.81}
$$

and reabsorbing terms in (5.80) we conclude with

$$
\kappa\lambda\le\frac{c}{\varepsilon^{2/q-1}}\left(\fint_{2\mathcal{B}_j}U^q d\mu\right)^{1/q}\Longrightarrow\mu(\mathcal{B}_j)\le\frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j}U^q d\mu\,,
$$

where c depends on n,α,Λ. Now, select a number $\kappa_3>0$, also using (4.2) we estimate

$$
\begin{aligned}
&\frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j}U^q d\mu\\
&\quad\le\frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U\le\kappa_3\kappa\lambda\}}U^q d\mu+\frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}}U^q d\mu\\
&\quad\le\frac{\tilde{c}\mu(\mathcal{B}_j)\kappa_3^q}{\varepsilon^{2-q}}+\frac{\tilde{c}}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}}U^q d\mu\,,
\end{aligned} \tag{5.82}
$$

again for $\tilde{c}$ depending only n,α,Λ. By choosing

$$
\kappa_3\le\left(\frac{\varepsilon^{2-q}}{2\tilde{c}}\right)^{1/q}\,, \tag{5.83}
$$

we arrive at

$$\mu(\mathcal{B}_j) \le \frac{c_3}{(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}} U^q d\mu \qquad \text{where} \qquad c_3 := \frac{2\tilde{c}}{\varepsilon^{2-q}} \tag{5.84}$$

and $\tilde{c}$ is independent of ε but only depends on n, α, Λ.

Occurrence of (5.75)-(5.76). In case of (5.75), we have

$$\left(\frac{\kappa\lambda}{3}\right)^{2_*} \le \frac{H^{2_*}[\mu(\mathcal{B}_j)]^{2_*\eta-1}}{\varepsilon^{1-2_*/2}}\int_{\mathcal{B}_j} F^{2_*}\, d\mu$$

which readily implies

$$\mu(\mathcal{B}_j) \le \left(\frac{3H}{\varepsilon^{1/2_*-1/2}\kappa\lambda}\right)^{2_*/(1-2_*\eta)} \left(\int_{\mathcal{B}_j} F^{2_*}\, d\mu\right)^{1/(1-2_*\eta)} .$$

Observe that by the definitions given in (4.16) we have that $2_*\eta < 1/2$. With $\kappa_4 \in (0,1)$ being a positive number to be chosen in a few lines, we further split the support of the integral of the right hand side integral as already done in (5.82):

$$\begin{aligned} \left(\int_{\mathcal{B}_j} F^{2_*}\, d\mu\right)^{1/(1-2_*\eta)} &\le \left[\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\, d\mu + (\kappa_4\kappa\lambda)^{2_*}\mu(\mathcal{B}_j)\right]^{1/(1-2_*\eta)} \\ &\le 2^{2_*\eta/(1-2_*\eta)}\left(\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\, d\mu\right)^{1/(1-2_*\eta)} \\ &\qquad +[2(L+1)]^{2_*/(1-2_*\eta)}(\kappa_4\kappa\lambda)^{2_*/(1-2_*\eta)}\mu(\mathcal{B}_j) . \end{aligned}$$

Observe that in view of $\mathcal{B}_j \subset \mathcal{B}(x_0, 2\varrho_0)$ and (5.18), we have estimated

$$[\mu(\mathcal{B}_j)]^{1/(1-2_*\eta)} \le [\mu(\mathcal{B}(x_0,2\varrho_0))]^{2_*\eta/(1-2_*\eta)}\mu(\mathcal{B}_j) \le L^{2_*\eta/(1-2_*\eta)}\mu(\mathcal{B}_j) . \tag{5.85}$$

We now take $\kappa_4 \in (0,1)$ in order to satisfy

$$\left[\frac{6H(L+1)\kappa_4}{\varepsilon^{1/2_*-1/2}}\right]^{2_*/(1-2_*\eta)} \le \frac{1}{2} \Longrightarrow \kappa_4 \le \left(\frac{1}{2}\right)^{(1-2_*\eta)/2_*} \frac{\varepsilon^{1/2_*-1/2}}{6H(L+1)} . \tag{5.86}$$

Using this choice and combining the content of the last four displays (and recalling that $2_*\eta/(1-2_*\eta) \le 1$) then yields that

$$\mu(\mathcal{B}_j) \le 4\left(\frac{3H}{\varepsilon^{1/2_*-1/2}\kappa\lambda}\right)^{2_*/(1-2_*\eta)} \left(\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\, d\mu\right)^{1/(1-2_*\eta)} .$$

Now, by means of (5.78)-(5.79), we have

$$\begin{aligned} \int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\, d\mu &\le (\kappa_4\kappa\lambda)^{2_*}\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}}\left(\frac{F}{\kappa_4\kappa\lambda}\right)^{2_*+\delta_f} d\mu \\ &\le \frac{\mu(2^{m-1}\mathcal{B}_j)}{(\kappa_4\kappa\lambda)^{\delta_f}} \fint_{2^{m-1}\mathcal{B}_j} F^{2_*+\delta_f}\, d\mu \\ &\le \frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{(\kappa_4\kappa\lambda)^{\delta_f}}[\Upsilon_0(2^{m-1}\mathcal{B}_j)]^{2_*+\delta_f} \le \frac{L\tilde{\lambda}_0^{2_*+\delta_f}}{(\kappa_4\kappa\lambda)^{\delta_f}} \end{aligned} \tag{5.87}$$

and hence

$$\mu(\mathcal{B}_j) \le \frac{c_4\tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}}\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\, d\mu \tag{5.88}$$

where

$$c_4 := 4\left[\frac{3H(L+1)}{\varepsilon^{1/2_*-1/2}}\right]^{2_*/(1-2_*\eta)} \tag{5.89}$$

and H has been defined in (5.81). A similar argument can be used in case that (5.76) holds. Specifically, we have

$$\mu(\mathcal{B}_j) \leq \left(\frac{3M}{\varepsilon^{1/p-1/2}\kappa\lambda}\right)^{p/(1-p\theta)} \left(\int_{\mathcal{B}_j} G^p \, d\mu\right)^{1/(1-p\theta)}$$

and then

$$\begin{aligned}\left(\int_{\mathcal{B}_j} G^p \, d\mu\right)^{1/(1-p\theta)} \quad &\leq \quad 2^{p\theta/(1-p\theta)} \left(\int_{\mathcal{B}_j \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu\right)^{1/(1-p\theta)} \\ &\quad + [2(L+1)]^{p/(1-p\theta)} (\kappa_5\kappa\lambda)^{p/(1-p\theta)} \mu(\mathcal{B}_j)\,.\end{aligned}$$

This time we select number $\kappa_5 \in (0,1)$ such that

$$\kappa_5 \leq \left(\frac{1}{2}\right)^{(1-p\theta)/p} \frac{\varepsilon^{1/p-1/2}}{6M(L+1)} \tag{5.90}$$

and recall Remark 4.1 in order to get

$$\mu(\mathcal{B}_j) \leq 2^{\Lambda_\theta+1} \left(\frac{3M}{\varepsilon^{1/p-1/2}\kappa\lambda}\right)^{p/(1-p\theta)} \left(\int_{\mathcal{B}_j \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu\right)^{1/(1-p\theta)} .$$

We then estimate as in (5.87) thereby obtaining

$$\int_{\mathcal{B}_j \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu \leq \frac{\mu(\mathcal{B}(x_0, 2\varrho_0))}{(\kappa_5\kappa\lambda)^{\delta_g}} [\Upsilon_0(2^{m-1}\mathcal{B}_j)]^{p+\delta_g} \leq \frac{L\tilde{\lambda}_0^{p+\delta_g}}{(\kappa_5\kappa\lambda)^{\delta_g}} \,.$$

and we conclude with

$$\mu(\mathcal{B}_j) \leq \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}_j \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu \,, \tag{5.91}$$

where

$$c_5 := 2^{\Lambda_\theta+1} \left[\frac{3M(L+1)}{\varepsilon^{1/p-1/2}}\right]^{p/(1-p\theta)} . \tag{5.92}$$

All in all, taking (5.84), (5.88) and (5.91) into account we obtain

$$\begin{aligned}\mu(\mathcal{B}_j) \quad \leq \quad & \frac{c_3}{(\kappa\lambda)^q} \int_{2\mathcal{B}_j \cap \{U > \kappa_3 \kappa\lambda\}} U^q d\mu \\ & + \frac{c_4 \tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}} \int_{\mathcal{B}_j \cap \{F > \kappa_4 \kappa\lambda\}} F^{2_*} \, d\mu \\ & + \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}_j \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu \,.\end{aligned}$$

Since $\{2\mathcal{B}_j\}_j$ is a disjoint family and all members belong to $\mathcal{B}(x_0, s)$, we have that

$$\begin{aligned}\sum_{j \in J_D} \mu(\mathcal{B}_j) \quad \leq \quad & \frac{c_3}{(\kappa\lambda)^q} \int_{\mathcal{B}(x_0,s) \cap \{U > \kappa_3 \kappa\lambda\}} U^q d\mu \\ & + \frac{c_4 \tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}} \int_{\mathcal{B}(x_0,s) \cap \{F > \kappa_4 \kappa\lambda\}} F^{2_*} \, d\mu \\ & + \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}(x_0,s) \cap \{G > \kappa_5 \kappa\lambda\}} G^p \, d\mu \,.\end{aligned} \tag{5.93}$$

The constants c_3, c_4, c_5 have been defined in (5.84), (5.89) and (5.92), respectively, while the numbers $\kappa, \kappa_3, \kappa_4, \kappa_5 \in (0,1)$ must be taken in order to satisfy (5.55), (5.83), (5.86) and (5.90), respectively.

5.12. **Conclusion of the proof.** We start combining (5.73) and (5.93). Employing the elementary estimation

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\kappa_3\kappa\lambda\}} U^2\,d\mu \le \lambda^{2-q}\int_{\mathcal{B}(x_0,t)\cap\{U>\kappa_3\kappa\lambda\}} U^q\,d\mu + \int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu\,,$$

(5.73) and (5.93) yield, after a few elementary manipulations, the following estimate:

$$\begin{aligned}\int_{\mathcal{B}(x_0,t)\cap\{U>\kappa_3\kappa\lambda\}} U^2\,d\mu \quad &\le \quad \frac{c}{(\kappa_3\kappa)^{2-q}}(\kappa_3\kappa\lambda)^{2-q}\int_{\mathcal{B}(x_0,s)\cap\{U>\kappa_3\kappa\lambda\}} U^q d\mu\\ &\quad +\frac{c_4\tilde\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\kappa_4^2(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\,d\mu\\ &\quad +\frac{c_5\tilde\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\kappa_5^2(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{G>\kappa_5\kappa\lambda\}} G^p\,d\mu\,.\end{aligned}\tag{5.94}$$

The constant c appearing above depends on n,α,Λ, but is still independent of ε, and we have also used the fact that $\kappa,\kappa_3\in(0,1)$. We can therefore reformulate estimate (5.94) as follows:

$$\begin{aligned}\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \quad &\le \quad \frac{c\lambda^{2-q}}{(\kappa_3\kappa\varepsilon)^{2-q}}\int_{\mathcal{B}(x_0,s)\cap\{U>\lambda\}} U^q d\mu\\ &\quad +\frac{c_6\tilde\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{F>\kappa_4\lambda/\kappa_3\}} F^{2_*}\,d\mu\\ &\quad +\frac{c_7\tilde\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{G>\kappa_5\lambda/\kappa_3\}} G^p\,d\mu\,.\end{aligned}\tag{5.95}$$

The constant $c\equiv c(n,\alpha,\Lambda)$ is independent of ε, while $c_6\equiv c_6(n,\alpha,\Lambda,L,\varepsilon)$ and $c_7\equiv c_7(n,\alpha,\Lambda,\beta,\gamma,p,L,\varepsilon)$; the constant c_7 exhibits a blow-up behaviour with respect to p as described in (3.5). Since estimate (5.94) holds for $\lambda\ge\lambda_1$ - and λ_1 has been defined in (5.33) - we have that (5.95) holds whenever $\lambda\ge\kappa\kappa_3\lambda_1$. We remark that the previous inequality hold for a choice of $\kappa,\kappa_3,\kappa_4,\kappa_5\in(0,1)$ that satisfy (5.55), (5.83), (5.86) and (5.90), respectively. In order to conclude with (5.7) we now need to estimate a few constants. We are primarily interested in an explicit dependence on ε in the second integral appearing in (5.95). We therefore look at (5.55) and (5.83) and we infer we can in fact choose κ,κ_3 in order to have

$$\kappa_3\kappa\approx\frac{\varepsilon^{3/q-1}}{c_*}\,,\tag{5.96}$$

for a constant c_* which is now independent of ε, but just depends on n,α,Λ. We next find an upper bound for the numbers $\tilde\lambda_0$ and λ_1 introduced in (5.10) and (5.33), respectively; this will allow to verify estimate (5.7) in the range dictated by (5.8). Let us notice that if $x\in B(x_0,t)$ and $(s-t)/40^n\le\varrho\le\varrho_0/2$, then $\mathcal{B}(x,\varrho)\subset\mathcal{B}(x_0,2\varrho_0)$. Therefore, recalling (4.2), whenever $\tilde U$ is a μ-integrable function we can estimate

$$\begin{aligned}\fint_{\mathcal{B}(x,\varrho)}\tilde U\,d\mu \quad &\le \quad \frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{\mu(\mathcal{B}(x,\varrho))}\fint_{\mathcal{B}(x_0,2\varrho_0)}\tilde U\,d\mu\\ &\le \quad c\left(\frac{\varrho_0}{s-t}\right)^{n+2\varepsilon}\fint_{\mathcal{B}(x_0,2\varrho_0)}\tilde U\,d\mu\end{aligned}\tag{5.97}$$

for a constant c depending on n but independent of ε. Applying the inequality in the last display to U^2,G^p, F^{2_*}, $G^{p+\delta_g}$ and $F^{2_*+\delta_f}$ - and eventually on different balls $2^k\mathcal{B}(x,\varrho)\subset 2^k\mathcal{B}(x_0,2\varrho_0)$ - yields

$$
\begin{aligned}
&\kappa^{-1}\left\{\Psi_{H,M}(x,\varrho)+\Upsilon_0(x,\varrho)+\Upsilon_1(x,\varrho)+\Upsilon_{2,M}(x,\varrho)\right\}\\
&\leq \frac{c}{\varepsilon^{1/q}}\left(\frac{\varrho_0}{s-t}\right)^{n+2\varepsilon}\left\{\Psi_{H,M}(x_0,2\varrho_0)+\Upsilon_0(x_0,2\varrho_0)+\Upsilon_1(x_0,2\varrho_0)\right.\\
&\qquad\qquad\qquad\left.+\Upsilon_{2,M}(x_0,2\varrho_0)\right\}\\
&\leq \frac{c}{\varepsilon^{1/q}}\left(\frac{\varrho_0}{s-t}\right)^{n+2\varepsilon}\mathrm{ADD}(x_0,2\varrho_0)\,. \qquad (5.98)
\end{aligned}
$$

In order, we have also used (5.56), (5.81) to get rid of the presence of M and H and that $\varrho_0/(s-t)$ is bounded away from zero. We recall that the functional $\mathrm{ADD}(\cdot)$ has been introduced in (5.5). We now obtain an upper bound for λ_1 defined in (5.33). The quantity appearing on the right hand side of (5.98) provides an upper bound on $\tilde{\lambda}_0$. In a similar way, if $\mathcal{K}=K_1\times K_2\in\Xi_{k_0}$, with k_0 as in (5.19), then $\mathcal{K}\subset\mathcal{B}(x_0,s)\subset\mathcal{B}(x_0,2\varrho_0)$ and therefore we have

$$
\mu(\mathcal{K})\geq\frac{c}{\varrho_0^{n-2\varepsilon}}\int_{K_1}\int_{K_2}dx\,dy=\frac{c(s-t)^{2n}}{\varrho_0^{n-2\varepsilon}}\,.
$$

Hence, as for (5.97), we have

$$
\fint_{\mathcal{K}}\tilde{U}\,d\mu\leq\frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{\mu(\mathcal{K})}\fint_{\mathcal{B}(x_0,2\varrho_0)}\tilde{U}\,d\mu\leq\frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n}\fint_{\mathcal{B}(x_0,2\varrho_0)}\tilde{U}\,d\mu\,. \qquad (5.99)
$$

By using (5.98)-(5.99), and recalling that $\varepsilon<1$, we get

$$
\lambda_1\leq\frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n}\mathrm{ADD}(x_0,2\varrho_0)
$$

where c depends only on $n,\alpha,\Lambda,\beta,p,\gamma,\varepsilon$. Summarizing the content of the above manipulations we can finally arrive at (5.7), with the restriction on λ described in (5.8). Specifically, we use (5.96) to estimate the constant in front of the second integral appearing in (5.95), and the bounds found for $\tilde{\lambda}_0$ and λ_1 to conclude with the admissible range of values $\lambda\geq\lambda_0$ described via (5.8). Needless to say, we are taking $\kappa_f:=\kappa_4/\kappa_3$ and $\kappa_g:=\kappa_5/\kappa_3$.

6. Self-improving inequalities

This section is dedicated to the proof of a fractional reverse Hölder type inequality on diagonal balls with increasing supports, that is the estimate (6.1) below. This will eventually imply Theorem 1.1 at the end of the section.

Theorem 6.1 (Reverse Hölder type inequality)**.** *Let $u\in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to (1.14) under the assumptions of Theorem 1.1, in particular, (3.1) and (3.3) are in force. Define the functions U,F and G as in (4.5). Then there exist positive constants $\varepsilon\in(0,1-\alpha)$, $\delta\in(0,1)$ and $c_8\geq 1$, depending on $n,\alpha,\Lambda,\beta,p,\gamma,\delta_1$, such that the following inequality holds whenever $\mathcal{B}\equiv\mathcal{B}(x_0,\varrho_0)\subset\mathbb{R}^{2n}$:*

$$
\begin{aligned}
\left(\fint_{\mathcal{B}}U^{2+\delta}\,d\mu\right)^{1/(2+\delta)} \leq\ & c_8\sum_{k=1}^{\infty}2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}}U^2\,d\mu\right)^{1/2}\\
&+c_8\varrho_0^{\alpha-\varepsilon}\left(\fint_{2\mathcal{B}}F^{2_*+\delta_0}\,d\mu\right)^{1/(2_*+\delta_0)}\\
&+c_8\varrho_0^{\gamma-2\beta+\alpha+\varepsilon(2/p-1)}\left(\fint_{2\mathcal{B}}G^{p(1+\delta_1)}\,d\mu\right)^{1/[p(1+\delta_1)]}
\end{aligned}
$$

$$+c_8 \varrho_0^{\gamma-2\beta+\alpha+\varepsilon(2/p-1)} \sum_{k=1}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \left(\fint_{2^k\mathcal{B}} G^p \, d\mu \right)^{1/p} . \tag{6.1}$$

All the terms on the right hand side of the previous inequality are finite.

Proof. Step 1: Determining the exponents. Let us observe that, whenever $\varepsilon \in (0, \alpha/2)$, we have

$$\frac{8\varepsilon}{n+2\varepsilon} < \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)} .$$

Therefore, we can always find two positive numbers $\varepsilon \in (0, \alpha/2)$ and $\delta_f > 0$, satisfying (4.6) and $\delta_f \leq \delta_0$, respectively, such that

$$\frac{8\varepsilon}{n+2\varepsilon} < \delta_f \leq \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)} \qquad \text{and} \qquad \varepsilon < 1-\alpha . \tag{6.2}$$

We recall that $F \in L^{2_*+\delta_f}_{\rm loc}(\mathbb{R}^n; \mu)$ by (4.7). Next, we determine the positive number $\delta > 0$ by imposing different restrictions on it; we indeed start assuming that

$$\delta \leq \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} \qquad \text{and} \qquad \delta \leq \frac{(\gamma-2\beta+\alpha)\delta_g}{4n} . \tag{6.3}$$

Let us briefly discuss a few consequences of the two conditions above, starting by the first one. Specifically, we start showing that

$$\delta \leq \delta_f \left[\frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)} \right] - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} \tag{6.4}$$

holds. Indeed, using the first inequality in (6.3), we have

$$\begin{aligned} \delta \leq \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} \quad &= \quad \frac{8\varepsilon}{n+2\varepsilon} \frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)} - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} \\ &\leq \quad \delta_f \left[\frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)} \right] - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} . \end{aligned}$$

Next, the definition in (4.16) and the fact that $\varepsilon < \alpha/2$ gives that $1 > \theta > (\gamma - 2\beta+\alpha)/(n+\alpha)$. Then, the fact that the function $t \to t/(1-t)$ is increasing in the interval $(0,1)$, allows to estimate

$$\frac{\gamma-2\beta+\alpha}{2n} \leq \frac{\gamma-2\beta+\alpha}{n-\gamma+2\beta} \leq \frac{\theta}{1-\theta} < \frac{p\theta}{1-p\theta}$$

so that from the second inequality in (6.3), it follows that

$$\delta < \frac{(\gamma-2\beta+\alpha)\delta_g}{4n} \leq \frac{\delta_g}{2} \frac{p\theta}{1-p\theta} . \tag{6.5}$$

Finally, for $t \in (0,1)$, we define the function

$$S(t) := \frac{2c_s(n+4)}{4\alpha t^6} \geq \frac{2c_s}{(2-q)t^{3(2-q)/q}} , \tag{6.6}$$

where c_s is the constant introduced in Proposition 5.1 and q has been introduced in (4.13); in the last estimation we have used that $\varepsilon \in (0, \alpha/2)$. We then impose the last restriction on δ, that is

$$\delta S(\varepsilon) \leq 1/4 . \tag{6.7}$$

All in all, the choices made in (6.3) and (6.7), allow to determine δ as a positive number depending only on $n, \alpha, \Lambda, \beta, p, \gamma, \delta_1$, as required in the statement of Theorem 6.1. In the subsequent Step 2, by applying Proposition 5.1 with the above choice of the numbers $\varepsilon, \delta, \delta_f$, we are going to prove that $U \in L^{2+\delta}_{\rm loc}(\mathbb{R}^{2n}; \mu)$.

Step 2: Reverse Hölder type inequalities. The finiteness of the terms on the right hand hand side of has already been discussed in Section 4.3. First of all, we show that we can reduce to the case $\varrho_0 = 1$ and $\mathcal{B} = B(0,1) \times B(0,1)$; this eventually allows to apply Proposition 5.1. Indeed, notice that the rescaled functions

$$\tilde{u}(x) := u(x_0 + \varrho_0 x)\,, \quad \tilde{g}(x) := \varrho_0^{2\alpha-2\beta} g(x_0 + \varrho_0 x)\,, \quad \tilde{f}(x) := \varrho_0^{2\alpha} f(x_0 + \varrho_0 x)$$

still solve equation (1.14). Therefore, applying (6.1) in this case and in $B(0,1) \times B(0,1)$, and scaling back to the original functions and to the original diagonal ball $\mathcal{B}$, leads to establish (6.1) in the general case. We now pass to the proof of (6.1) when $\mathcal{B} = B(0,1) \times B(0,1)$. We define the truncated function $U_m := \min\{U, m\}$ for m being a positive integer, and the measure $d\nu = U^2 d\mu$. Moreover, we abbreviate the notation $\mathcal{B}_s := \mathcal{B}(0,s)$. With the aim of applying Proposition 5.1, we then consider balls $\mathcal{B} \equiv \mathcal{B}_1 \subset \mathcal{B}_t \subset \mathcal{B}_s \subset \mathcal{B}_2$ as in (5.6), while λ_0 is accordingly defined as in (5.8). We shall derive uniform higher integrability for the functions U_m and will recover the final result by letting $m \to \infty$. With $\delta \in (0,1)$ being the number determined in Step 1, by Cavalieri's principle we have that

$$\begin{aligned}
\int_{\mathcal{B}_t} U_m^\delta U^2\, d\mu &= \int_{\mathcal{B}_t} U_m^\delta\, d\nu \\
&= \delta \int_0^\infty \lambda^{\delta-1} \nu(\mathcal{B}_t \cap \{U_m > \lambda\})\, d\lambda \\
&= \delta \int_0^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda \\
&\leq \lambda_0^\delta \int_{\mathcal{B}_t} U^2\, d\mu + \delta \int_{\lambda_0}^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda\,. \qquad (6.8)
\end{aligned}$$

The second-last integral appearing in the above display can be easily estimated by recalling the identity of λ_0 in (5.8) and that $\varrho_0/(s-t) \geq 1$ and using (4.2):

$$\lambda_0^\delta \int_{\mathcal{B}_t} U^2\, d\mu \leq \mu(\mathcal{B}_2) \lambda_0^\delta \fint_{2\mathcal{B}} U^2\, d\mu \leq c\mu(\mathcal{B}_1) \lambda_0^{2+\delta}\,. \qquad (6.9)$$

We proceed with the remaining term in (6.8); using (5.7) we gain

$$\begin{aligned}
&\delta \int_{\lambda_0}^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda \\
&\leq \frac{c_s \delta}{\varepsilon^{3(2-q)/q}} \int_{\lambda_0}^m \lambda^{\delta+1-q} \int_{\mathcal{B}_s \cap \{U > \lambda\}} U^q\, d\mu\, d\lambda \\
&\quad + c_f \delta \int_{\lambda_0}^m \frac{\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)-1-\delta}} \int_{\mathcal{B}_s \cap \{F > \kappa_f \lambda\}} F^{2_*}\, d\mu\, d\lambda \\
&\quad + c_g \delta \int_{\lambda_0}^m \frac{\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)-1-\delta}} \int_{\mathcal{B}_s \cap \{G > \kappa_g \lambda\}} G^p\, d\mu\, d\lambda \\
&=: \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3\,. \qquad (6.10)
\end{aligned}$$

Using (6.6)-(6.7) and Fubini's theorem, we get

$$\begin{aligned}
\mathcal{J}_1 &\leq \frac{c_s \delta}{\varepsilon^{3(2-q)/q}} \int_0^\infty \lambda^{\delta+1-q} \int_{\mathcal{B}_s \cap \{U_m > \lambda\}} U^q\, d\mu\, d\lambda \\
&= \frac{c_s \delta}{(\delta+2-q)\varepsilon^{3(2-q)/q}} \int_{\mathcal{B}_s} U_m^{2+\delta-q} U^q\, d\mu \\
&\leq \delta S(\varepsilon) \int_{\mathcal{B}_s} U_m^\delta U^2\, d\mu \\
&\leq \frac{1}{4} \int_{\mathcal{B}_s} U_m^\delta U^2\, d\mu\,. \qquad (6.11)
\end{aligned}$$

We next estimate $\mathcal{J}_2$. Changing variables, using Fubini's theorem, and recalling the dependence $\kappa_f \equiv \kappa_f(n,\alpha,\Lambda,\varepsilon)$, we have

$$\begin{aligned}
&\int_{\lambda_0}^{m} \lambda^{\delta+1-(1+\eta\delta_f)2_*/(1-2_*\eta)} \int_{\mathcal{B}_s\cap\{F>\kappa_f\lambda\}} F^{2_*}\,d\mu\,d\lambda \\
&\quad\le c\int_0^{\infty} \lambda^{\delta+1-(1+\eta\delta_f)2_*/(1-2_*\eta)} \int_{\mathcal{B}_s\cap\{F>\lambda\}} F^{2_*}\,d\mu\,d\lambda \\
&\quad= \frac{c\mu(\mathcal{B}_2)}{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)} \fint_{\mathcal{B}_2} F^{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}\,d\mu \\
&\quad\le \frac{c\mu(\mathcal{B}_2)}{\delta} \fint_{\mathcal{B}_2} F^{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}\,d\mu\,, \qquad (6.12)
\end{aligned}$$

again for a constant depending on n,α,Λ and ε. In writing the last inequality we have used that (6.2) is in force and the fact that

$$\delta_f \le \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)} \Longleftrightarrow 2-\frac{(1+\eta\delta_f)2_*}{1-2_*\eta}\ge 0\,.$$

The last integral appearing in (6.12) is finite if $\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*\le 2_*+\delta_f$, and a lengthy computation shows that this is equivalent to (6.4). Therefore, using Hölder's inequality, we can estimate

$$\begin{aligned}
\mathcal{J}_2 &\le c\mu(\mathcal{B}_2)\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}\left(\fint_{\mathcal{B}_2} F^{2_*+\delta_f}\,d\mu\right)^{\frac{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}{2_*+\delta_f}} \\
&\le c\mu(\mathcal{B}_1)\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)+\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*} \\
&= c\mu(\mathcal{B}_1)\lambda_0^{2+\delta}\,, \qquad (6.13)
\end{aligned}$$

where c depends only on n,α,Λ and ε. We finally come to the estimation of $\mathcal{J}_3$. For this we notice that the definitions of p and θ give, independently of ε, that

$$p \ge \frac{2n}{n+2(\gamma-2\beta+\alpha)} \Longleftrightarrow \frac{p}{1-p\theta}\ge 2 \qquad (6.14)$$

and then, recalling that $\kappa_g \equiv \kappa_g(n,\alpha,\Lambda,\varepsilon,\gamma,\beta,p)$, we have

$$\begin{aligned}
&\int_{\lambda_0}^{m} \lambda^{\delta+1-(1+\theta\delta_g)p/(1-p\theta)} \int_{\mathcal{B}_s\cap\{G>\kappa_g\lambda\}} G^{p}\,d\mu\,d\lambda \\
&\quad\le \int_{\lambda_0}^{\infty} \lambda^{\delta+1-(1+\theta\delta_g)p/(1-p\theta)}\,d\lambda \int_{\mathcal{B}_s} G^p\,d\mu \\
&\quad\le \frac{c\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)}\mu(\mathcal{B}_2)}{(1+\theta\delta_g)p/(1-p\theta)-\delta-2} \fint_{\mathcal{B}_2} G^p\,d\mu \\
&\quad\le \frac{c\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)}\mu(\mathcal{B}_2)}{\theta\delta_g p/(1-p\theta)-\delta}\left(\fint_{\mathcal{B}_2} G^{p+\delta_g}\,d\mu\right)^{p/(p+\delta_g)} \\
&\quad\le \frac{c}{\delta}\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)+p}\mu(\mathcal{B}_2)\,.
\end{aligned}$$

Observe that in order to perform the last two estimations we have also used (6.14) and (6.5), respectively. Therefore we can estimate as in (6.13), that is

$$\mathcal{J}_3 \le c\mu(\mathcal{B}_2)\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)+\delta+2-(1+\theta\delta_g)p/(1-p\theta)+p} = c\mu(\mathcal{B}_2)\lambda_0^{2+\delta} \qquad (6.15)$$

with $c\equiv c(n,\alpha,\Lambda,\varepsilon,\gamma,\beta,p)$. Connecting (6.11), (6.13) and (6.15) to (6.10), and combining the resulting inequality with (6.8) and (6.9), we get

$$\int_{\mathcal{B}_t} U_m^{\delta}U^2\,d\mu \le \frac{1}{4}\int_{\mathcal{B}_s} U_m^{\delta}U^2\,d\mu + c\mu(\mathcal{B}_1)\lambda_0^{2+\delta}\,.$$

By recalling the identity of λ_0 in (5.8), and using several times the doubling property of μ, after a few elementary manipulations we come to

$$\left(\fint_{\mathcal{B}_t} U_m^\delta U^2\,d\mu\right)^{1/(2+\delta)} \leq \frac{1}{2}\left(\fint_{\mathcal{B}_s} U_m^\delta U^2\,d\mu\right)^{1/(2+\delta)} + \frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n}\mathrm{ADD}(2\mathcal{B})\,.$$

We can therefore rewrite the above inequality as

$$\phi(t) \leq \frac{1}{2}\phi(s) + \frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n}\mathrm{ADD}(2\mathcal{B})$$

for a constant $c \equiv c(n,\alpha,\Lambda,\varepsilon,\gamma,\beta,p)$ which is still independent of $m \in \mathbb{N}$, and where, obviously, we have set

$$\phi(\varrho) := \left(\fint_{\mathcal{B}_\varrho} U_m^\delta U^2\,d\mu\right)^{1/(2+\delta)}$$

for $\varrho \in [\varrho_0, (3/2)\varrho_0]$. We are therefore in position to apply the standard iteration Lemma 6.1 below, that gives, after returning to the full notation:

$$\left(\fint_{\mathcal{B}} U_m^\delta U^2\,d\mu\right)^{1/(2+\delta)} \leq c\mathrm{ADD}(2\mathcal{B})\,.$$

The previous inequality holds for a constant $c \equiv c(n,\alpha,\Lambda,\varepsilon,\gamma,\beta,p)$ which is independent of $m \in \mathbb{N}$. Therefore letting $m \to \infty$ yields

$$\left(\fint_{\mathcal{B}} U^{2+\delta}\,d\mu\right)^{1/(2+\delta)} \leq c\mathrm{ADD}(2\mathcal{B})\,.$$

At this point (6.1) follows by recalling the definition of $\mathrm{ADD}(2\mathcal{B})$ in (5.5) and using a few elementary manipulations involving Hölder's inequality. In particular, we use the fact that $2_* + \delta_f \leq 2_* + \delta_0$ and $p + \delta_g \leq p(1+\delta_1)$; see Lemma 4.1. □

Lemma 6.1. *Let* $\phi\colon [\varrho_0, 3\varrho_0/2] \to [0,\infty)$ *be a function such that*

$$\phi(t) \leq \frac{1}{2}\phi(s) + \frac{A}{(s-t)^\gamma}$$

holds whenever $\varrho_0 < t < s < (3/2)\varrho_0$*, where* A *and* γ *are positive constants. Then the inequality*

$$\phi(\varrho_0) \leq \frac{cA}{\varrho_0^\gamma}\,.$$

holds for a constant $c \equiv c(\gamma)$*.*

For a proof of the previous lemma see for instance [17, Chapter 6].

Proof of Theorem 1.1. The proof is now a simple consequence of Theorem 6.1, that gives that $U \in L^{2+\delta}(\mathcal{B};\mu)$ whenever $\mathcal{B} = B \times B$ and $B \subset \mathbb{R}^n$ is a ball (that for simplicity we take to be centred at the origin). We now translate this information in terms of fractional norms of the original function u. In fact this means that, whenever $B \subset \mathbb{R}^n$ is a ball centred at the origin, we have

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

Re-writing the last integral we find

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

whenever $B \subset \mathbb{R}^n$ is a ball, and this means that $u \in W^{\alpha+\varepsilon\delta/(2+\delta),2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$; observe that since $\varepsilon < 1-\alpha$ then $\alpha+\varepsilon\delta/(2+\delta) < 1$. We have therefore improved the regularity of u both in the fractional and in the differentiability scale, and Theorem 1.1 follows by suitably renaming (via embedding theorems) the number δ considered in its statement. □

Proof of Theorem 1.2. The proof is just a consequence of the arguments developed to prove Theorem 6.1. In fact the only thing needed there is Proposition 4.2, whose content is now considered as an assumption in (1.23), provided we are taking $F = G = 0$; the rest of the argument then remains unchanged. □

References

[1] Bass R. F. & Kassmann M.: Hölder continuity of harmonic functions with respect to operators of variable order. *Comm. PDE* 30 (2005), 1249–1259.
[2] Bass R. F. & Ren H.: Meyers inequality and strong stability for stable-like operators. *J. Funct. Anal.* 265 (2013), 28–48.
[3] Bjorland C & Caffarelli L. & Figalli A.: Non-local gradient dependent operators. *Adv. Math.* 230 (2012), 1859–1894.
[4] Bojarski B. & Iwaniec T.: Analytical foundations of the theory of quasiconformal mappings in $\mathbb{R}^n$. *Ann. Acad. Sci. Fenn. Ser. A I Math.* 8 (1983), 257–324.
[5] Cabré X. & Cinti E.: Sharp energy estimates for nonlinear fractional diffusion equations. *Calc. Var. & PDE*, to appear.
[6] Cabré X. & Roquejoffre J.M. The influence of fractional diffusion in Fisher-KPP equations. *Comm. Math. Phys.* 320 (2013), 679–722.
[7] Caffarelli L. & Chan C. H. & Vasseur A.: Regularity theory for parabolic nonlinear integral operators. *J. Amer. Math. Soc.* 24 (2011), 849–869.
[8] Caffarelli L. & Silvestre L.: The Evans-Krylov theorem for nonlocal fully nonlinear equations. *Ann. of Math. (2)* 174 (2011), 1163–1187.
[9] Di Castro A. & Kuusi T. & Palatucci G.: Local behavior of fractional p-minimizers. *Submitted paper.*
[10] Di Castro A. & Kuusi T. & Palatucci G.: Nonlocal Harnack inequalities. *J. Funct. Anal.*, 267 (2014), 1807–1836.
[11] Di Nezza E. & Palatucci P. & Valdinoci E.: Hitchhikers guide to the fractional Sobolev spaces. *Bull. Sci. Math.* 136 (2012), 521–573.
[12] Dorronsoro J.R.: A characterization of potential spaces. *Proc. Amer. Math. Soc.* 95 (1985), 21–31.
[13] Elcrat A. & Meyers N.: Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions. *Duke Math. J.* 42 (1975), 121–136.
[14] Felsinger M. & Kassmann M.: Local regularity for parabolic nonlocal operators. *Comm. PDE* 38 (2013), 1539–1573.
[15] Fusco N. & Sbordone C.: Higher integrability of the gradient of minimizers of functionals with nonstandard growth conditions. *Comm. Pure Appl. Math.* 43 (1990), 673–683.
[16] Gehring F. W.: The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* 130 (1973), 265–277.
[17] Giusti E.: *Direct methods in the calculus of variations.* World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
[18] Iwaniec T.: The Gehring lemma. *Quasiconformal mappings and analysis (Ann Arbor, MI, 1995)*, 181–204, Springer, New York, 1998.
[19] Kassmann M.: A priori estimates for integro-differential operators with measurable kernels. *Calc. Var. & PDE* 34 (2009), 1–21.
[20] Kinnunen J. & Lewis J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* 102 (2000), 253–271.
[21] Maz'ya V.: *Sobolev spaces with applications to elliptic partial differential equations. Second, revised and augmented edition.* Grund. der Math. Wiss., 342. Springer, Heidelberg, 2011. xxviii+866 pp.
[22] Kuusi T. & Mingione G.: Universal potentials estimates. *J. Funct. Anal.* 262 (2012), 4205–4269.
[23] Kuusi T. & Mingione G. & Sire Y.: A Fractional Gehring lemma, with applications to nonlocal equations. *Rend. Lincei - Mat. Appl.* (2014).
[24] Meyers N. G.: An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa C. Sci. (3)* 17 (1963), 189–206.

[25] Mingione G.: The singular set of solutions to non-differentiable elliptic systems. *Arch. Ration. Mech. Anal.* 166 (2003), 287–301.
[26] Stein E. M.: The characterization of functions arising as potentials. *Bull. Amer. Math. Soc.* 67 (1961), 102–104.
[27] Stein E. M.: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton Math. Series, 43. Princeton University Press, Princeton, NJ, 1993.

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

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

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