Ann. Inst. H. Poincaré
Anal. Non Linéaire (Online first)
DOI 10.4171/AIHPC/37

Regularity theory for nonlocal equations with VMO coefficients

Simon Nowak

Abstract. We prove higher regularity for nonlinear nonlocal equations with possibly discontinuous coefficients of VMO type in fractional Sobolev spaces. While for corresponding local elliptic equations with VMO coefficients it is only possible to obtain higher integrability, in our nonlocal setting we are able to also prove a substantial amount of higher differentiability, so that our result is in some sense of purely nonlocal type. By embedding, we also obtain higher Hölder regularity for such nonlocal equations.

1. Introduction

1.1. Setting

In this work, we are dealing with nonlinear nonlocal integro-differential equations of the form

$$L_A^\Phi u = f \quad \text{in } \Omega \subset \mathbb{R}^n, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is a domain (= open set) and $f: \Omega \to \mathbb{R}$ is a given function, while $A: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a coefficient and $\Phi: \mathbb{R} \to \mathbb{R}$ is a nonlinearity with properties to be specified below. Moreover, for some fixed $s \in (0,1)$ the nonlocal operator L_A^Φ is formally defined by

$$L_A^\Phi u(x) := \text{p.v.} \int_{\mathbb{R}^n} \frac{A(x,y)}{|x-y|^{n+2s}} \Phi(u(x)-u(y))\, dy, \quad x \in \Omega. \tag{1.2}$$

For the sake of simplicity, throughout the paper we assume that $n > 2s$. Moreover, we assume that the coefficient A is measurable and that there exists a constant $\Lambda \geq 1$ such that

$$\Lambda^{-1} \leq A(x,y) \leq \Lambda \quad \text{for almost all } x, y \in \mathbb{R}^n. \tag{1.3}$$

In addition, we require A to be symmetric, that is,

$$A(x,y) = A(y,x) \quad \text{for almost all } x, y \in \mathbb{R}^n. \tag{1.4}$$

2020 Mathematics Subject Classification. Primary 35R09; Secondary 35B65, 35D30, 46E35, 47G20.
Keywords. Nonlocal operator, nonlocal equations, Sobolev regularity, Hölder regularity, Calderón–Zygmund estimates.

We define $\mathcal{L}_0(\Lambda)$ as the class of all such measurable coefficients A that satisfy conditions (1.3) and (1.4). Furthermore, we require that the nonlinearity Φ satisfies $\Phi(0) = 0$ and the following Lipschitz continuity and monotonicity assumptions, namely

$$|\Phi(t) - \Phi(t')| \leq \Lambda |t - t'| \quad \text{for all } t, t' \in \mathbb{R} \tag{1.5}$$

and

$$(\Phi(t) - \Phi(t'))(t - t') \geq \Lambda^{-1}(t - t')^2 \quad \text{for all } t, t' \in \mathbb{R}, \tag{1.6}$$

where for simplicity we use the same constant $\Lambda \geq 1$ as in (1.3). The above conditions are for instance satisfied by any C^1 function Φ with $\Phi(0) = 0$ such that the image of the first derivative Φ' of Φ is contained in $[\Lambda^{-1}, \Lambda]$. Consider the fractional Sobolev space

$$W^{s,2}(\mathbb{R}^n) = \left\{ u \in L^2(\mathbb{R}^n) \;\middle|\; \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \tfrac{|u(x)-u(y)|^2}{|x-y|^{n+2s}} \, dy \, dx < \infty \right\}$$

and denote by $W^{s,2}_c(\Omega)$ the set of all functions that belong to $W^{s,2}(\mathbb{R}^n)$ and are compactly supported in Ω. We are now in a position to define weak solutions of equation (1.1) as follows.

Definition. Given $f \in L^{\frac{2n}{n+2s}}_{\mathrm{loc}}(\Omega)$, we say that $u \in W^{s,2}(\mathbb{R}^n)$ is a weak solution of the equation $L^{\Phi}_A u = f$ in Ω, if

$$\begin{aligned} &\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{A(x,y)}{|x-y|^{n+2s}} \Phi(u(x) - u(y))(\varphi(x) - \varphi(y)) \, dy \, dx \\ &\quad = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in W^{s,2}_c(\Omega). \end{aligned} \tag{1.7}$$

We remark that the right-hand side of (1.7) is finite in view of using Hölder's inequality with Hölder conjugates $\frac{2n}{n+2s}$ and $\frac{2n}{n-2s}$ and the fractional Sobolev embedding (see Proposition 2.3).

In our main results, we require A to be of vanishing mean oscillation close to the diagonal in the following sense.

Definition. Let $\delta > 0$ and $A \in \mathcal{L}_0(\Lambda)$. We say that A is δ-vanishing in a ball $B \subset \mathbb{R}^n$ if, for any $r > 0$ and all $x_0, y_0 \in B$ with $B_r(x_0) \subset B$ and $B_r(y_0) \subset B$, we have

$$\fint_{B_r(x_0)} \fint_{B_r(y_0)} |A(x,y) - \bar{A}_{r,x_0,y_0}| \, dy \, dx \leq \delta,$$

where $\bar{A}_{r,x_0,y_0} := \fint_{B_r(x_0)} \fint_{B_r(y_0)} A(x,y) \, dy \, dx$.

Moreover, we say that A is (δ, R)-BMO in a domain $\Omega \subset \mathbb{R}^n$ and for some $R > 0$, if for any $z \in \Omega$ and any $0 < r \leq R$ with $B_r(z) \Subset \Omega$, A is δ-vanishing in $B_r(z)$.

Finally, we say that A is VMO in Ω, if for any $\delta > 0$, there exists some $R > 0$ such that A is (δ, R)-BMO in Ω.

If A belongs to the classical space of functions with vanishing mean oscillation $\mathrm{VMO}(\mathbb{R}^{2n})$ (see e.g. [30, Section 2.1.1], [15] or [39]), then A is also VMO in $\mathbb{R}^n$ in the above sense. Nevertheless, our assumption that A is VMO in Ω is more general, as it roughly speaking only means that A is of vanishing mean oscillation in some arbitrarily small open neighborhood of the diagonal in $\Omega \times \Omega$, while away from the diagonal in $\Omega \times \Omega$ and outside $\Omega \times \Omega$ the behavior of A is allowed to be more general. In particular, if A is continuous in an open neighborhood of the diagonal in $\Omega \times \Omega$, then A is clearly VMO in Ω. Nevertheless, continuity close to the diagonal is not essential, as there are plenty of VMO functions that are discontinuous. For example, assuming that Ω contains the origin, if for some $\alpha \in (0,1)$ we have

$$A(x,y) = \begin{cases} \sin\left(\left|\log(|x|+|y|)\right|^{\alpha}\right) + 2 & \text{if } x \neq 0 \text{ or } y \neq 0, \\ 0 & \text{if } x = y = 0, \end{cases} \tag{1.8}$$

or

$$A(x,y) = \begin{cases} \sin\left(\log\left|\log(|x|+|y|)\right|\right) + 2 & \text{if } x \neq 0 \text{ or } y \neq 0, \\ 0 & \text{if } x = y = 0, \end{cases} \tag{1.9}$$

in an open neighborhood of $\mathrm{diag}(\Omega \times \Omega)$, then A is VMO in Ω. However, in both cases A is discontinuous at $x = y = 0$.

1.2. Main results

Our first main result is concerned with Sobolev regularity.

Theorem 1.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $s \in (0,1)$ and $p \in [2,\infty)$. Moreover, fix some t such that*

$$s \leq t < \min\left\{2s\left(1-\frac{1}{p}\right), 1-\frac{2-2s}{p}\right\} = \begin{cases} 2s\left(1-\frac{1}{p}\right) & \text{if } s \leq 1/2 \\ 1-\dfrac{2-2s}{p} & \text{if } s > 1/2 \end{cases} =: t_{\sup}. \tag{1.10}$$

If $A \in \mathcal{L}_0(\Lambda)$ is VMO in Ω and if Φ satisfies conditions (1.5) and (1.6) with respect to Λ, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation

$$L_A^{\Phi} u = f \quad \text{in } \Omega,$$

we have the implication

$$f \in L^{\frac{np}{n+(2s-t)p}}_{\mathrm{loc}}(\Omega) \Rightarrow u \in W^{t,p}_{\mathrm{loc}}(\Omega).$$

Remark 1.2. In fact, in order to arrive at the conclusion that $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$ for some t and some p in Theorem 1.1, it is actually enough to assume that A is (δ, R)-BMO in Ω for some arbitrarily small $R > 0$ and some small enough $\delta > 0$ depending only on p, t, n, s and Λ; see Theorem 9.1 below. This is in line with corresponding results for local elliptic equations; see e.g. [6].

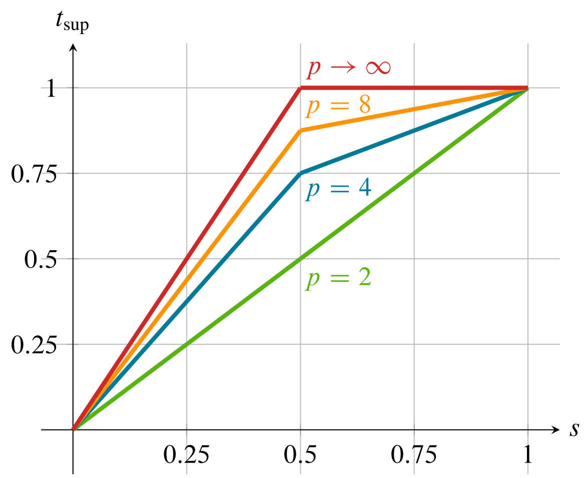

Figure 1. Higher differentiability in relation to s and p.

An interesting feature of Theorem 1.1 is that the differentiability gain indicated by the number $t_{\sup}$ depends on the gain of integrability by relation (1.10). This relation is visualized in the Figure 1.

In particular, on the one hand we observe that in the case when p is close to 2, that is, in the case of a small gain of integrability, Theorem 1.1 also implies only a small gain of differentiability. On the other hand, in the limit case when $p \to \infty$ we obtain differentiability in the whole range $s \leq t < \min\{2s, 1\}$, which we expect to be sharp in the case when A is merely VMO. An interesting question is whether also in the case of smaller values of p the differentiability gain in Theorem 1.1 can be improved beyond $t_{\sup}$ to the full range $s \leq t < \min\{2s, 1\}$, or whether counterexamples that contradict such an improvement can be constructed. This is because such an improved gain of differentiability was in fact observed in the recent paper [32]. However, in [32] this improved regularity is only proved in the linear case when $\Phi(t) = t$ and under some Hölder continuity assumption on A, which in particular does not include many examples of discontinuous VMO coefficients like (1.8) and (1.9); see Section 1.4 for more details.

In Theorem 1.1, we stated the result in terms of the higher integrability exponent p at which we arrive, which has the advantage that the statement of Theorem 1.1 is relatively clean. However, an interesting question is how much higher integrability and differentiability we gain if we instead prescribe the integrability of the source function f. This question leads to the following reformulation of Theorem 1.1.

Theorem 1.3. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $s \in (0,1)$ and $f \in L^q_{\mathrm{loc}}(\Omega)$ for some $q \in (\frac{2n}{n+2s}, \infty)$. In addition, assume that $A \in \mathcal{L}_0(\Lambda)$ is VMO in Ω and that Φ satisfies conditions (1.5) and (1.6) with respect to Λ. Then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation $L_A^{\Phi} u = f$ in Ω, the following is true. Fix some t such that $s \leq t < 1$:*

- *If t satisfies*

$$2s - \frac{n}{q} < t < \begin{cases} 2s\Big(1 - \dfrac{n}{(n+2s)q}\Big) & \text{if } s \leq 1/2, \\ 1 - \dfrac{(2-2s)(n+q-2sq)}{(n+2-2s)q} & \text{if } s > 1/2, \end{cases} \tag{1.11}$$

then we have $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$, where $p = \frac{nq}{n-(2s-t)q}$.

- *If t satisfies*

$$t \leq 2s - \frac{n}{q}, \tag{1.12}$$

then we have $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$ for any $p \in (1,\infty)$.

Note that in the first case of Theorem 1.3 we always have $\frac{nq}{n-(2s-t)q} > 2$, so that we always gain integrability beyond the initial integrability exponent 2 as well as differentiability beyond the initial differentiability parameter s.

Moreover, we note that in the case when $2s - \frac{n}{q} < 1$, it is relatively easy to see that we always have

$$2s - \frac{n}{q} < \begin{cases} 2s\Big(1 - \dfrac{n}{(n+2s)q}\Big) & \text{if } s \leq 1/2, \\ 1 - \dfrac{(2-2s)(n+q-2sq)}{(n+2-2s)q} & \text{if } s > 1/2, \end{cases}$$

so that in this case the range of t given by (1.11) is always nonempty.

Also, we remark that in the case when $2s - \frac{n}{q} \geq 1$, Theorem 1.3 implies that u belongs to $W^{t,p}_{\mathrm{loc}}(\Omega)$ for any t in the range $s \leq t < 1$ and any $p \in (1,\infty)$.

By embedding, Theorem 1.3 also implies the following higher Hölder regularity result.

Theorem 1.4. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $s \in (0,1)$ and $f \in L^q_{\mathrm{loc}}(\Omega)$ for some $q > \frac{n}{2s}$. If $A \in \mathcal{L}_0(\Lambda)$ is VMO in Ω and Φ satisfies conditions (1.5) and (1.6) with respect to Λ, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation*

$$L_A^\Phi u = f \quad \text{in } \Omega,$$

we have

$$u \in \begin{cases} C^{2s-\frac{n}{q}}_{\mathrm{loc}}(\Omega) & \text{if } 2s - \dfrac{n}{q} < 1, \\ C^\alpha_{\mathrm{loc}}(\Omega)\ \forall \alpha \in (0,1) & \text{if } 2s - \dfrac{n}{q} \geq 1. \end{cases} \tag{1.13}$$

While, as mentioned, it is up to further investigation whether the differentiability gain in Theorems 1.1 and 1.3 is optimal, we nevertheless expect the Hölder regularity in Theorem 1.4 to be sharp in the case of VMO coefficients or even continuous coefficients, since even the mentioned improved gain of differentiability along the Sobolev scale in the range $s \leq t < \{2s, 1\}$ would still only lead to the same amount of Hölder regularity obtained in Theorem 1.4.

1.3. Local elliptic equations with VMO coefficients

For the sake of comparison, let us briefly discuss corresponding regularity results for local elliptic equations in divergence form of the type

$$\operatorname{div}(B\nabla u) = f \quad \text{in } \Omega, \tag{1.14}$$

where the matrix of coefficients $B = \{b_{ij}\}_{i,j=1}^n$ is assumed to be uniformly elliptic and bounded. In the linear case when $\Phi(t) = t$, equation (1.14) can in some sense be thought of as a local analogue of the nonlocal equation (1.1) corresponding to the limit case $s = 1$. For some rigorous results in this direction, we refer to [21]. It is known that if the coefficients b_{ij} belong to VMO(Ω) and $f \in L^{\frac{np}{n+p}}_{\mathrm{loc}}(\Omega)$ for some $p > 2$, then weak solutions $u \in W^{1,2}_{\mathrm{loc}}(\Omega)$ of equation (1.14) belong to $W^{1,p}_{\mathrm{loc}}(\Omega)$; see e.g. [6, 15, 23] and also [1, 2, 17, 26] for more general developments in this direction. This corresponds to our Theorem 1.1 in the case when $t = s$. On the other hand, in order to gain any amount of differentiability along the Sobolev scale in the context of local equations, a corresponding amount of differentiability has to be imposed on the coefficients, so that in the case of VMO coefficients in general no differentiability gain at all is attainable. Therefore, the additional differentiability gain in Theorem 1.1 is in some sense a purely nonlocal phenomenon.

This nonlocal differential stability effect is also visible in the context of Hölder regularity, although in this case it is somewhat more subtle to recognize it. In fact, embedding the above $W^{1,p}$ regularity result implies that for any weak solution $u \in W^{1,2}_{\mathrm{loc}}(\Omega)$ of (1.14) with $f \in L^q_{\mathrm{loc}}(\Omega)$ for some $q > \frac{n}{2}$ we indeed have

$$u \in \begin{cases} C^{2-\frac{n}{q}}_{\mathrm{loc}}(\Omega) & \text{if } q < n, \\ C^{\alpha}_{\mathrm{loc}}(\Omega) \ \forall \alpha \in (0,1) & \text{if } q \geq n, \end{cases}$$

which at first sight directly corresponds to Theorem 1.4. However, there is an important difference in the case when q is large, which is due to the differentiability gain in Theorem 1.3. In order to illustrate this difference, note that in the case when $f \in L^{\infty}_{\mathrm{loc}}(\Omega)$, for any weak solution $u \in W^{1,2}_{\mathrm{loc}}(\Omega)$ of (1.14) we have $C^{\alpha}_{\mathrm{loc}}(\Omega)$ for any $\alpha \in (0,1)$. Since in some sense the order of equation (1.1) is s times the order of equation (1.14), one might therefore be tempted to guess that weak solutions $u \in W^{s,2}(\mathbb{R}^n)$ of (1.1) should in general not exceed C^s regularity. However, Theorem 1.4 shows that any such weak solution to (1.1) indeed belongs to $C^{\alpha}_{\mathrm{loc}}(\Omega)$ for any $0 < \alpha < \min\{2s, 1\}$ whenever $f \in L^{\infty}_{\mathrm{loc}}(\Omega)$, exceeding C^s regularity. In particular, in the case when $s \geq 1/2$, such weak solutions to nonlocal equations with VMO coefficients and locally bounded right-hand side enjoy the same amount of Hölder regularity as weak solutions to corresponding local equations with VMO coefficients, despite the fact that the order of such nonlocal equations is lower.

1.4. Previous results

In recent years, the regularity theory for weak solutions to nonlocal equations of type (1.1) has seen a great amount of progress, in particular concerning regularity results of purely nonlocal type, in the sense that, as above, the obtained regularity is better than one might expect when considering corresponding results for local elliptic equations.

Regarding such results for general coefficients $A \in \mathcal{L}_0(\Lambda)$, in [28] and [40] it is demonstrated that weak solutions to nonlocal equations of type (1.1) have slightly higher differentiability and higher integrability along the scale of fractional Sobolev spaces, which is a phenomenon not shared by local elliptic equations of type (1.14) with merely measurable coefficients, where it is only possible to obtain higher integrability.

Concerning higher Sobolev regularity of purely nonlocal type, in [32] the authors in particular show that in the linear case when $\Phi(t) = t$, if $\Omega = \mathbb{R}^n$ and if the mapping $x \mapsto A(x, y)$ is globally Hölder continuous with some arbitrary Hölder exponent, then the statement of Theorem 1.1 holds for t in the improved range $s \leq t < \min\{2s, 1\}$. As we discussed briefly in Section 1.2, an interesting question is therefore whether the regularity obtained in [32] can be replicated in our general setting of possibly nonlinear equations with VMO coefficients posed on general domains $\Omega \subset \mathbb{R}^n$, in particular since many examples of discontinuous VMO coefficients like (1.8) or (1.9) are not covered by the Hölder continuity assumption in [32].

On the other hand, regarding higher Hölder regularity, by the Sobolev embedding the mentioned Sobolev regularity result in [32] implies exactly the same amount of Hölder regularity given by (1.13) from Theorem 1.4 under the mentioned assumptions imposed in [32]. In other words, although in comparison to [32] in general we obtain less differentiability along the Sobolev scale, we nevertheless gain enough differentiability in order to obtain the same amount of Hölder regularity by embedding in our general setting. A similar Hölder regularity result was obtained in [19], again in the case of linear equations, but allowing for coefficients that are merely continuous. Concerning higher Hölder regularity for possibly nonlinear equations, in [35] it is in particular proved that if Φ satisfies assumptions (1.5) and (1.6) and A is continuous in Ω, then weak solutions u of (1.1) belong to $C^{\alpha}_{\mathrm{loc}}(\Omega)$ for any $0 < \alpha < \min\{2s - \frac{n}{q}, 1\}$ whenever $f \in L^q_{\mathrm{loc}}(\Omega)$ for some $q > \frac{n}{2s}$, which almost matches the regularity (1.13) obtained in Theorem 1.4. In addition, while in comparison with Theorem 1.4 the result in [35] does not include general discontinuous coefficients of VMO type, it in fact holds for a slightly larger class of coefficients than simply continuous ones, including in particular coefficients that are translation invariant inside Ω. Nevertheless, our approach can easily be modified in order to prove our main results under the assumption on A from [35]; see Remark 9.3. In addition, the Hölder regularity result in [35] holds for a slightly larger class of weak solutions called local weak solutions, essentially only assuming that $u \in W^{s,2}_{\mathrm{loc}}(\Omega)$ and the finiteness of the nonlocal tails of u; see [35]. While we believe that our approach can be modified in order to generalize our main results to this setting of local weak solutions, we decided not to insist on this point, in particular since this would also require a revision of the previous work [28].

Let us also mention that in [36] (see also [34]), Theorem 1.1 was proved in the case when $t = s$, that is, without the additional differentiability gain, under essentially the same assumptions on A and Φ as in [35].

More results concerning Sobolev regularity for nonlocal equations are for example proved in [3, 4, 12, 18, 22, 25, 31, 44], while some more results on Hölder regularity are proved in [5, 7, 9–11, 13, 14, 20, 24, 29, 38, 41]. Furthermore, for various regularity results regarding nonlocal equations similar to (1.1) in the more general setting of measure data, we refer to [27].

1.5. Approach

Our approach is mainly influenced by techniques introduced in [8] and [28]. Namely, in [8] techniques were developed allowing us to prove higher integrability of the gradient ∇u of weak solutions to local equations with VMO coefficients of type (1.14), which corresponds to the $W^{1,p}$ regularity theory briefly discussed in Section 1.2.

The approach can be summarized as follows. The first step is to use the assumption that the coefficients b_{ij} are VMO in order to locally approximate the gradient of some weak solution u of (1.14) by the gradient of a weak solution v to a suitable homogeneous equation with constant coefficients. In order to include discontinuous coefficients of VMO type into the analysis, one uses the fact that ∇u is known to satisfy an $L^{2+\gamma}_{\mathrm{loc}}$ estimate for some small $\gamma > 0$, which can be proved in the general setting of merely bounded measurable coefficients by means of so-called Gehring-type lemmas. One then exploits the fact that the approximate solution v, which in the local case up to a change of coordinates is simply a harmonic function, is already known to satisfy a local Lipschitz estimate in order to transfer some regularity from v to u. This transfer of regularity is achieved by covering the level sets of the Hardy–Littlewood maximal function of $|\nabla u|^2$ of the form $\{\mathcal{M}(|\nabla u|^2) > \lambda^2\}$ by dyadic cubes that are chosen by means of an exit time argument and form a so-called Calderón–Zygmund covering, essentially meaning that the cubes in the covering have in some sense good density properties with respect to the level set that is covered by them. Combined with the fact that ∇u can be approximated by the gradient of a harmonic and therefore very regular function, these good density properties are then exploited in order gain control of the measures of the cubes in the covering by means of so-called good-λ inequalities. By standard arguments from measure theory, this then allows us to prove the desired higher integrability of ∇u, which then implies the desired $W^{1,p}_{\mathrm{loc}}$ estimate.

Adapting this approach in order to prove higher Sobolev regularity for nonlocal equations of type (1.1) comes with a number of obstacles. In particular, a main challenge in the nonlocal context is to find a suitable replacement for the gradient ∇u which is used in the local context. In [34] and [36], the above approach was executed for weak solutions u to nonlocal equations of type (1.1) with the local gradient replaced by the nonlocal gradient-type operator

$$\nabla^s u(x) = \left(\int_{\mathbb{R}^n} \frac{(u(x)-u(y))^2}{|x-y|^{n+2s}} \, dy \right)^{\frac{1}{2}}. \tag{1.15}$$

In view of an alternative characterization of Bessel potential spaces, the obtained higher integrability of $\nabla^s u$ then leads to $W^{s,p}_{\mathrm{loc}}$ regularity, which corresponds to the case of no differentiability gain as in the setting of local equations. However, as no local higher integrability estimate for small exponents is known for $\nabla^s u$ for weak solutions u to (1.1), the main result in [36] does not include the case of VMO coefficients. In addition, while this result corresponds to the $W^{1,p}_{\mathrm{loc}}$ estimate obtained in the setting of local equations, considering the gradient-type operator (1.15) does not lead to any higher differentiability.

In order to also gain higher differentiability and include the case of VMO coefficients, we instead use another nonlocal-type gradient operator which is inspired by [28]. Fix some $\theta \in (0, \frac{1}{2})$. We define a Borel measure μ on $\mathbb{R}^{2n}$ as follows. For any measurable set $E \subset \mathbb{R}^{2n}$, set

$$\mu(E) := \int_E \frac{dx\, dy}{|x-y|^{n-2\theta}}. \tag{1.16}$$

Moreover, for any function $u\colon \mathbb{R}^n \to \mathbb{R}$ and $(x,y) \in \mathbb{R}^{2n}$ with $x \neq y$, we define the function

$$U(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{s+\theta}}. \tag{1.17}$$

For any domain $\Omega \subset \mathbb{R}^n$, we then clearly have $u \in W^{s,2}(\Omega)$ if and only if $u \in L^2(\Omega)$ and $U \in L^2(\Omega \times \Omega, \mu)$, so that U and μ are in some sense in duality. Regarding larger exponents, by a simple computation for any $p > 2$ and $s_\theta := s + \theta(1 - \frac{2}{p}) > s$, we have

$$u \in W^{s_\theta,p}(\Omega) \quad \text{if and only if} \quad u \in L^p(\Omega) \text{ and } U \in L^p(\Omega \times \Omega, \mu). \tag{1.18}$$

Therefore, in contrast to the gradient-type operator ∇^s, by proving higher integrability of the gradient-type function U with respect to the measure μ, we do not only gain regularity along the integrability scale of fractional Sobolev spaces, but also a substantial amount of higher differentiability! However, proving this higher integrability of U in the case when A is merely VMO in Ω comes with a number of additional difficulties. In order to accomplish this, we combine nonlocal adaptations of the approximation and covering techniques from [8] with adaptations of some further covering and combinatorial techniques from [28].

First of all, in order to include equations with VMO coefficients, we need a local higher integrability result for U for small exponents, which is proved in [28] for $\theta > 0$ small enough, which is sufficient for our purposes.

Furthermore, in contrast to the functions ∇u and $\nabla^s u$ which are defined on $\mathbb{R}^n$, the function U is defined on $\mathbb{R}^{2n}$. In particular, the level sets of the maximal function with respect to μ of U^2, that is, the sets of the form $\{\mathcal{M}(U^2) > \lambda^2\}$, are subsets of $\mathbb{R}^{2n}$ instead of $\mathbb{R}^n$. Therefore, in this setting we need to run an exit time argument in $\mathbb{R}^{2n}$ instead of $\mathbb{R}^n$ in order to cover the level set of U by Calderón–Zygmund cubes in $\mathbb{R}^{2n}$. In other words, every dyadic cube $\mathcal{K}$ in the corresponding Calderón–Zygmund covering is of the form $\mathcal{K} = K_1 \times K_2$, where K_1 and K_2 are dyadic cubes in $\mathbb{R}^n$. A major technical issue that arises at this point is that for cubes $\mathcal{K} = K_1 \times K_2$ that are far away from the diagonal in the sense that $\operatorname{dist}(K_1, K_2)$ is large, the information that u solves a nonlocal equation of type (1.1) cannot be used effectively. For this reason, we additionally construct an auxiliary diagonal cover consisting of diagonal balls $\mathcal{B} = B \times B$ that once again have nice density properties with respect to the level set $\{\mathcal{M}(U^2) > \lambda^2\}$. Since close to the diagonal the information given by the equation can be used much more efficiently, we construct this auxiliary cover in such a way that the exit time at which the balls are chosen is somewhat smaller than the corresponding exit time at which the corresponding Calderón–Zygmund cubes are chosen, so that the balls in the auxiliary cover tend to be somewhat larger than the corresponding Calderón–Zygmund cubes. All in all, roughly speaking we have

$$\{\mathcal{M}(U^2) > \lambda^2\} \subset \bigcup \mathcal{B} \cup \bigcup \mathcal{K},$$

where the balls $\mathcal{B}$ are diagonal balls with good density properties and the cubes $\mathcal{K}$ are Calderón–Zygmund cubes that are far away from the diagonal.

The measures of the balls in the auxiliary diagonal cover can then be estimated by approximating U by a corresponding function V in small enough balls, which is given as in (1.17) with u replaced by a weak solution v of a corresponding equation of the form $L^{\Phi}_{\tilde{A}} v = 0$, where the coefficient A is locally replaced by a suitable constant, while the global behavior of A has to be left unchanged, since our assumption that A is VMO is local in nature. This leads to the issue that proving a strong enough estimate for v, enabling us to transfer enough regularity to u, is much more difficult than in the setting of linear local equations, where as mentioned, the approximate solution is effectively simply a harmonic function. Nevertheless, in [35] it is proved that such weak solutions v to $L^{\Phi}_{\tilde{A}} v = 0$ satisfy a $C^{s+\theta}_{\mathrm{loc}}$ estimate in the restricted range $0 < \theta < \min\{s, 1-s\}$. This Hölder estimate directly implies that V satisfies an L^∞_{loc} estimate, which is sufficient in order to control the measures of the balls in the auxiliary diagonal cover.

As already indicated, the task of controlling the measures of the off-diagonal Calderón–Zygmund cubes requires additional ideas, since far from the diagonal the information provided by the equation is only of very limited use. In order to bypass this problem in the context of proving higher integrability and differentiability of u for small exponents, in [28, Lemma 5.3] it was noted that on cubes that are far away from the diagonal, L^2-reverse Hölder-type inequalities hold for U without relying on the equation. Since we want to prove higher integrability for large exponents as well, we overcome this problem by noticing that such reverse Hölder-type inequalities for off-diagonal cubes also hold for larger exponents. However, as in [28, Lemma 5.3], these reverse Hölder-type inequalities come with additional diagonal correction terms involving diagonal cubes that do not belong to the original Calderón–Zygmund covering, leading to serious difficulties. These difficulties are bypassed by an involved combinatorial argument inspired by a corresponding one in [28], enabling us to also control the measures of the off-diagonal Calderón–Zygmund cubes.

By combining the estimates for the measures of the diagonal balls and off-diagonal cubes, we are then able to estimate the measure of the level set $\{\mathcal{M}(U^2) > \lambda^2\}$ for λ large enough, which by a standard application of Fubini's theorem and standard properties of the Hardy–Littlewood maximal function implies the desired L^p_{loc} estimate for U in the form of an a priori estimate, which is then used in order to prove the desired regularity by standard smoothing techniques based on mollifiers.

1.6. Outline of the paper

The paper is organized as follows. In Section 2 we formally introduce the fractional Sobolev spaces $W^{s,p}$ and mention some important results concerning these spaces.

In Section 3.1 we turn to discussing some simple properties of the measure μ introduced in the previous Section 1.5, while in Section 3.2 we define the Hardy–Littlewood maximal function with respect to the measure μ and mention some important properties of it.

In Section 4 we then discuss some preliminary estimates for nonlocal equations which are essentially known. More precisely, in Section 4.1 we briefly recall the mentioned higher Hölder regularity result from [35], while in Section 4.2 we recall the mentioned Sobolev regularity result for small exponents contained in [28]. In Section 4.3 we then state a result about $H^{2s,p}$ estimates for the homogeneous Dirichlet problem involving the fractional Poisson-type equation $(-\Delta)^s g = f$, where $(-\Delta)^s$ is the fractional Laplacian. This estimate allows us to focus on proving regularity for nonlocal equations of the type $L^\Phi_A u = (-\Delta)^s g$ instead of (1.1), since once we are able to transfer a sufficient amount of regularity from g to u, Theorem 1.1 follows by first transferring the regularity from f to some solution g of $(-\Delta)^s g = f$ and then from g to weak solutions u of (1.1).

The rest of the paper is then devoted to the proof of our main results. Namely, in Section 5 we prove a comparison estimate enabling us to carry out the approximation argument and also the smoothing procedure mentioned in Section 1.5. Section 6 is devoted to proving good-λ inequalities, both at the diagonal and far away from the diagonal. In Section 7 we then set up the mentioned covering argument and use the good-λ inequalities from Section 6 in order to estimate the measure of the level sets of $\mathcal{M}(U^2)$. In Section 8, this level set estimate is then used in order to prove the desired regularity in the form of an a priori estimate. Finally, in Section 9 we then use smoothing techniques in order to deduce our main results from the a priori estimates obtained in Section 8.

1.7. Some definitions and notation

For convenience, let us fix some notation which we use throughout the paper. By C, c and $C_i, c_i, i \in \mathbb{N}_0$, we always denote positive constants, while dependences on parameters of the constants will be shown in parentheses. As usual, by

$$B_r(x_0) := \left\{x \in \mathbb{R}^n \;\middle|\; |x - x_0| < r\right\}$$

we denote the open euclidean ball with center $x_0 \in \mathbb{R}^n$ and radius $r > 0$. We also set $B_r := B_r(0)$. In addition, by

$$Q_r(x_0) := \left\{x \in \mathbb{R}^n \,\middle|\, |x - x_0|_\infty < r/2\right\}$$

we denote the open cube with center $x_0 \in \mathbb{R}^n$ and sidelength $r > 0$. Moreover, if $E \subset \mathbb{R}^n$ is measurable, then by $|E|$ we denote the n-dimensional Lebesgue measure of E. If $0 < |E| < \infty$, then for any $u \in L^1(E)$ we define

$$\bar{u}_E := \fint_E u(x)\,dx := \frac{1}{|E|}\int_E u(x)\,dx.$$

Throughout this paper, we often consider integrals and functions on $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$. Instead of dealing with the usual euclidean balls in $\mathbb{R}^{2n}$, for this purpose it is more convenient for us to use the balls generated by the norm

$$\|(x_0, y_0)\| := \max\{|x_0|, |y_0|\}, \quad (x_0, y_0) \in \mathbb{R}^{2n}.$$

These balls with center $(x_0, y_0) \in \mathbb{R}^{2n}$ and radius $r > 0$ are denoted by $\mathcal{B}_r(x_0, y_0)$ and are of the form

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

In the case when $x_0 = y_0$ we also write $\mathcal{B}_r(x_0) := \mathcal{B}_r(x_0, x_0)$, and we call such balls diagonal balls. We also set $\mathcal{B}_r := \mathcal{B}_r(0)$. Similarly, for $x_0, y_0 \in \mathbb{R}^n$ and $r > 0$ we define $\mathcal{Q}_r(x_0, y_0) := Q_r(x_0) \times Q_r(y_0)$ and $\mathcal{Q}_r(x_0) := \mathcal{Q}_r(x_0, x_0)$ and also $\mathcal{Q}_r := \mathcal{Q}_r(0)$.

2. Fractional Sobolev spaces

Definition. Let $\Omega \subset \mathbb{R}^n$ be a domain. For $p \in [1, \infty)$ and $s \in (0, 1)$ we define the fractional Sobolev space

$$W^{s,p}(\Omega) := \left\{u \in L^p(\Omega) \,\middle|\, \int_\Omega \int_\Omega \tfrac{|u(x)-u(y)|^p}{|x-y|^{n+sp}}\,dy\,dx < \infty\right\}$$

with norm

$$\|u\|_{W^{s,p}(\Omega)} := (\|u\|^p_{L^p(\Omega)} + [u]^p_{W^{s,p}(\Omega)})^{1/p},$$

where

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

Moreover, we define the corresponding local fractional Sobolev spaces by

$$W^{s,p}_{\text{loc}}(\Omega) := \left\{u \in L^p_{\text{loc}}(\Omega) \,\middle|\, u \in W^{s,p}(\Omega') \text{ for any domain } \Omega' \Subset \Omega\right\}.$$

Also, we define the space

$$W^{s,p}_0(\Omega) := \left\{u \in W^{s,2}(\mathbb{R}^n) \,\middle|\, u = 0 \quad \text{in } \mathbb{R}^n \setminus \Omega\right\}.$$

We use the following fractional Poincaré inequality; see [33, Section 4].

Lemma 2.1 (Fractional Poincaré inequality). *Let $s \in (0,1)$, $p \in [1,\infty)$, $r > 0$ and $x_0 \in \mathbb{R}^n$. For any $u \in W^{s,p}(B_r(x_0))$ we have*

$$\int_{B_r(x_0)} |u(x) - \bar{u}_{B_r(x_0)}|^p \, dx \leq C r^{sp} \int_{B_r(x_0)} \int_{B_r(x_0)} \frac{|u(x) - u(y)|^p}{|x-y|^{n+sp}} \, dy \, dx,$$

where $C = C(s,p) > 0$.

We also use another Poncaré-type inequality; see [35, Lemma 2.3].

Lemma 2.2 (Fractional Friedrichs–Poincaré inequality). *Let $s \in (0,1)$ and consider a bounded domain $\Omega \subset \mathbb{R}^n$. For any $u \in W_0^{s,2}(\Omega)$ we have*

$$\int_{\mathbb{R}^n} |u(x)|^2 \, dx \leq C |\Omega|^{\frac{2s}{n}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}} \, dy \, dx, \tag{2.1}$$

where $C = C(n,s) > 0$.

For the following embedding results we refer to [16, Theorems 6.7, 6.10, 8.2].

Proposition 2.3. *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain, $s \in (0,1)$ and $p \in [1,\infty)$:*

- *If $sp < n$, then we have the continuous embedding*

$$W^{s,p}(\Omega) \hookrightarrow L^{\frac{np}{n-sp}}(\Omega).$$

- *If $sp = n$, then for any $q \in [1,\infty)$ we have the continuous embedding*

$$W^{s,p}(\Omega) \hookrightarrow L^q(\Omega).$$

- *If $sp > n$, then we have the continuous embedding*

$$W^{s,p}(\Omega) \hookrightarrow C^{s-\frac{n}{p}}(\Omega).$$

By combining Proposition 2.3 with Lemma 2.1 and a scaling argument, it is easy to deduce the following result.

Lemma 2.4 (Fractional Sobolev–Poincaré inequality). *Let $s \in (0,1)$, $p \in [1,\infty)$, $r > 0$ and $x_0 \in \mathbb{R}^n$. In addition, let*

$$q \in \begin{cases} \left[1, \dfrac{np}{n-sp}\right] & \text{if } sp < n, \\ [1,\infty) & \text{if } sp \geq n. \end{cases}$$

Then for any $u \in W^{s,p}(B_r(x_0))$ we have

$$\left(\fint_{B_r(x_0)} |u(x) - \bar{u}_{B_r(x_0)}|^q \, dx \right)^{\frac{1}{q}} \leq C r^s \left(\fint_{B_r(x_0)} \int_{B_r(x_0)} \frac{|u(x) - u(y)|^p}{|x-y|^{n+sp}} \, dy \, dx \right)^{\frac{1}{p}},$$

where $C = C(n,s,p,q) > 0$.

For $p \in (1, \infty)$ and $s \in (0, 2)$, denote by $H^{s,p}(\Omega)$ the standard Bessel potential spaces on Ω; see e.g. [34, Section 3]. The following embedding result follows from [43, Theorem 2.5], where it is given in the more general context of Besov and Triebel–Lizorkin spaces

Proposition 2.5. *Let $1 < p_0 < p < p_1 < \infty$, $s \in (0, 2)$, $s_0, s_1 \in (0, 1)$ and assume that $\Omega \subset \mathbb{R}^n$ is a smooth domain. If $s_0 - \frac{n}{p_0} = s - \frac{n}{p} = s_1 - \frac{n}{p_1}$, then*

$$W^{s_0,p_0}(\Omega) \hookrightarrow H^{s,p}(\Omega) \hookrightarrow W^{s_1,p_1}(\Omega).$$

Unlike the classical Sobolev spaces $W^{1,p}(\Omega)$ on a bounded domain $\Omega \subset \mathbb{R}^n$, the fractional Sobolev spaces $W^{s,p}(\Omega)$ are not contained in each other as the integrability exponent p decreases. Nevertheless, we have the following result, essentially stating that the mentioned inclusions are almost true.

Proposition 2.6. *Let $1 < p_0 \leq p < \infty$, $s \in (0, 1)$ and assume that $\Omega \subset \mathbb{R}^n$ is a smooth bounded domain. Then for any $0 < \varepsilon < \min\{1 - s, \frac{2n}{p}, 2n(1 - \frac{1}{p_0})\}$ we have*

$$W^{s+\varepsilon,p}(\Omega) \hookrightarrow W^{s,p_0}(\Omega).$$

Proof. By Proposition 2.5, we have $W^{s+\varepsilon,p}(\Omega) \hookrightarrow H^{s+\varepsilon/2,\tilde{p}}(\Omega)$, where $\tilde{p} := \frac{np}{n-\varepsilon p/2}$. Now since for $\hat{p}_0 := \frac{np_0}{n+\varepsilon p_0/2}$ we have $1 < \hat{p}_0 < p_0 \leq p < \tilde{p}$, by [43, Theorem 2.85(ii)] we have $H^{s+\varepsilon/2,\tilde{p}}(\Omega) \hookrightarrow H^{s+\varepsilon/2,\hat{p}_0}(\Omega)$. Since in addition by Proposition 2.5 we have $H^{s+\varepsilon/2,\hat{p}_0}(\Omega) \hookrightarrow W^{s,p_0}(\Omega)$, the proof is finished by combining the above three embeddings. ∎

3. The measure μ

3.1. Basic properties of μ

For the rest of this paper, we fix some $s \in (0, 1)$ along with some parameter θ in the range

$$0 < \theta < \min\{s, 1 - s\} \tag{3.1}$$

and let the measure μ be defined by (1.16). Moreover, for any function $u : \mathbb{R}^n \to \mathbb{R}$, let the function U be given by (1.17). The following relation can be deduced by a simple computation.

Lemma 3.1. *Let $p \geq 2$ and set $s_\theta := s + \theta(1 - \frac{2}{p})$. Then we have*

$$u \in W^{s_\theta,p}(\Omega) \quad \textit{if and only if} \quad u \in L^p(\Omega) \textit{ and } U \in L^p(\Omega \times \Omega, \mu)$$

and

$$\|U\|_{L^p(\Omega\times\Omega,d\mu)} = [u]_{W^{s_\theta,p}(\Omega)}.$$

The next proposition contains some further important properties of the measure μ which are straightforward to deduce by applying changes of variables. We will use these properties frequently throughout the paper, usually without explicit reference.

Proposition 3.2. (i) *For all $r > 0$ and $x_0 \in \mathbb{R}^n$, we have*

$$\mu(\mathcal{B}_r(x_0)) = \mu(\mathcal{B}_r) = c r^{n+2\theta},$$

where $c = c(n,\theta) > 0$.

(ii) *(Volume doubling property). For any $(x_0, y_0) \in \mathbb{R}^{2n}$, any $r > 0$ and any $M > 0$, we have*

$$\mu(\mathcal{B}_{Mr}(x_0, y_0)) = M^{n+2\theta} \mu(\mathcal{B}_r(x_0, y_0)).$$

3.2. The Hardy–Littlewood maximal function

Another tool we use is the Hardy–Littlewood maximal function with respect to the measure μ.

Definition. Let $F \in L^1_{\mathrm{loc}}(\mathbb{R}^{2n}, \mu)$. We define the Hardy–Littlewood maximal function $\mathcal{M}F : \mathbb{R}^{2n} \to [0, \infty]$ of F by

$$\mathcal{M}(F)(x, y) := \sup_{\rho > 0} \fint_{\mathcal{B}_\rho(x,y)} |F| \, d\mu,$$

where

$$\fint_{\mathcal{B}_\rho(x,y)} |F| \, d\mu := \frac{1}{\mu(\mathcal{B}_\rho(x, y))} \int_{\mathcal{B}_\rho(x,y)} |F| \, d\mu.$$

Moreover, for any open set $E \subset \mathbb{R}^{2n}$, we define

$$\mathcal{M}_E(F) := \mathcal{M}(F \chi_E),$$

where χ_E is the characteristic function of E. In addition, for any $r > 0$ we define

$$\mathcal{M}_{\geq r}(F)(x, y) := \sup_{\rho \geq r} \fint_{\mathcal{B}_\rho(x,y)} |F| \, d\mu, \quad \mathcal{M}_{\geq r, E}(F) := \mathcal{M}_{\geq r}(F \chi_E).$$

The following result shows that the Hardy–Littlewood maximal function behaves nicely in the context of L^p spaces. Since by Proposition 3.2, μ is a doubling measure with doubling constant $2^{n+2\theta}$, the result follows directly from [42, Chapter 1, Section 3, Theorem 1].

Proposition 3.3. *Let E be an open subset of $\mathbb{R}^{2n}$:*

(i) *(Weak p–p estimates). If $F \in L^p(E, \mu)$ for some $p \geq 1$ and $\lambda > 0$, then*

$$\mu\big(\{x \in E \mid \mathcal{M}_E(F)(x) > \lambda\}\big) \leq \frac{C}{\lambda^p} \int_E |F|^p \, d\mu,$$

where C depends only on n, θ and p.

(ii) *(Strong p–p estimates). If $F \in L^p(E, \mu)$ for some $p \in (1, \infty]$, then*

$$\|\mathcal{M}_E(F)\|_{L^p(E, d\mu)} \leq C \|F\|_{L^p(E, d\mu)},$$

where C depends only on n, θ and p.

For the following result we also refer to [42, Chapter 1, Section 3].

Proposition 3.4 (Lebesgue differentiation theorem). *If $F \in L^1_{\mathrm{loc}}(\mathbb{R}^{2n}, \mu)$, then for almost every $(x, y) \in \mathbb{R}^{2n}$, we have*

$$\lim_{r \to 0} \fint_{\mathcal{B}_r(x,y)} F \, d\mu = F(x, y).$$

An immediate corollary of the Lebesgue differentiation theorem is given as follows.

Corollary 3.5. *Let $F \in L^1_{\mathrm{loc}}(\mathbb{R}^{2n}, \mu)$. Then for almost every $(x, y) \in \mathbb{R}^{2n}$ we have*

$$|F(x, y)| \leq \mathcal{M}(F)(x, y).$$

In addition, for any open set $E \subset \mathbb{R}^{2n}$ and any $p \in [1, \infty]$ we have

$$\|F\|_{L^p(E,d\mu)} \leq \|\mathcal{M}_E(F)\|_{L^p(E,d\mu)}.$$

4. Some preliminary estimates

4.1. Higher Hölder regularity

The following result on higher Hölder regularity plays an essential role in our approach and follows from [35, Theorem 1.1].

Theorem 4.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain and let $f \in L^\infty_{\mathrm{loc}}(\Omega)$. Consider a coefficient $A \in \mathcal{L}_0(\Lambda)$ that is continuous in $\Omega \times \Omega$ and suppose that Φ satisfies (1.5) and (1.6) with respect to Λ. Moreover, assume that $u \in W^{s,2}(\mathbb{R}^n)$ is a weak solution of the equation $L_A^\Phi u = f$ in Ω. Then for any $0 < \alpha < \min\{2s, 1\}$, we have $u \in C^\alpha_{\mathrm{loc}}(\Omega)$. Furthermore, for all $R > 0$, $x_0 \in \mathbb{R}^n$ such that $B_R(x_0) \Subset \Omega$ and any $\sigma \in (0, 1)$, we have*

$$[u]_{C^\alpha(B_{\sigma R}(x_0))} \leq \frac{C}{R^\alpha} \Bigg(R^{-\frac{n}{2}} \|u\|_{L^2(B_R(x_0))} + R^{2s} \int_{\mathbb{R}^n \setminus B_R(x_0)} \frac{|u(y)|}{|x_0 - y|^{n+2s}} \, dy + R^{2s} \|f\|_{L^\infty(B_R(x_0))} \Bigg),$$

where $C = C(n, s, \Lambda, \alpha, \sigma) > 0$ and

$$[u]_{C^\alpha(B_{\sigma R}(x_0))} := \sup_{\substack{x,y \in B_{\sigma R}(x_0) \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

Remark 4.2. In [35], Theorem 4.1 is proved in the more general context of so-called local weak solutions; see [35, Section 1.1] for a precise definition. From [36, Lemma 3.5] it follows that any local weak solution is a weak solution in our sense, so that Theorem 4.1 indeed follows from [35, Theorem 1.1]. Moreover, it is immediate from the definition of local weak solutions in [35, Section 1.1] that for any local weak solution u of the equation $L_A^\Phi u = 0$ in Ω and any constant $c \in \mathbb{R}$, $u - c$ is also a local weak solution of the same equation. Therefore, in the setting of Theorem 4.1 for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of $L_A^\Phi u = 0$ in Ω and any $c \in \mathbb{R}$, we have the estimate

$$[u]_{C^\alpha(B_{\sigma R}(x_0))} \leq \frac{C}{R^\alpha}\left(R^{-\frac{n}{2}}\|u-c\|_{L^2(B_R(x_0))} + R^{2s}\int_{\mathbb{R}^n\setminus B_R(x_0)} \frac{|u(y)-c|}{|x_0-y|^{n+2s}}\,dy\right),$$

where $C = C(n,s,\Lambda,\alpha,\sigma) > 0$.

4.2. Higher integrability of U for small exponents

For technical reasons, we also study equations with a more general right-hand side than in (1.1).

Definition. Let $2_\star := \frac{2n}{n+2s}$. Given $f \in L^{2_\star}(\Omega)$ and $g \in W^{s,2}(\mathbb{R}^n)$, we say that $u \in W^{s,2}(\mathbb{R}^n)$ is a weak solution of the equation $L_A^\Phi u = (-\Delta)^s g + f$ in Ω if

$$\begin{aligned}&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{A(x,y)}{|x-y|^{n+2s}}\Phi(u(x)-u(y))(\varphi(x)-\varphi(y))\,dy\,dx\\ &\quad = C_{n,s}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{g(x)-g(y)}{|x-y|^{n+2s}}(\varphi(x)-\varphi(y))\,dy\,dx + \int_\Omega f\varphi\,dx \quad \forall \varphi \in W_0^{s,2}(\Omega).\end{aligned}$$

Here $(-\Delta)^s g$ is the fractional Laplacian of g (see Section 4.3) and $C_{n,s}$ is a constant depending on n and s whose exact value is not important for our purposes.

Throughout this work, whenever we deal with functions u and g as in the above definition, for $(x,y) \in \mathbb{R}^{2n}$ with $x \neq y$ we define the functions

$$U(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{s+\theta}}, \quad G(x,y) := \frac{|g(x)-g(y)|}{|x-y|^{s+\theta}}.$$

The following higher integrability result is essentially given by [28, Theorem 6.1], where it is stated under the stronger assumptions that the equation holds on the whole space $\mathbb{R}^n$ and that g is higher differentiable and integrable in the whole $\mathbb{R}^n$. Nevertheless, in [28] the equation in only used to prove the Caccioppoli-type inequality [28, Theorem 3.1], where the equation is tested with test functions that are supported in the ball where the estimate is proved. Therefore, it is enough to assume that the equation holds locally. Moreover, as indicated by the estimate below, in comparison with [28] it is also sufficient to prescribe the higher differentiability and integrability on g locally.

Theorem 4.3. *Let $r > 0$, $x_0 \in \mathbb{R}^n$ and $\sigma_0 > 0$. Moreover, consider a coefficient $A \in \mathcal{L}_0(\Lambda)$ and assume that the Borel function $\Phi\colon \mathbb{R} \to \mathbb{R}$ satisfies*

$$|\Phi(t)| \leq \Lambda t, \quad \Phi(t)t \geq \Lambda^{-1}t^2 \quad \forall t \in \mathbb{R}. \tag{4.1}$$

In addition, assume that $u \in W^{s,2}(\mathbb{R}^n)$ is a weak solution of the equation $L_A^\Phi u = (-\Delta)^s g + f$ in $B_{2r}(x_0)$. Then there exist small enough positive constants $\gamma = \gamma(n,s,\Lambda,\sigma_0) \in (0,\frac{s}{2})$

and $\sigma = \sigma(n, s, \Lambda, \sigma_0) \in (0, \sigma_0)$ *such that if* $f \in L^{2_\star+\sigma_0}(B_{2r}(x_0))$ *and* $g \in W^{s,2}(\mathbb{R}^n) \cap W^{s_\gamma,2+\sigma_0}(B_{2r}(x_0))$ *for* $s_\gamma := s + \gamma(1 - \frac{2}{2+\sigma_0})$*, then*

$$\begin{aligned}\left(\fint_{\mathcal{B}_r(x_0)} U_\gamma^{2+\sigma}\, d\mu_\gamma\right)^{\frac{1}{2+\sigma}} &\\ \leq C \sum_{k=1}^{\infty} 2^{-k(s-\gamma)} &\left(\fint_{\mathcal{B}_{2^k r}(x_0)} U_\gamma^2\, d\mu_\gamma\right)^{\frac{1}{2}} + C\left(\fint_{\mathcal{B}_{2r}(x_0)} G_\gamma^{2+\sigma_0}\, d\mu_\gamma\right)^{\frac{1}{2+\sigma_0}}\\ &+ C \sum_{k=1}^{\infty} 2^{-k(s-\gamma)}\left(\fint_{\mathcal{B}_{2^k r}(x_0)} G_\gamma^2\, d\mu_\gamma\right)^{\frac{1}{2}}\\ &+ C r^{s-\gamma}\left(\fint_{\mathcal{B}_{2r}(x_0)} F^{2_\star+\sigma_0}\, d\mu_\gamma\right)^{\frac{1}{2_\star+\sigma_0}},\end{aligned}$$

where $C = C(n, s, \Lambda, \sigma_0) > 0$*. Here we denote*

$$U_\gamma(x,y) := \frac{|u(x)-u(y)|}{|x-y|^{s+\gamma}}, \quad G_\gamma(x,y) := \frac{|g(x)-g(y)|}{|x-y|^{s+\gamma}}, \quad F(x,y) := f(x),$$

while the measure μ_γ *is defined on measurable sets* $E \subset \mathbb{R}^{2n}$ *by*

$$\mu_\gamma(E) := \int_E \frac{dx\, dy}{|x-y|^{n-2\gamma}}.$$

We note that the assumptions in (4.1) are clearly implied by the assumptions $\Phi(0) = 0$, (1.5) and (1.6) which are used in our main results. Since working with the measure μ_γ and the functions U_γ and G_γ is inconvenient for us, we note that the right-hand side of the estimate from Theorem 4.3 can be rewritten in terms of the measure μ and the functions U and G. More precisely, by using the relevant definitions and taking into account Lemma 3.1, it is straightforward to deduce from Theorem 4.3 the following version of the estimate in Theorem 4.3 for a different C as in Theorem 4.3 depending only on n, s, Λ, θ, γ and σ_0:

$$\begin{aligned}\left(\fint_{\mathcal{B}_r(x_0)} U_\gamma^{2+\sigma}\, d\mu_\gamma\right)^{\frac{1}{2+\sigma}} &\\ \leq C r^{\theta-\gamma}\Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} &\left(\fint_{\mathcal{B}_{2^k r}(x_0)} U^2\, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{2r}(x_0)} G^{2+\sigma_0}\, d\mu\right)^{\frac{1}{2+\sigma_0}}\\ &+ \sum_{k=1}^{\infty} 2^{-k(s-\theta)}\left(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2\, d\mu\right)^{\frac{1}{2}}\Bigg)\\ + C r^{s-\gamma}&\left(\fint_{B_{2r}(x_0)} f^{2_\star+\sigma_0}\, dx\right)^{\frac{1}{2_\star+\sigma_0}}. \qquad (4.2)\end{aligned}$$

4.3. $H^{2s,p}$ estimates for the fractional Laplacian

For any regular enough function $u\colon \mathbb{R}^n \to \mathbb{R}$ and $s \in (0,1)$, the fractional Laplacian of u is formally defined by

$$(-\Delta)^s u(x) = C_{n,s}\,\text{p.v.} \int_{\mathbb{R}^n} \frac{u(x)-u(y)}{|x-y|^{n+2s}}\,dy,$$

where, as in Section 4.2, $C_{n,s}$ is a certain constant depending on n and s. In other words, $(-\Delta)^s$ corresponds to the operator L_A^{Φ} in the special case when $\Phi(t) = t$ and $A = C_{n,s}$.

The following local regularity result for weak solutions of the Dirichlet problem associated to the fractional Laplacian is essentially proved in [3]; see also [25]. The main idea is to multiply the solution by an appropriate cutoff function in order to reduce the problem to a corresponding one which is posed on the whole space $\mathbb{R}^n$, for which the desired estimate can be inferred by classical techniques from Fourier analysis; see [25, Lemma 3.5]. We note that while in [3] estimate (4.4) is not explicitly stated, it can be deduced by keeping track of the estimates in the proofs in [3]. Also, for a formal definition of weak solutions to nonlocal Dirichlet problems as considered below, we refer to [36, Section 4].

Theorem 4.4. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, $s \in (0,1)$ and $p \in (\frac{2n}{n+2s}, \infty)$. If $f \in L^p(\Omega) \cap L^2(\Omega)$, then the unique weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the Dirichlet problem*

$$\begin{cases} (-\Delta)^s u = f & \text{in } \Omega, \\ u = 0 & \text{a.e. in } \mathbb{R}^n \setminus \Omega, \end{cases} \tag{4.3}$$

belongs to $H^{2s,p}_{\mathrm{loc}}(\Omega)$. Moreover, for any open set $\Omega' \Subset \Omega$, we have the estimate

$$\|u\|_{H^{2s,p}(\Omega')} \leq C \|f\|_{L^p(\Omega)}, \tag{4.4}$$

where $C = C(n,s,p,\Omega',\Omega) > 0$.

Also, for some more local and global regularity results for the fractional Laplacian, we refer to [22, 37].

5. Comparison estimates

The following lemma relates the nonlocal tail of a function u to the corresponding function U.

Lemma 5.1. *Let $R > 0$ and $x_0 \in \mathbb{R}^n$. For any function $u \in W^{s,2}(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n \setminus B_R(x_0)} \frac{|u(y) - \bar{u}_{B_R(x_0)}|}{|x_0 - y|^{n+2s}}\,dy \leq C R^{-s+\theta} \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\,d\mu \right)^{\frac{1}{2}}, \tag{5.1}$$

where $C = C(n,s,\theta) > 0$.

Proof. First of all, splitting the integral on the left-hand side into annuli yields

$$\begin{aligned}\int_{\mathbb{R}^n\setminus B_R(x_0)} \frac{|u(y)-\bar{u}_{B_R(x_0)}|}{|x_0-y|^{n+2s}}\,dy &= \sum_{j=0}^{\infty}\int_{B_{2^{j+1}R}(x_0)\setminus B_{2^jR}(x_0)} \frac{|u(y)-\bar{u}_{B_R(x_0)}|}{|x_0-y|^{n+2s}}\,dy\\ &\leq \sum_{j=0}^{\infty}(2^jR)^{-n-2s}\int_{B_{2^{j+1}R}(x_0)} |u(y)-\bar{u}_{B_R(x_0)}|\,dy\\ &= C_1\sum_{j=0}^{\infty}(2^jR)^{-2s}\fint_{B_{2^{j+1}R}(x_0)} |u(y)-\bar{u}_{B_R(x_0)}|\,dy,\end{aligned}$$

where $C_1 = C_1(n)$. Using the Cauchy–Schwarz inequality, we deduce

$$\begin{aligned}&\fint_{B_{2^{j+1}R}(x_0)} |u(y)-\bar{u}_{B_R(x_0)}|\,dy\\ &\quad\leq \fint_{B_{2^{j+1}R}(x_0)} |u(y)-\bar{u}_{B_{2^{j+1}R}(x_0)}|\,dy + \sum_{k=0}^{j}|\bar{u}_{B_{2^{k+1}R}(x_0)}-\bar{u}_{B_{2^kR}(x_0)}|\\ &\quad\leq \fint_{B_{2^{j+1}R}(x_0)} |u(y)-\bar{u}_{B_{2^{j+1}R}(x_0)}|\,dy + 2^n\sum_{k=0}^{j}\fint_{B_{2^{k+1}R}(x_0)}|u(y)-\bar{u}_{B_{2^{k+1}R}(x_0)}|\,dy\\ &\quad\leq 2^{n+1}\sum_{k=1}^{j+1}\fint_{B_{2^kR}(x_0)}|u(y)-\bar{u}_{B_{2^kR}(x_0)}|\,dy\\ &\quad\leq 2^{n+1}\sum_{k=1}^{j+1}\left(\fint_{B_{2^kR}(x_0)}|u(y)-\bar{u}_{B_{2^kR}(x_0)}|^2\,dy\right)^{\frac{1}{2}}.\end{aligned}$$

In order to further estimate the right-hand side of the previous computation, we use the fractional Poincaré inequality (Lemma 2.1) in order to obtain

$$\begin{aligned}\fint_{B_{2^kR}(x_0)}|u(y)-\bar{u}_{B_{2^kR}(x_0)}|^2\,dx &\leq C_2(2^kR)^{2s}\int_{B_{2^kR}(x_0)}\fint_{B_{2^kR}(x_0)}\frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dy\,dx\\ &= C_3(2^kR)^{2(s+\theta)}\fint_{\mathcal{B}_{2^kR}(x_0)}U^2\,d\mu,\end{aligned}$$

where C_2 and C_3 depend only on n, s and θ. Combining the previous two calculations leads to

$$\fint_{B_{2^{j+1}R}(x_0)}|u(y)-\bar{u}_{B_R(x_0)}|\,dy \leq C_4\sum_{k=1}^{j+1}(2^kR)^{(s+\theta)}\left(\fint_{\mathcal{B}_{2^kR}(x_0)}U^2\,d\mu\right)^{\frac{1}{2}},$$

where $C_4 = C_4(n, s, \theta) > 0$. Next, combining the previous estimate with the first display in this proof yields

$$\int_{\mathbb{R}^n \setminus B_R(x_0)} \frac{|u(y) - \bar{u}_{B_R(x_0)}|}{|x_0 - y|^{n+2s}}\, dy \\ \leq C_1 C_4 R^{-s+\theta} \sum_{j=0}^{\infty} \sum_{k=1}^{j+1} 2^{-2sj} 2^{k(s+\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}}. \tag{5.2}$$

By reversing the order of summation, we obtain

$$\begin{aligned} &\sum_{j=0}^{\infty} \sum_{k=1}^{j+1} 2^{-2sj} 2^{k(s+\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}} \\ &\quad = \sum_{k=1}^{\infty} 2^{k(s+\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}} \sum_{j=k-1}^{\infty} 2^{-2sj} \\ &\quad \leq 4^{2s} \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}} \sum_{j=1}^{\infty} 2^{-2sj}. \end{aligned}$$

Since the sum $\sum_{j=1}^{\infty} 2^{-2sj}$ is finite, we conclude that

$$\sum_{j=0}^{\infty} \sum_{k=1}^{j+1} 2^{-2sj} 2^{k(s+\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}} \leq C_5 \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k R}(x_0)} U^2\, d\mu \right)^{\frac{1}{2}},$$

where $C_5 = C_5(s) > 0$. Finally, by combining the last display with (5.2), we arrive at (5.1). ■

A crucial tool for the proof of the higher integrability of U is given by the following comparison estimate. Essentially, it will allow us to transfer some regularity from the solution of a more well-behaved equation to the solution of the original equation.

Proposition 5.2. *Let $x_0 \in \mathbb{R}^n$, $r > 0$, $g \in W^{s,2}(\mathbb{R}^n) \cap W^{s_\theta, 2+\sigma_0}(B_{2r}(x_0))$, $f \in L^{2_\star+\sigma_0}(B_{2r}(x_0))$, $\tilde{f} \in L^{2_\star}(B_{2r}(x_0))$ and $A \in \mathcal{L}_0(\Lambda)$. In addition, assume that Φ satisfies conditions (1.5) and (1.6). Moreover, let $u \in W^{s,2}(\mathbb{R}^n)$ be a weak solution of the equation*

$$L_A^\Phi u = (-\Delta)^s g + f \quad \text{in } B_{2r}(x_0), \tag{5.3}$$

and let $v \in W^{s,2}(\mathbb{R}^n)$ be the unique weak solution of the Dirichlet problem

$$\begin{cases} L_{\tilde{A}}^\Phi v = \tilde{f} & \text{in } B_{2r}(x_0), \\ v = u & \text{a.e. in } \mathbb{R}^n \setminus B_{2r}(x_0), \end{cases} \tag{5.4}$$

where $\tilde{A}$ is another coefficient of class $\mathcal{L}_0(\Lambda)$ such that $\tilde{A} = A$ in $\mathbb{R}^{2n} \setminus \mathcal{B}_r(x_0)$. Then the function $w := u - v \in W_0^{s,2}(B_{2r}(x_0))$ satisfies

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{(w(x)-w(y))^2}{|x-y|^{n+2s}}\,dy\,dx \\
&\quad\leq C\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}\mu(\mathcal{B}_r(x_0))\Bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)} U^2\,d\mu\Big)^{\frac{1}{2}}\Bigg)^2 \\
&\qquad + C\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}\mu(\mathcal{B}_r(x_0))\Big(\fint_{\mathcal{B}_{2r}(x_0)} G^{2+\sigma_0}\,d\mu\Big)^{\frac{2}{2+\sigma_0}} \\
&\qquad + C(\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}+1)\mu(\mathcal{B}_r(x_0))\Bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2\,d\mu\Big)^{\frac{1}{2}}\Bigg)^2 \\
&\qquad + C\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}r^{2(s-\theta)}\mu(\mathcal{B}_r(x_0))\Big(\fint_{B_{2r}(x_0)} |f|^{2_\star+\sigma_0}\,dx\Big)^{\frac{2}{2_\star+\sigma_0}} \\
&\qquad + Cr^{2(s-\theta)}\mu(\mathcal{B}_r(x_0))\Big(\fint_{B_{2r}(x_0)} |f-\tilde{f}|^{2_\star}\,dx\Big)^{\frac{2}{2_\star}},
\end{aligned}$$

where $C = C(n,s,\theta,\Lambda,\sigma_0) > 0$ and

$$\omega(A-\tilde{A},r,x_0) := \fint_{B_r(x_0)}\fint_{B_r(x_0)} |A(x,y)-\tilde{A}(x,y)|\,dy\,dx.$$

Proof. First of all, note that the function v that uniquely solves (5.4) exists by [36, Proposition 4.1]. Using w as a test function in (5.4) and also in (5.3), using (1.6) and taking into account that $A(x,y) = \tilde{A}(x,y)$ whenever $(x,y) \notin \mathcal{B}_r(x_0)$, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{(w(x)-w(y))^2}{|x-y|^{n+2s}}\,dy\,dx \\
&\quad\leq \Lambda\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \tilde{A}(x,y)\frac{\big((u(x)-u(y))-(v(x)-v(y))\big)^2}{|x-y|^{n+2s}}\,dy\,dx \\
&\quad\leq \Lambda^2\Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \tilde{A}(x,y)\frac{\Phi(u(x)-u(y))(w(x)-w(y))}{|x-y|^{n+2s}}\,dy\,dx \\
&\qquad\qquad - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \tilde{A}(x,y)\frac{\Phi(v(x)-v(y))(w(x)-w(y))}{|x-y|^{n+2s}}\,dy\,dx\Big) \\
&\quad= \Lambda^2\Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (\tilde{A}(x,y)-A(x,y))\frac{\Phi(u(x)-u(y))(w(x)-w(y))}{|x-y|^{n+2s}}\,dy\,dx \\
&\qquad\qquad + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} A(x,y)\frac{\Phi(u(x)-u(y))(w(x)-w(y))}{|x-y|^{n+2s}}\,dy\,dx \\
&\qquad\qquad - \int_{B_{2r}(x_0)} \tilde{f}(x)w(x)\,dx\Big)
\end{aligned}$$

$$= \Lambda^2 \Bigg(\underbrace{\int_{B_r(x_0)} \int_{B_r(x_0)} (\tilde{A}(x,y) - A(x,y)) \frac{\Phi(u(x)-u(y))(w(x)-w(y))}{|x-y|^{n+2s}} \, dy \, dx}_{=:I_1}$$

$$+ \underbrace{C_{n,s} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{(g(x)-g(y))(w(x)-w(y))}{|x-y|^{n+2s}} \, dy \, dx}_{:=I_2}$$

$$+ \underbrace{\int_{B_{2r}(x_0)} (f(x) - \tilde{f}(x)) w(x) \, dx}_{:=I_3} \Bigg).$$

Let $\sigma = \sigma(n,s,\Lambda,\sigma_0) > 0$ and $\gamma = \gamma(n,s,\Lambda,\sigma_0) > 0$ be given by Theorem 4.3. By using (1.5), the Cauchy–Schwarz inequality and then Hölder's inequality with conjugated exponents $\frac{2+\sigma}{\sigma}$ and $\frac{2+\sigma}{2}$, we estimate I_1 as

$$\begin{aligned}
I_1 &\leq \Lambda \left(\int_{\mathcal{B}_r(x_0)} (\tilde{A}(x,y) - A(x,y))^2 U_\gamma^2(x,y) \, d\mu_\gamma \right)^{\frac{1}{2}} \\
&\quad \times \left(\int_{B_r(x_0)} \int_{B_r(x_0)} \frac{(w(x)-w(y))^2}{|x-y|^{n+2s}} \, dy \, dx \right)^{\frac{1}{2}} \\
&\leq C_1 \left(\left(\int_{\mathcal{B}_r(x_0)} |\tilde{A}(x,y) - A(x,y)|^{\frac{4}{\sigma}+2} \, d\mu_\gamma \right)^{\frac{\sigma}{2+\sigma}} \left(\fint_{\mathcal{B}_r(x_0)} U_\gamma^{2+\sigma} \, d\mu_\gamma \right)^{\frac{2}{2+\sigma}} r^{\frac{2n+4\gamma}{2+\sigma}} \right)^{\frac{1}{2}} \\
&\quad \times \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{(w(x)-w(y))^2}{|x-y|^{n+2s}} \, dy \, dx \right)^{\frac{1}{2}},
\end{aligned}$$

where $C_1 = C_1(n,\Lambda,\sigma,\gamma) > 0$. By using Hölder's inequality with conjugated exponents $\frac{n-\gamma}{n-2\gamma}$ and $\frac{n-\gamma}{\gamma}$, we obtain

$$\begin{aligned}
&\int_{\mathcal{B}_r(x_0)} |\tilde{A}(x,y) - A(x,y)|^{\frac{4}{\sigma}+2} \, d\mu_\gamma \\
&\quad \leq \left(\int_{B_r(x_0)} \int_{B_r(x_0)} |\tilde{A}(x,y) - A(x,y)|^{(\frac{4}{\sigma}+2)(\frac{n-\gamma}{\gamma})} \, dy \, dx \right)^{\frac{\gamma}{n-\gamma}} \\
&\quad \quad \times \left(\int_{B_r(x_0)} \int_{B_r(x_0)} \frac{dy \, dx}{|x-y|^{n-\gamma}} \right)^{\frac{n-2\gamma}{n-\gamma}} \\
&\quad \leq (2\Lambda)^{\frac{4}{\sigma}+2} \left(\fint_{B_r(x_0)} \fint_{B_r(x_0)} |\tilde{A}(x,y) - A(x,y)| \, dy \, dx \right)^{\frac{\gamma}{n-\gamma}} \\
&\quad \quad \times C_2 r^{2n\frac{\gamma}{n-\gamma}} \mu_{\gamma/2}(\mathcal{B}_r(x_0))^{\frac{n-2\gamma}{n-\gamma}} \\
&\quad \leq C_3 \omega(A - \tilde{A}, r, x_0)^{\frac{\gamma}{n-\gamma}} r^{\frac{2n\gamma}{n-\gamma} + (n+\gamma)\frac{n-2\gamma}{n-\gamma}},
\end{aligned}$$

where $C_2 = C_2(n,s,\gamma) > 0$ and $C_3 = C_3(n,s,\gamma,\sigma,\Lambda) > 0$.

By combining the last two displays with estimate (4.2) and the fact that

$$\frac{\sigma}{2+\sigma}\Big(\frac{2n\gamma}{n-\gamma}+(n+\gamma)\frac{n-2\gamma}{n-\gamma}\Big)+\frac{2n+4\gamma}{2+\sigma}+2(\theta-\gamma)=n+2\theta,$$

we arrive at

$$\begin{aligned}
I_1 \leq C_4\Bigg(&\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}\mu(\mathcal{B}_r(x_0))\Bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)}U^2\,d\mu\Big)^{\frac{1}{2}}\Bigg)^2\\
&+\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}\mu(\mathcal{B}_r(x_0))\Big(\fint_{\mathcal{B}_{2r}(x_0)}G^{2+\sigma_0}\,d\mu\Big)^{\frac{2}{2+\sigma_0}}\\
&+\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}\mu(\mathcal{B}_r(x_0))\Bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)}G^2\,d\mu\Big)^{\frac{1}{2}}\Bigg)^2\\
&+\omega(A-\tilde{A},r,x_0)^{\frac{\gamma}{n-\gamma}}r^{2(s-\theta)}\mu(\mathcal{B}_r(x_0))\Big(\fint_{B_{2r}(x_0)}|f|^{2_\star+\sigma_0}\,dx\Big)^{\frac{2}{2+\sigma_0}}\Bigg),
\end{aligned}$$

where $C_4=C_4(n,s,\gamma,\theta,\sigma,\sigma_0,\Lambda)>0$.

In order to estimate I_2, we set $g_1:=g-\bar{g}_{B_{2r}(x_0)}$ and split the integral as follows:

$$\begin{aligned}
I_2 &\leq \int_{B_{4r}(x_0)}\int_{B_{4r}(x_0)}\frac{|g(x)-g(y)|\,|w(x)-w(y)|}{|x-y|^{n+2s}}\,dy\,dx\\
&\quad+2\int_{B_{2r}(x_0)}\int_{\mathbb{R}^n\setminus B_{4r}(x_0)}\frac{|g_1(x)-g_1(y)|\,|w(x)|}{|x-y|^{n+2s}}\,dy\,dx\\
&\leq \underbrace{\int_{B_{4r}(x_0)}\int_{B_{4r}(x_0)}\frac{|g(x)-g(y)|\,|w(x)-w(y)|}{|x-y|^{n+2s}}\,dy\,dx}_{=:I_{2,1}}\\
&\quad+2\underbrace{\int_{B_{2r}(x_0)}\int_{\mathbb{R}^n\setminus B_{4r}(x_0)}\frac{|g_1(x)|\,|w(x)|}{|x-y|^{n+2s}}\,dy\,dx}_{=:I_{2,2}}\\
&\quad+2\underbrace{\int_{B_{2r}(x_0)}\int_{\mathbb{R}^n\setminus B_{4r}(x_0)}\frac{|g_1(y)|\,|w(x)|}{|x-y|^{n+2s}}\,dy\,dx}_{=:I_{2,3}}.
\end{aligned}$$

By using the Cauchy–Schwarz inequality, we estimate $I_{2,1}$ as

$$I_{2,1}\leq C_5\mu(\mathcal{B}_r(x_0))^{\frac{1}{2}}\Big(\fint_{\mathcal{B}_{4r}(x_0)}G^2\,d\mu\Big)^{\frac{1}{2}}\Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{(w(x)-w(y))^2}{|x-y|^{n+2s}}\,dy\,dx\Big)^{\frac{1}{2}},$$

where $C_5=C_5(n,\theta)>0$. For any $x\in B_{2r}(x_0)$ and any $y\in\mathbb{R}^n\setminus B_{4r}(x_0)$ we have

$$\begin{aligned}
|x_0-y|\leq|x_0-x|+|x-y|&<\Big(\frac{2r}{|x-y|}+1\Big)|x-y|\\
&\leq\Big(\frac{2r}{2r}+1\Big)|x-y|=2|x-y|. \qquad (5.5)
\end{aligned}$$

Moreover, in view of integration in polar coordinates we have

$$\int_{\mathbb{R}^n\setminus B_{4r}(x_0)} \frac{dy}{|x_0-y|^{n+2s}} = \int_{\mathbb{R}^n\setminus B_{4r}} \frac{dy}{|y|^{n+2s}} = \frac{C_6}{r^{2s}}, \tag{5.6}$$

where $C_6 = C_6(n,s) > 0$. Therefore, by using (5.5), (5.6), the Cauchy–Schwarz inequality, the fractional Poincaré inequality (Lemma 2.1) and the fractional Friedrichs–Poincaré inequality (Proposition 2.2), we obtain

$$\begin{aligned} I_{2,2} &\leq 2^{n+2s} \int_{B_{2r}(x_0)} \int_{\mathbb{R}^n\setminus B_{4r}(x_0)} \frac{|g_1(x)|\,|w(x)|}{|x_0-y|^{n+2s}}\,dy\,dx \\ &= C_7 r^{-2s} \int_{B_{2r}(x_0)} |g_1(x)|\,|w(x)|\,dx \\ &\leq C_7 r^{-2s} \Big(\int_{B_{2r}(x_0)} |g_1(x)|^2\,dx\Big)^{\frac{1}{2}} \Big(\int_{B_{2r}(x_0)} |w(x)|^2\,dx\Big)^{\frac{1}{2}} \\ &\leq C_8 \mu(\mathcal{B}_r(x_0))^{\frac{1}{2}} \Big(\fint_{\mathcal{B}_{2r}(x_0)} G^2\,d\mu\Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|w(x)-w(y)|^2}{|x-y|^{n+2s}}\,dy\,dx\Big)^{\frac{1}{2}}, \end{aligned}$$

where all constants depend only on n, s, θ and Λ. Next, by (5.5), the Cauchy–Schwarz inequality, the fractional Friedrichs–Poincaré inequality (Proposition 2.2) and Lemma 5.1, we obtain

$$\begin{aligned} I_{2,3} &\leq 2^{n+2s} \Big(\int_{\mathbb{R}^n\setminus B_{2r}(x_0)} \frac{|g_1(y)|}{|x_0-y|^{n+2s}}\,dy\Big)\Big(\int_{B_{2r}(x_0)} |w(x)|\,dx\Big) \\ &\leq C_9 \Big(r^{-s+\theta} \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Big(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2\,d\mu\Big)^{\frac{1}{2}}\Big) |B_{2r}|^{\frac{1}{2}} \Big(\int_{B_{2r}(x_0)} |w(x)|^2\,dx\Big)^{\frac{1}{2}} \\ &\leq C_{10} \mu(\mathcal{B}_r(x_0))^{\frac{1}{2}} \Big(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Big(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2\,d\mu\Big)^{\frac{1}{2}}\Big) \\ &\quad \times \Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|w(x)-w(y)|^2}{|x-y|^{n+2s}}\,dy\,dx\Big)^{\frac{1}{2}}, \end{aligned}$$

C_9 and C_{10} depend only on n, s and θ. Next, by Hölder's inequality and the fractional Sobolev inequality (see [16, Theorem 6.5]), for I_3 we get

$$\begin{aligned} I_3 &\leq \Big(\int_{B_{2r}(x_0)} |f(x)-\tilde{f}(x)|^{2_\star}\,dx\Big)^{\frac{1}{2_\star}} \Big(\int_{B_{2r}(x_0)} |w(x)|^{\frac{2n}{n-2s}}\,dx\Big)^{\frac{n-2s}{2n}} \\ &\leq C_{11} \Big(\int_{B_{2r}(x_0)} |f(x)-\tilde{f}(x)|^{2_\star}\,dx\Big)^{\frac{1}{2_\star}} \Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|w(x)-w(y)|^2}{|x-y|^{n+2s}}\,dy\,dx\Big)^{\frac{1}{2}}, \end{aligned}$$

where $C_{11} = C_{11}(n,s) > 0$. Combining all the above estimates along with squaring both sides of the resulting inequality and then dividing by $\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|w(x)-w(y)|^2}{|x-y|^{n+2s}}\,dy\,dx$ on both sides now finishes the proof. ∎

From now on, we fix some $p \in (2,\infty)$, some $\Lambda \geq 1$ and some coefficient $A \in \mathcal{L}_0(\Lambda)$ that is δ-vanishing in B_{4n}, where $\delta > 0$ remains to be chosen small enough later. Moreover, we fix another number $q \in [2, p)$ and define

$$q^\star := \begin{cases} \dfrac{nq}{n-sq} & \text{if } n > sq, \\ 2p & \text{if } n \leq sq. \end{cases} \tag{5.7}$$

In addition, we choose the number $\sigma_0 > 0$ small enough that $2 + \sigma_0 < \min\{(q+q^\star)/2, p\}$ and set

$$q_0 := \max\{2+\sigma_0, q\} < \min\{(q+q^\star)/2, p\}. \tag{5.8}$$

Furthermore, we fix some $g \in W^{s,2}(\mathbb{R}^n)$ and a weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation

$$L_A^\Phi u = (-\Delta)^s g \quad \text{in } B_{4n} \tag{5.9}$$

and set

$$\begin{aligned} \lambda_0 := M_0 \Bigg(& \sum_{k=1}^\infty 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} U^2 \, d\mu \right)^{\frac{1}{2}} \\ & + \delta^{-1} \sum_{k=1}^\infty 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} G^2 \, d\mu \right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{4n}} G^{q_0} \, d\mu \right)^{\frac{1}{q_0}} \Bigg), \end{aligned} \tag{5.10}$$

where $M_0 \geq 1$ remains to be chosen large enough.

Lemma 5.3. *Let $M > 0$, $x_0 \in B_{\frac{\sqrt{n}}{2}}$, $r \in (0, \frac{\sqrt{n}}{2})$ and $\lambda \geq \lambda_0$. Moreover, consider the coefficient*

$$\tilde{A}(x,y) := \begin{cases} \bar{A}_{3r,x_0,x_0} & \text{if } (x,y) \in \mathcal{B}_{3r}(x_0), \\ A(x,y) & \text{if } (x,y) \notin \mathcal{B}_{3r}(x_0). \end{cases}$$

Then for any $\varepsilon_0 > 0$, there exists some small enough $\delta = \delta(\varepsilon_0, n, s, \theta, \Lambda, M) \in (0,1)$ such that, under the assumptions made above along with

$$\mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x_0) \leq M\lambda^2, \quad \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x_0) \leq M\lambda^{q_0}\delta^{q_0}, \tag{5.11}$$

for the unique weak solution $v \in W^{s,2}(\mathbb{R}^n)$ of the Dirichlet problem

$$\begin{cases} L_{\tilde{A}}^\Phi v = 0 & \text{in } B_{6r}(x_0), \\ v = u & \text{a.e. in } \mathbb{R}^n \setminus B_{6r}(x_0), \end{cases} \tag{5.12}$$

and the function

$$W(x,y) := \frac{|u(x)-v(x)-u(y)+v(y)|}{|x-y|^{s+\theta}}, \quad (x,y) \in \mathbb{R}^{2n},$$

we have

$$\int_{\mathbb{R}^{2n}} W^2 d\mu \leq \varepsilon^2 \lambda^2 \mu(\mathcal{B}_r(x_0)). \tag{5.13}$$

Moreover, the function

$$V(x,y) := \frac{|v(x)-v(y)|}{|x-y|^{s+\theta}}, \quad (x,y) \in \mathbb{R}^{2n}$$

satisfies the estimate

$$\|V\|_{L^\infty(\mathcal{B}_{2r}(x_0),d\mu)} \leq N_0 \lambda \tag{5.14}$$

for some constant $N_0 = N_0(n,s,\theta,\Lambda,M) > 0$.

Proof. Fix $x_0 \in B_{\frac{\sqrt{n}}{2}}$ and $r \in (0, \frac{\sqrt{n}}{2})$ and note that $\tilde{A} = A$ in $\mathbb{R}^{2n} \setminus \mathcal{B}_{3r}(x_0)$. Moreover, since A is δ-vanishing in B_{4n}, we have

$$\omega(A-\tilde{A}, 3r, x_0) = \fint_{B_{3r}(x_0)} \fint_{B_{3r}(x_0)} |A(x,y) - \bar{A}_{3r,x_0,x_0}| \, dy\, dx \leq \delta. \tag{5.15}$$

First, we prove (5.13). Let $m \in \mathbb{N}$ be determined by $2^{m-1} r < \sqrt{n} \leq 2^m r$; note that $m \geq 2$. Then for any $k < m$, by (5.11) we have

$$\begin{aligned} &\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \leq M \lambda^2, \\ &\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^{q_0} \, d\mu \leq M \lambda^{q_0} \delta^{q_0}. \end{aligned} \tag{5.16}$$

On the other hand, in view of (5.10) and the inclusions

$$\mathcal{B}_{2^k \sqrt{n}}(x_0) \subset \mathcal{B}_{2^{k+m-1} 3r}(x_0) \subset \mathcal{B}_{2^k 3\sqrt{n}}(x_0) \subset \mathcal{B}_{2^k 4n},$$

we have

$$\begin{aligned} &\sum_{k=m}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} \\ &= 2^{-(m-1)(s-\theta)} \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^{k+m-1} 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} \\ &\leq \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\frac{\mu(\mathcal{B}_{2^k 4n})}{\mu(\mathcal{B}_{2^k \sqrt{n}})} \fint_{\mathcal{B}_{2^k 4n}} U^2 \, d\mu \right)^{\frac{1}{2}} \\ &= C_1 \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} U^2 \, d\mu \right)^{\frac{1}{2}} \leq C_1 \lambda_0, \end{aligned}$$

where $C_1 = C_1(n,\theta) > 0$. Together with (5.16) and the facts that $\theta < s$ and $\lambda \geq \lambda_0$, we arrive at

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} \\
&\quad \leq \sum_{k=1}^{m-1} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} + \sum_{k=m}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} \\
&\quad \leq M^{\frac{1}{2}} \lambda \sum_{k=1}^{\infty} 2^{-k(s-\theta)} + C_1 \lambda_0 \leq C_2 \lambda, \qquad (5.17)
\end{aligned}$$

where $C_2 = C_2(n,s,\theta,M) > 0$. By similar reasoning to above, we have

$$\sum_{k=m}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^2 \, d\mu \right)^{\frac{1}{2}} \leq C_1 \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} G^2 \, d\mu \right)^{\frac{1}{2}} \leq C_1 \lambda_0 \delta$$

and therefore along with Hölder's inequality,

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^2 \, d\mu \right)^{\frac{1}{2}} \\
&\quad \leq \sum_{k=1}^{m-1} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^{q_0} \, d\mu \right)^{\frac{1}{q_0}} + \sum_{k=m}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^2 \, d\mu \right)^{\frac{1}{2}} \\
&\quad \leq M^{\frac{1}{q_0}} \lambda \delta \sum_{k=1}^{\infty} 2^{-k(s-\theta)} + C_1 \lambda_0 \delta \leq C_2 \lambda \delta. \qquad (5.18)
\end{aligned}$$

By combining (5.15), (5.16), (5.17) and (5.18) with Proposition 5.2, the fact that $2 + \sigma_0 \leq q_0$ and Hölder's inequality, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^{2n}} W^2 d\mu = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{(w(x)-w(y))^2}{|x-y|^{n+2s}} \, dy \, dx \\
&\leq C_3 \omega(A-\tilde{A},3r,x_0)^{\frac{\gamma}{n-\gamma}} \mu(\mathcal{B}_r(x_0)) \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2 \, d\mu \right)^{\frac{1}{2}} \right)^2 \\
&\quad + C_3 \omega(A-\tilde{A},3r,x_0)^{\frac{\gamma}{n-\gamma}} \mu(\mathcal{B}_r(x_0)) \left(\fint_{\mathcal{B}_{6r}(x_0)} G^{q_0} \, d\mu \right)^{\frac{2}{q_0}} \\
&\quad + C_3 (\omega(A-\tilde{A},3r,x_0)^{\frac{\gamma}{n-\gamma}} + 1) \mu(\mathcal{B}_r(x_0)) \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} G^2 \, d\mu \right)^{\frac{1}{2}} \right)^2 \\
&\leq C_4 \mu(\mathcal{B}_r(x_0)) \lambda^2 \delta^{\frac{2\gamma}{n-\gamma}} < \varepsilon^2 \lambda^2 \mu(\mathcal{B}_r(x_0)),
\end{aligned}$$

where $C_4 = C_4(n,s,\theta,\Lambda,M) > 0$ and the last inequality was obtained by choosing δ sufficiently small. This proves (5.13).

Let us now prove estimate (5.14). Since $\tilde{A}$ is constant and therefore continuous in $\mathcal{B}_{3r}(x_0)$, by Theorem 4.1 and Remark 4.2 with $\alpha = s + \theta \in (0, \min\{2s, 1\})$ and $c = \bar{v}_{B_{3r}(x_0)}$, we have

$$\begin{aligned}
&\|V\|^2_{L^\infty(\mathcal{B}_{2r}(x_0), d\mu)} \\
&\quad \leq [v]^2_{C^\alpha(B_{2r}(x_0))} \\
&\quad \leq \frac{C_5}{r^{2(s+\theta)}}\left(r^{-n}\|v_1\|^2_{L^2(B_{3r}(x_0))} + \left(r^{2s}\int_{\mathbb{R}^n \setminus B_{3r}(x_0)} \frac{|v_1(y)|}{|x_0 - y|^{n+2s}}\,dy\right)^2\right) \\
&\quad \leq \frac{C_6}{r^{2(s+\theta)}}\left(r^{2s-n}[v]^2_{W^{s,2}(B_{3r}(x_0))} + \left(r^{2s}\int_{\mathbb{R}^n \setminus B_{3r}(x_0)} \frac{|v_1(y)|}{|x_0 - y|^{n+2s}}\,dy\right)^2\right),
\end{aligned}$$

where $v_1 := v - \bar{v}_{B_{3r}(x_0)}$ and C_5 and C_6 depend only on n, s and Λ. Here we also used Lemma 2.1 in order to obtain the last inequality. By using (5.16) and (5.13), we further estimate the first term on the right-hand side as

$$\begin{aligned}
r^{2s-n}[v]^2_{W^{s,2}(B_{3r}(x_0))} &\leq 2r^{2s-n}\left(\int_{B_{3r}(x_0)}\int_{B_{3r}(x_0)} \frac{|u(x) - u(y)|^2}{|x - y|^{n+2s}}\,dy\,dx\right. \\
&\qquad\qquad \left. + \int_{B_{3r}(x_0)}\int_{B_{3r}(x_0)} \frac{|w(x) - w(y)|^2}{|x - y|^{n+2s}}\,dy\,dx\right) \\
&\leq 2r^{2s-n}\left(\lambda^2\mu(\mathcal{B}_{3r}(x_0)) + \varepsilon^2\lambda^2\mu(\mathcal{B}_r(x_0))\right) \leq C_7\lambda^2 r^{2(s+\theta)},
\end{aligned}$$

where $C_7 = C_7(n, \theta) > 0$. Moreover, by taking into account that $v = u$ in $\mathbb{R}^n \setminus B_{3r}(x_0)$, we split the tail term as

$$\begin{aligned}
&\left(r^{2s}\int_{\mathbb{R}^n \setminus B_{3r}(x_0)} \frac{|v_1(y)|}{|x_0 - y|^{n+2s}}\,dy\right)^2 \\
&\quad \leq 2\underbrace{\left(r^{2s}\int_{\mathbb{R}^n \setminus B_{3r}(x_0)} \frac{|u(y) - \bar{u}_{B_{3r}(x_0)}|}{|x_0 - y|^{n+2s}}\,dy\right)^2}_{=:J_1} \\
&\qquad + 2\underbrace{\left(r^{2s}\int_{\mathbb{R}^n \setminus B_{3r}(x_0)} \frac{|\bar{u}_{B_{3r}(x_0)} - \bar{v}_{B_{3r}(x_0)}|}{|x_0 - y|^{n+2s}}\,dy\right)^2}_{=:J_2}.
\end{aligned}$$

In view of Lemma 5.1 and (5.17), for J_1 we have

$$J_1 \leq C_8\left(r^{s+\theta}\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\left(\fint_{\mathcal{B}_{2^k 3r}(x_0)} U^2\,d\mu\right)^{\frac{1}{2}}\right)^2 \leq C_8\lambda^2 r^{2(s+\theta)}.$$

Moreover, by using (5.6), applying Jensen's inequality twice, using the fractional Friedrichs–Poincaré inequality with respect to $w = u - v \in W_0^{s,2}(B_{2r}(x_0))$ and again taking into account (5.16) and (5.13), for J_2 we obtain

$$
\begin{aligned}
J_2 &= C_9 \left| \fint_{B_{3r}(x_0)} (u(x) - \bar{v}_{B_{3r}(x_0)})\, dx \right|^2 \\
&\leq C_9 \fint_{B_{3r}(x_0)} |u(x) - \bar{v}_{B_{3r}(x_0)}|^2\, dx \\
&\leq C_9 \fint_{B_{3r}(x_0)} \fint_{B_{3r}(x_0)} |u(x) - v(y)|^2\, dy\, dx \\
&\leq 2C_{10} \Big(r^{-2n} \int_{B_{3r}(x_0)} \int_{B_{3r}(x_0)} |u(x) - u(y)|^2\, dy\, dx + r^{-n} \int_{B_{3r}(x_0)} |w(y)|^2\, dy \Big) \\
&\leq C_{11} r^{2s-n} \Big(\int_{B_{3r}(x_0)} \int_{B_{3r}(x_0)} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}}\, dy\, dx \\
&\qquad + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|w(x) - w(y)|^2}{|x-y|^{n+2s}}\, dy\, dx \Big) \\
&\leq C_{12} \lambda^2 r^{2(s+\theta)},
\end{aligned}
$$

where all the constants depend only on n, s, θ, Λ and M. Combining the last five displays now shows that estimate (5.14) holds for some $N_0 = N_0(n, s, \theta, \Lambda, M) > 0$, which finishes the proof. ∎

6. Good-λ inequalities

6.1. Diagonal good-λ inequalities

The following result is a consequence of the above approximation lemma and roughly speaking shows that if the set where the maximal function of U is very large has a large enough density in a ball, then in this ball the maximal functions of U and G cannot be too small.

Lemma 6.1. *There is a constant $N_d = N_d(n, s, \theta, \Lambda) \geq 1$, such that the following holds. For any $\varepsilon > 0$ and any $\kappa > 0$ there exists some small enough $\delta = \delta(\varepsilon, \kappa, n, s, \theta, \Lambda) \in (0, 1)$, such that for any $\lambda \geq \lambda_0$, any $r \in (0, \frac{\sqrt{n}}{2})$ and any point $x_0 \in Q_1$ with*

$$
\mu\big(\big\{(x, y) \in \mathcal{B}_r(x_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_d^2 \lambda^2\big\}\big) \geq \kappa \varepsilon \mu(\mathcal{B}_r(x_0)), \tag{6.1}
$$

we have

$$
\begin{aligned}
\mathcal{B}_r(x_0) \subset \big\{(x, y) \in \mathcal{B}_r(x_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > \lambda^2\big\} \\
\cap \big\{(x, y) \in \mathcal{B}_r(x_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x, y) > \lambda^{q_0} \delta^{q_0}\big\},
\end{aligned} \tag{6.2}
$$

Proof. Let $\varepsilon_0 > 0$ and $M > 0$ be chosen and consider the corresponding $\delta = \delta(\varepsilon_0, n, s, \theta, \Lambda, M) \in (0, 1)$ given by Lemma 5.3. Fix $\varepsilon, \kappa > 0$, $r \in (0, \frac{\sqrt{n}}{2})$, $x_0 \in Q_1$ and assume that

(6.1) holds, but that (6.2) is false, so that there exists a point $(x', y') \in \mathcal{B}_r(x_0)$ such that

$$\mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x', y') \leq \lambda^2, \quad \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x', y') \leq \lambda^{q_0} \delta^{q_0}.$$

Therefore, for any $\rho > 0$ we have

$$\fint_{\mathcal{B}_\rho(x', y')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \leq \lambda^2, \quad \fint_{\mathcal{B}_\rho(x', y')} \chi_{\mathcal{B}_{4n}} G^{q_0} \, d\mu \leq \lambda^{q_0} \delta^{q_0}. \tag{6.3}$$

Note that for any $\rho \geq r$ we have $\mathcal{B}_\rho(x_0) \subset \mathcal{B}_{2\rho}(x', y') \subset \mathcal{B}_{3\rho}(x_0)$. Together with (6.3), this observation yields

$$\begin{aligned}
\fint_{\mathcal{B}_\rho(x_0)} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu &\leq \frac{\mu(\mathcal{B}_{2\rho}(x', y'))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x', y')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \\
&\leq \frac{\mu(\mathcal{B}_{3\rho}(x_0))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x', y')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \\
&\leq 3^{n+2\theta} \lambda^2
\end{aligned}$$

and similarly,

$$\begin{aligned}
\fint_{\mathcal{B}_\rho(x_0)} \chi_{\mathcal{B}_{4n}} G^{q_0} \, d\mu &\leq \frac{\mu(\mathcal{B}_{2\rho}(x', y'))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x', y')} \chi_{\mathcal{B}_{4n}} G^{q_0} \, d\mu \\
&\leq \frac{\mu(\mathcal{B}_{3\rho}(x_0))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x', y')} \chi_{\mathcal{B}_{4n}} G^{q_0} \, d\mu \leq 3^{n+2\theta} \lambda^{q_0} \delta^{q_0}
\end{aligned}$$

so that U and G satisfy condition (5.11) with $M = 3^{n+2\theta}$. Therefore, by Lemma 5.3 the unique weak solution $v \in W^{s,2}(\mathbb{R}^n)$ of the Dirichlet problem

$$\begin{cases} L_{\tilde{A}}^{\Phi} v = 0 & \text{weakly in } B_{6r}(x_0), \\ v = u & \text{a.e. in } \mathbb{R}^n \setminus B_{6r}(x_0), \end{cases}$$

satisfies

$$\int_{\mathbb{R}^{2n}} W^2 \, d\mu \leq \varepsilon_0^2 \lambda^2 \mu(\mathcal{B}_r(x_0)), \tag{6.4}$$

where W is given as in Lemma 5.3. Moreover, by Lemma 5.3 there exists a constant $N_0 = N_0(n, s, \theta, \Lambda) > 0$ such that

$$\|V\|^2_{L^\infty(\mathcal{B}_{2r}(x_0))} \leq N_0^2 \lambda^2. \tag{6.5}$$

Next we define $N_d := (\max\{4N_0^2, 5^{n+2\theta}\})^{1/2} > 1$ and claim that

$$\begin{aligned}
&\left\{(x, y) \in \mathcal{B}_r(x_0) \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_d^2 \lambda^2\right\} \\
&\qquad \subset \left\{(x, y) \in \mathcal{B}_r(x_0) \mid \mathcal{M}_{\mathcal{B}_{2r}(x_0)}(W^2)(x, y) > N_0^2 \lambda^2\right\}.
\end{aligned} \tag{6.6}$$

To see this, assume that

$$(x_1, y_1) \in \left\{x \in \mathcal{B}_r(x_0) \mid \mathcal{M}_{\mathcal{B}_{2r}(x_0)}(W^2)(x, y) \leq N_0^2 \lambda^2\right\}. \tag{6.7}$$

For $\rho < r$, we have $\mathcal{B}_\rho(x_1, y_1) \subset \mathcal{B}_r(x_1, y_1) \subset \mathcal{B}_{2r}(x_0)$, so that together with (6.7) and (6.5) we deduce

$$\begin{aligned}\fint_{\mathcal{B}_\rho(x_1,y_1)} U^2\, d\mu &\leq 2 \fint_{\mathcal{B}_\rho(x_1,y_1)} (W^2 + V^2)\, d\mu \\ &\leq 2 \fint_{\mathcal{B}_\rho(x_1,y_1)} W^2\, d\mu + 2\|V\|^2_{L^\infty(\mathcal{B}_\rho(x_1,y_1))} \\ &\leq 2\mathcal{M}_{\mathcal{B}_{2r}(x_0)}(W^2)(x_1, y_1) + 2\|V\|^2_{L^\infty(\mathcal{B}_{2r}(x_0))} \leq 4N_0^2\lambda^2.\end{aligned}$$

On the other hand, for $\rho \geq r$ we have $\mathcal{B}_\rho(x_1, y_1) \subset \mathcal{B}_{3\rho}(x', y') \subset \mathcal{B}_{5\rho}(x_1, y_1)$, so that (6.3) implies

$$\begin{aligned}\fint_{\mathcal{B}_\rho(x_1,y_1)} \chi_{\mathcal{B}_{4n}} U^2\, d\mu &\leq \frac{\mu(\mathcal{B}_{3\rho}(x', y'))}{\mu(\mathcal{B}_\rho(x_1, y_1))} \fint_{\mathcal{B}_{3\rho}(x',y')} \chi_{\mathcal{B}_{4n}} U^2\, d\mu \\ &\leq \frac{\mu(\mathcal{B}_{5\rho}(x_1, y_1))}{\mu(\mathcal{B}_\rho(x_1, y_1))} \fint_{\mathcal{B}_{3\rho}(x',y')} \chi_{\mathcal{B}_{4n}} U^2\, d\mu \leq 5^{n+2\theta}\lambda^2.\end{aligned}$$

Thus, we have

$$(x_1, y_1) \in \left\{(x, y) \in \mathcal{B}_r(x_0, y_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) \leq N_d^2\lambda^2\right\},$$

which implies (6.6). Now using (6.6), the weak 1–1 estimate from Proposition 3.3 and (6.4), we conclude that there exists some constant $C = C(n, \theta) > 0$ such that

$$\begin{aligned}&\mu\left(\left\{(x, y) \in \mathcal{B}_r(x_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_d^2\lambda^2\right\}\right) \\ &\quad \leq \mu\left(\left\{(x, y) \in \mathcal{B}_r(x_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{2r}(x_0)}(W^2)(x, y) > N_0^2\lambda^2\right\}\right) \\ &\quad \leq \frac{C}{N_0^2\lambda^2} \int_{\mathbb{R}^{2n}} W^2\, d\mu \\ &\quad \leq \frac{C}{N_0^2} \mu(\mathcal{B}_r(x_0))\varepsilon_0^2 < \varepsilon\kappa\mu(\mathcal{B}_r(x_0)),\end{aligned}$$

where the last inequality is obtained by choosing ε_0 and thus also δ sufficiently small. This contradicts (6.1) and thus finishes our proof. ∎

6.2. Off-diagonal reverse Hölder inequalities

Our next goal is to prove an analogue of the above diagonal good-λ inequality on balls that are far away from the diagonal. In order to prove Lemma 6.1, the main tool was given by the comparison estimates from Section 5. Unfortunately, far away from the diagonal the equation cannot be used very efficiently, since the further away we are from the diagonal, the less the estimates available reflect the scale we are working at. In particular, far away from the diagonal no useful comparison estimates are available.

In order to bypass this loss of information, we replace the comparison estimates used in the diagonal setting by certain off-diagonal reverse Hölder inequalities with diagonal correction terms. Although such reverse Hölder inequalities lead to, in some sense, weaker good-λ inequalities than comparison estimates, by using combinatorial techniques and an iteration argument at the end we will nevertheless be able to deduce the desired regularity.

For this reason, from now on we assume that for any $r > 0$, $x_0 \in \mathbb{R}^n$ with $B_r(x_0) \subset B_{4n}$, U satisfies an estimate of the form

$$\left(\fint_{\mathcal{B}_{r/2}(x_0)} U^q \, d\mu\right)^{\frac{1}{q}} \leq C_q \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} U^2 \, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_r(x_0)} G^{q_0} \, d\mu\right)^{\frac{1}{q_0}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2 \, d\mu\right)^{\frac{1}{2}} \Bigg), \tag{6.8}$$

where C_q depends only on q, n, s, θ and Λ.

Proposition 6.2. *Let $r > 0$, $x_0, y_0 \in \mathbb{R}^n$ and suppose that for some $m \in (0, 1]$ we have $\operatorname{dist}(B_r(x_0), B_r(y_0)) \geq mr$. Then we have*

$$\begin{aligned}
&\left(\fint_{\mathcal{B}_r(x_0,y_0)} U^{q^\star} \, d\mu\right)^{\frac{1}{q^\star}} \\
&\quad \leq C_{nd} \left(\fint_{\mathcal{B}_r(x_0,y_0)} U^2 \, d\mu\right)^{\frac{1}{2}} \\
&\quad + C_{nd} \left(\frac{r}{\operatorname{dist}(B_r(x_0), B_r(y_0))}\right)^{s+\theta} \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} U^2 \, d\mu\right)^{\frac{1}{2}} \\
&\qquad + \left(\fint_{\mathcal{B}_{2r}(x_0)} G^{q_0} \, d\mu\right)^{\frac{1}{q_0}} \\
&\qquad + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2 \, d\mu\right)^{\frac{1}{2}} \Bigg) \\
&\quad + C_{nd} \left(\frac{r}{\operatorname{dist}(B_r(x_0), B_r(y_0))}\right)^{s+\theta} \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(y_0)} U^2 \, d\mu\right)^{\frac{1}{2}} \\
&\qquad + \left(\fint_{\mathcal{B}_{2r}(y_0)} G^{q_0} \, d\mu\right)^{\frac{1}{q_0}} \\
&\qquad + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(y_0)} G^2 \, d\mu\right)^{\frac{1}{2}} \Bigg),
\end{aligned}$$

where $C_{nd} = C_{nd}(n, s, \theta, \Lambda, m, q, p) \geq 1$ and $q^\star$ is given by (5.7).

Proof. Choose points $x_1 \in \overline{B}_r(x_0)$ and $y_1 \in \overline{B}_r(y_0)$ such that $\operatorname{dist}(B_r(x_0), B_r(y_0)) = |x_1 - y_1|$. For any $(x, y) \in \mathcal{B}_r(x_0, y_0)$, we obtain

$$\begin{aligned}|x-y| &\leq |x_1 - y_1| + |x_1 - x| + |y_1 - y| \\ &\leq \operatorname{dist}(B_r(x_0), B_r(y_0)) + 2r \leq 3 \operatorname{dist}(B_r(x_0), B_r(y_0))/m.\end{aligned}$$

Together with the definition of $\operatorname{dist}(B_r(x_0), B_r(y_0))$, it follows that for any $(x, y) \in \mathcal{B}_r(x_0, y_0)$, we have

$$1 \leq \frac{|x-y|}{\operatorname{dist}(B_r(x_0), B_r(y_0))} \leq 3/m. \tag{6.9}$$

Therefore, by taking into account the definition of the measure μ, we conclude that

$$\frac{c_1 r^{2n}}{\operatorname{dist}(B_r(x_0), B_r(y_0))^{n-2\theta}} \leq \mu(\mathcal{B}_r(x_0, y_0)) \leq \frac{C_1 r^{2n}}{\operatorname{dist}(B_r(x_0), B_r(y_0))^{n-2\theta}}, \tag{6.10}$$

where $c_1 = c_1(n, m, \theta) \in (0, 1)$ and $C_1 = C_1(n) \geq 1$. By (6.10) and (6.9), we have

$$\begin{aligned}&\left(\fint_{\mathcal{B}_r(x_0,y_0)} U^{q^\star}\, d\mu\right)^{\frac{1}{q^\star}} \\ &\quad \leq \left(\frac{\operatorname{dist}(B_r(x_0), B_r(y_0))^{n-2\theta}}{c_1 r^{2n}} \int_{B_r(x_0)} \int_{B_r(y_0)} \frac{|u(x)-u(y)|^{q^\star}}{|x-y|^{n-2\theta+q^\star(s+\theta)}}\, dy\, dx\right)^{\frac{1}{q^\star}} \\ &\quad \leq C_2 \operatorname{dist}(B_r(x_0), B_r(y_0))^{-(s+\theta)} \left(\fint_{B_r(x_0)} \fint_{B_r(y_0)} |u(x)-u(y)|^{q^\star}\, dy\, dx\right)^{\frac{1}{q^\star}},\end{aligned}$$

where $C_2 = C_2(n, m, \theta) \geq 1$. By using Minkowski's inequality, we further estimate the integral on the right-hand side as

$$\begin{aligned}&\left(\fint_{B_r(x_0)} \fint_{B_r(y_0)} |u(x)-u(y)|^{q^\star}\, dy\, dx\right)^{\frac{1}{q^\star}} \\ &\quad \leq \underbrace{\left(\fint_{B_r(x_0)} |u(x) - \bar{u}_{B_r(x_0)}|^{q^\star}\, dy\, dx\right)^{\frac{1}{q^\star}}}_{=:I_1} + \underbrace{\left(\fint_{B_r(y_0)} |u(x) - \bar{u}_{B_r(y_0)}|^{q^\star}\, dy\, dx\right)^{\frac{1}{q^\star}}}_{=:I_2} \\ &\qquad + \underbrace{|\bar{u}_{B_r(x_0)} - \bar{u}_{B_r(y_0)}|}_{=:I_3}.\end{aligned}$$

By using the fractional Sobolev–Poincaré inequality (Lemma 2.4) and then estimate (6.8), for I_1 we obtain

$$\begin{aligned}I_1 &\leq C_3 r^s \left(\frac{1}{r^n} \int_{B_r(x_0)} \int_{B_r(x_0)} \frac{|u(x)-u(y)|^q}{|x-y|^{n+sq}}\, dy\, dx\right)^{\frac{1}{q}} \\ &= C_4 r^s \left(\frac{r^{2\theta}}{\mu(\mathcal{B}_r(x_0))} \int_{B_r(x_0)} \int_{B_r(x_0)} \frac{|u(x)-u(y)|^q |x-y|^{(q-2)\theta}}{|x-y|^{n-2\theta+q(s+\theta)}}\, dy\, dx\right)^{\frac{1}{q}}\end{aligned}$$

$$
\begin{aligned}
&\leq C_5 r^{s+\theta}\Big(\fint_{\mathcal{B}_r(x_0)} U^q\,d\mu\Big)^{\frac{1}{q}}\\
&\leq C_q C_5 r^{s+\theta}\Big(\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)} U^2\,d\mu\Big)^{\frac{1}{2}}+\Big(\fint_{\mathcal{B}_{2r}(x_0)} G^{q_0}\,d\mu\Big)^{\frac{1}{q_0}}\\
&\qquad+\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(x_0)} G^2\,d\mu\Big)^{\frac{1}{2}}\Big),
\end{aligned}
$$

where C_3, C_4 and C_5 depend only on n, s and θ. In the same way, for I_2 we have

$$
\begin{aligned}
I_2 \leq C_q C_5 r^{s+\theta}\Big(&\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(y_0)} U^2\,d\mu\Big)^{\frac{1}{2}}+\Big(\fint_{\mathcal{B}_{2r}(y_0)} G^{q_0}\,d\mu\Big)^{\frac{1}{q_0}}\\
&+\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k r}(y_0)} G^2\,d\mu\Big)^{\frac{1}{2}}\Big).
\end{aligned}
$$

Finally, by the Cauchy–Schwarz inequality, (6.10) and (6.9), for I_3 we have

$$
\begin{aligned}
I_3 &\leq \Big(\fint_{B_r(x_0)}\fint_{B_r(y_0)} |u(x)-u(y)|^2\,dy\,dx\Big)^{\frac{1}{2}}\\
&\leq \Big(\frac{C_1}{\operatorname{dist}(B_r(x_0),B_r(y_0))^{n-2\theta}\mu(\mathcal{B}_r(x_0,y_0))}\int_{B_r(x_0)}\int_{B_r(y_0)} |u(x)-u(y)|^2\,dy\,dx\Big)^{\frac{1}{2}}\\
&\leq C_6\Big(\fint_{\mathcal{B}_r(x_0,y_0)} |u(x)-u(y)|^2\,d\mu\Big)^{\frac{1}{2}}\\
&\leq C_7 \operatorname{dist}(B_r(x_0),B_r(y_0))^{s+\theta}\Big(\fint_{\mathcal{B}_r(x_0,y_0)} U^2\,d\mu\Big)^{\frac{1}{2}},
\end{aligned}
$$

where $C_6 = C_6(n,m,\theta) \geq 1$ and $C_7 = C_7(n,m,\theta) \geq 1$. The claim now follows by combining the last five displays, so that the proof is finished. ■

6.3. Off-diagonal good-λ inequalities

In what follows, fix some $\varepsilon \in (0,1)$ to be chosen small enough and set

$$
N_{\varepsilon,q} := \frac{C_{nd} C_{s,\theta} N_d 10^{10n}}{\varepsilon^{1/q^\star}}, \tag{6.11}
$$

where $N_d = N_d(n,s,\theta,q,\Lambda) \geq 1$ is given by Lemma 6.1, $C_{nd} = C_{nd}(n,s,\theta,\Lambda,m,q) \geq 1$ is given by Proposition 6.2 with m to be chosen and

$$
1 \leq C_{s,\theta} := \sum_{k=1}^{\infty} 2^{-k(s-\theta)} < \infty. \tag{6.12}
$$

Moreover, for all $r \in (0, \frac{\sqrt{n}}{2})$ and all $(x_0, y_0) \in \mathcal{Q}_1$ we define

$$\tilde{\phi}(r, x_0, y_0) := \frac{r}{\operatorname{dist}(B_{\frac{r}{2}}(x_0), B_{\frac{r}{2}}(y_0))}. \tag{6.13}$$

Lemma 6.3. *For any $\lambda \geq \lambda_0$, $r \in (0, \frac{\sqrt{n}}{2})$ and any point $(x_0, y_0) \in \mathcal{Q}_1$ satisfying $|x_0 - y_0| \geq (3\sqrt{n} + 1)r$ and*

$$\mu\big(\big\{(x, y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_{\varepsilon,q}^2 \lambda^2\big\}\big) \geq \varepsilon \mu(\mathcal{B}_{\frac{r}{2}}(x_0, y_0)), \tag{6.14}$$

we have

$$\begin{aligned}
\mathcal{B}_{\frac{r}{2}}(x_0, y_0) \subset &\big\{(x, y) \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0) \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > \lambda^2\big\} \\
&\cup \big\{(x, y) \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x, x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2\big\} \\
&\cup \big\{(x, y) \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(y, y) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2\big\} \\
&\cup \big\{(x, y) \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(G^{q_0})(x, x) > 3^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}\big\} \\
&\cup \big\{(x, y) \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(G^{q_0})(y, y) > 3^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}\big\}.
\end{aligned}$$

Proof. Assume that (6.14) holds, but that the conclusion is false, so that there exists a point $(x', y') \in \mathcal{B}_{\frac{r}{2}}(x_0, y_0)$ such that

$$\begin{aligned}
\mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x', y') &\leq \lambda^2, \\
\mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x', x') &\leq 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2, \\
\mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(y', y') &\leq 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2 \\
\mathcal{M}_{\geq r, \mathcal{B}_{4n}}(G^{q_0})(x', x') &\leq 3^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}, \\
\mathcal{M}_{\geq r, \mathcal{B}_{4n}}(G^{q_0})(y', y') &\leq 3^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}.
\end{aligned}$$

Therefore, for any $\rho \geq r$ we have

$$\fint_{\mathcal{B}_\rho(x', y')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \leq \lambda^2, \tag{6.15}$$

$$\fint_{\mathcal{B}_\rho(x')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \leq 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2, \tag{6.16}$$

$$\fint_{\mathcal{B}_\rho(x')} \chi_{\mathcal{B}_{4n}} G^{q_0} \, d\mu \leq 3^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}. \tag{6.17}$$

Of course, (6.16) and (6.17) hold also with x' replaced by y'. Since for any $\rho \geq r$ we have $\mathcal{B}_\rho(x_0, y_0) \subset \mathcal{B}_{2\rho}(x', y') \subset \mathcal{B}_{3\rho}(x_0, y_0)$, from (6.15) we deduce

$$\begin{aligned}
\fint_{\mathcal{B}_\rho(x_0, y_0)} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu &\leq \frac{\mu(\mathcal{B}_{2\rho}(x', y'))}{\mu(\mathcal{B}_\rho(x_0, y_0))} \fint_{\mathcal{B}_{2\rho}(x')} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \\
&\leq \frac{\mu(\mathcal{B}_{3\rho}(x_0, y_0))}{\mu(\mathcal{B}_\rho(x_0, y_0))} \leq 3^{n+2\theta} \lambda^2.
\end{aligned} \tag{6.18}$$

Note also that for any $\rho \geq r$ we have $\mathcal{B}_\rho(x_0) \subset \mathcal{B}_{2\rho}(x')$. Together with (6.16) this observation yields

$$\begin{aligned} \fint_{\mathcal{B}_\rho(x_0)} \chi_{\mathcal{B}_{4n}} U^2\, d\mu &\leq \frac{\mu(\mathcal{B}_{2\rho}(x'))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x')} \chi_{\mathcal{B}_{4n}} U^2\, d\mu \\ &\leq 6^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)} \lambda^2 \end{aligned} \tag{6.19}$$

and similarly

$$\begin{aligned} \fint_{\mathcal{B}_\rho(x_0)} \chi_{\mathcal{B}_{4n}} G^{q_0}\, d\mu &\leq \frac{\mu(\mathcal{B}_{2\rho}(x'))}{\mu(\mathcal{B}_\rho(x_0))} \fint_{\mathcal{B}_{2\rho}(x')} \chi_{\mathcal{B}_{4n}} G^{q_0}\, d\mu \\ &\leq 6^{n+2\theta} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)} \lambda^{q_0}. \end{aligned} \tag{6.20}$$

By the same reasoning, (6.19) and (6.20) clearly also hold with x_0 replaced by y_0. Next we claim that

$$\begin{aligned} &\left\{(x, y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_{\varepsilon,q}^2 \lambda^2\right\} \\ &\qquad \subset \left\{(x, y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)}(U^2)(x, y) > N_{\varepsilon,q}^2 \lambda^2\right\}. \end{aligned} \tag{6.21}$$

To see this, assume that

$$(x_1, y_1) \in \left\{x \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)}(U^2)(x, y) \leq N_{\varepsilon,q}^2 \lambda^2\right\}. \tag{6.22}$$

For $\rho < \sqrt{n}r$, we have $\mathcal{B}_\rho(x_1, y_1) \subset \mathcal{B}_{\sqrt{n}r}(x_1, y_1) \subset \mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0, y_0)$, so that together with (6.22) we deduce

$$\fint_{\mathcal{B}_\rho(x_1,y_1)} U^2\, d\mu \leq \mathcal{M}_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)}(U^2)(x_1, y_1) \leq N_{\varepsilon,q}^2 \lambda^2.$$

On the other hand, for $\rho \geq \sqrt{n}r$ we have $\mathcal{B}_\rho(x_1, y_1) \subset \mathcal{B}_{3\rho}(x', y') \subset \mathcal{B}_{5\rho}(x_1, y_1)$, so that (6.15) implies

$$\fint_{\mathcal{B}_\rho(x_1,y_1)} \chi_{\mathcal{B}_{4n}} U^2\, d\mu \leq \frac{\mu(\mathcal{B}_{5\rho}(x_1, y_1))}{\mu(\mathcal{B}_\rho(x_1, y_1))} \fint_{\mathcal{B}_{3\rho}(x',y')} \chi_{\mathcal{B}_{4n}} U^2\, d\mu \leq 5^{n+2\theta} \lambda^2 \leq N_{\varepsilon,q}^2 \lambda^2.$$

Thus, we have

$$(x_1, y_1) \in \left\{(x, y) \in \mathcal{B}_r(x_0, y_0) \;\middle|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) \leq N_{\varepsilon,q}^2 \lambda^2\right\},$$

which implies (6.21). As in the proof of Lemma 5.3, let $m \in \mathbb{N}$ be determined by $2^{m-1}r < \sqrt{n} \leq 2^m r$; note that $m \geq 2$. Then for any $k < m$, by (6.19) and (6.20) we have

$$\begin{aligned} \fint_{\mathcal{B}_{2^k \frac{3\sqrt{n}}{2}r}(x_0)} U^2\, d\mu &\leq 6^{n+2\theta} N_d^2 \lambda^2 \tilde{\phi}(r, x_0, y_0)^{-2(s+\theta)}, \\ \fint_{\mathcal{B}_{2^k \frac{3\sqrt{n}}{2}r}(x_0)} G^{q_0}\, d\mu &\leq 6^{n+2\theta} \lambda^{q_0} \tilde{\phi}(r, x_0, y_0)^{-q_0(s+\theta)}. \end{aligned} \tag{6.23}$$

Moreover, in view of (5.10), the inclusions $\mathcal{B}_{2^k\frac{n}{2}}(x_0) \subset \mathcal{B}_{2^{k+m-1}\frac{3\sqrt{n}}{2}r}(x_0) \subset \mathcal{B}_{2^k\frac{3n}{2}}(x_0) \subset \mathcal{B}_{2^k 4n}$ and the fact that $\tilde{\phi}(r, x_0, y_0) \leq 1$, we have

$$\begin{aligned}
&\sum_{k=m}^{\infty} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} U^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \bigg(\frac{\mu(\mathcal{B}_{2^k 4n})}{\mu(\mathcal{B}_{2^k\frac{n}{2}})} \fint_{\mathcal{B}_{2^k 4n}} U^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq 8^{\frac{n}{2}+\theta} \lambda_0 \leq 8^{\frac{n}{2}+\theta} \lambda_0 \tilde{\phi}(r, x_0, y_0)^{-(s+\theta)}. \qquad (6.24)
\end{aligned}$$

Together with (6.23) and the assumption that $\lambda \geq \lambda_0$, we arrive at

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} U^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq \sum_{k=1}^{m-1} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} U^2 \, d\mu \bigg)^{\frac{1}{2}} + \sum_{k=m}^{\infty} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} U^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq 8^{\frac{n}{2}+\theta} C_{s,\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-(s+\theta)} \lambda. \qquad (6.25)
\end{aligned}$$

By similar reasoning to above, we note that estimate (6.24) also holds with U replaced by G. Therefore, along with Hölder's inequality we obtain

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} G^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq \sum_{k=1}^{m-1} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} G^{q_0} \, d\mu \bigg)^{\frac{1}{q_0}} + \sum_{k=m}^{\infty} 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)} G^2 \, d\mu \bigg)^{\frac{1}{2}} \\
&\quad \leq 8^{\frac{n}{2}+\theta} C_{s,\theta} \tilde{\phi}(r, x_0, y_0)^{-(s+\theta)} \lambda. \qquad (6.26)
\end{aligned}$$

Again, by the same arguments as above, (6.25) and (6.26) clearly also hold for x_0 replaced by y_0. Therefore, together with the weak $\frac{q^\star}{2}$–$\frac{q^\star}{2}$ estimate for the Hardy–Littlewood maximal function, Proposition 6.2 with $m = \frac{1}{3\sqrt{n}}$, (6.18), (6.25), (6.20), (6.26) and (6.11), we arrive at

$$\begin{aligned}
&\mu\big(\big\{(x, y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x, y) > N_{\varepsilon,q}^2 \lambda^2\big\}\big) \\
&\quad \leq \mu\big(\big\{(x, y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0, y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)}(U^2)(x, y) > N_{\varepsilon,q}^2 \lambda^2\big\}\big) \\
&\quad \leq N_{\varepsilon,q}^{-q^\star} \lambda^{-q^\star} \int_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)} U^{q^\star} \, d\mu
\end{aligned}$$

$$
\begin{aligned}
&\leq N_{\varepsilon,q}^{-q^\star}\lambda^{-q^\star}3^{q^\star}C_{nd}^{q^\star}\mu(\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0))\\
&\quad\times\Bigg[\bigg(\fint_{\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)}U^2\,d\mu\bigg)^{\frac{q^\star}{2}}\\
&\qquad+\Big(\frac{3\sqrt{n}r/2}{\operatorname{dist}(B_{\frac{3\sqrt{n}}{2}r}(x_0),B_{\frac{3\sqrt{n}}{2}r}(y_0))}\Big)^{q^\star(s+\theta)}\\
&\qquad\quad\times\bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)}U^2\,d\mu\bigg)^{\frac{1}{2}}+\bigg(\fint_{\mathcal{B}_{3\sqrt{n}r}(x_0)}G^{q_0}\,d\mu\bigg)^{\frac{1}{q_0}}\\
&\qquad\qquad+\sum_{k=1}^{\infty}2^{-k(s-\theta)}\bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(x_0)}G^2\,d\mu\bigg)^{\frac{1}{2}}\bigg)^{q^\star}\\
&\qquad+\Big(\frac{3\sqrt{n}r/2}{\operatorname{dist}(B_{\frac{3\sqrt{n}}{2}r}(x_0),B_{\frac{3\sqrt{n}}{2}r}(y_0))}\Big)^{q^\star(s+\theta)}\\
&\qquad\quad\times\bigg(\sum_{k=1}^{\infty}2^{-k(s-\theta)}\bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(y_0)}U^2\,d\mu\bigg)^{\frac{1}{2}}+\bigg(\fint_{\mathcal{B}_{3\sqrt{n}r}(y_0)}G^{q_0}\,d\mu\bigg)^{\frac{1}{q_0}}\\
&\qquad\qquad+\sum_{k=1}^{\infty}2^{-k(s-\theta)}\bigg(\fint_{\mathcal{B}_{2^k\frac{3\sqrt{n}}{2}r}(y_0)}G^2\,d\mu\bigg)^{\frac{1}{2}}\bigg)^{q^\star}\Bigg]\\
&\leq N_{\varepsilon,q}^{-q^\star}\lambda^{-q^\star}3^{q^\star}C_{nd}^{q^\star}(3\sqrt{n})^{n+2\theta}\mu(\mathcal{B}_{\frac{r}{2}}(x_0,y_0))\\
&\quad\times\big(3^{(\frac{n}{2}+\theta)q^\star}\lambda^{q^\star}\\
&\qquad+6^{q^\star}(9n)^{q^\star(s+\theta)}\tilde{\phi}(r,x_0,y_0)^{q^\star(s+\theta)}8^{nq^\star}C_{s,\theta}^{q^\star}N_d^{q^\star}\tilde{\phi}(r,x_0,y_0)^{-q^\star(s+\theta)}\lambda^{q^\star}\big)\\
&<\varepsilon\mu(\mathcal{B}_{\frac{r}{2}}(x_0,y_0)),
\end{aligned}
$$

which contradicts (6.14) and thus finishes the proof. ■

Since we are going to use a Calderón–Zygmund cube decomposition, we next prove a version of the previous Lemma with balls replaced by cubes. For notational convenience, in analogy to the quantity $\tilde{\phi}(r,x_0,y_0)$ defined in (6.13), for any $r\in(0,\frac{\sqrt{n}}{2})$ and all $x_0,y_0\in\mathbb{R}^n$ with $|x_0-y_0|>\sqrt{n}r$, we define the quantity

$$
\phi(r,x_0,y_0):=\frac{r}{\operatorname{dist}(Q_r(x_0),Q_r(y_0))}. \tag{6.27}
$$

Note that since $B_{r/2}(x_0)\subset Q_r(x_0)$ and $B_{r/2}(y_0)\subset Q_r(y_0)$, the two quantities are related by

$$
\tilde{\phi}(r,x_0,y_0)\leq\phi(r,x_0,y_0). \tag{6.28}
$$

Corollary 6.4. *For any $\lambda \geq \lambda_0$, $r \in (0, \frac{\sqrt{n}}{2})$ and any point $(x_0, y_0) \in \mathcal{Q}_1$ satisfying $|x_0 - y_0| \geq (3\sqrt{n}+1)r$ and*

$$\mu\Big(\big\{(x,y) \in \mathcal{Q}_r(x_0,y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2 \lambda^2\big\}\Big) > \varepsilon \mu(\mathcal{Q}_r(x_0,y_0)), \tag{6.29}$$

we have

$$\begin{aligned}
&\mu(\mathcal{Q}_r(x_0,y_0)) \\
&\leq (\sqrt{n})^{n+2\theta} \Big(\mu\big(\big\{(x,y) \in \mathcal{Q}_r(x_0,y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) \\
&\qquad + \phi(r,x_0,y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\big\{(x,y) \in \mathcal{Q}_r(x_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(r,x_0,y_0)^{-2(\theta+s)} \lambda^2\big\}\big) \\
&\qquad + \phi(r,x_0,y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\big\{(x,y) \in \mathcal{Q}_r(y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(r,x_0,y_0)^{-2(\theta+s)} \lambda^2\big\}\big) \\
&\qquad + \phi(r,x_0,y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\big\{(x,y) \in \mathcal{Q}_r(x_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(r,x_0,y_0)^{-q_0(\theta+s)} \lambda^{q_0}\big\}\big) \\
&\qquad + \phi(r,x_0,y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\big\{(x,y) \in \mathcal{Q}_r(y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(r,x_0,y_0)^{-q_0(\theta+s)} \lambda^{q_0}\big\}\big)\Big).
\end{aligned}$$

Proof. First of all, note that $(x_0,y_0) \in \mathcal{Q}_1 \subset \mathcal{B}_{\frac{\sqrt{n}}{2}}$. By assumption (6.29) and the fact that $\mathcal{Q}_r(x_0,y_0) \subset \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0,y_0)$, we have

$$\begin{aligned}
&\mu\Big(\big\{(x,y) \in \mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0,y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2\big\}\Big) \\
&\quad \geq \mu\Big(\big\{(x,y) \in \mathcal{Q}_r(x_0,y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2\big\}\Big) \\
&\quad > \varepsilon\mu(\mathcal{Q}_r(x_0,y_0)) \geq \varepsilon\mu(\mathcal{B}_{r/2}(x_0,y_0)).
\end{aligned}$$

Therefore, assumption (6.14) from Lemma 6.3 is satisfied, so that by the volume doubling property of μ, Lemma 6.3 and the inclusion $\mathcal{B}_{r/2}(x_0,y_0) \subset \mathcal{Q}_r(x_0,y_0)$, we obtain

$$\begin{aligned}
&\mu(\mathcal{Q}_r(x_0,y_0)) \leq \mu(\mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0,y_0)) \\
&\leq (\sqrt{n})^{n+2\theta} \mu(\mathcal{B}_{r/2}(x_0,y_0)) \\
&\leq (\sqrt{n})^{n+2\theta} \Big(\mu\big(\big\{(x,y) \in \mathcal{Q}_r(x_0,y_0) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) \\
&\quad + \mu\big(\big\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \;\big|\; \mathcal{M}_{\geq r,\mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r,x_0,y_0)^{-2(\theta+s)} \lambda^2\big\}\big) \\
&\quad + \mu\big(\big\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \;\big|\; \mathcal{M}_{\geq r,\mathcal{B}_{4n}}(U^2)(y,y) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r,x_0,y_0)^{-2(\theta+s)} \lambda^2\big\}\big) \\
&\quad + \mu\big(\big\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \;\big|\; \mathcal{M}_{\geq r,\mathcal{B}_{4n}}(G^{q_0})(x,x) > 3^{n+2\theta} \tilde{\phi}(r,x_0,y_0)^{-q_0(\theta+s)} \lambda^{q_0}\big\}\big) \\
&\quad + \mu\big(\big\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \;\big|\; \mathcal{M}_{\geq r,\mathcal{B}_{4n}}(G^{q_0})(y,y) > 3^{n+2\theta} \tilde{\phi}(r,x_0,y_0)^{-q_0(\theta+s)} \lambda^{q_0}\big\}\big)\Big).
\end{aligned}$$

We proceed by further estimating the second term on the right-hand side of the last display and note that the last three terms can be estimated similarly. Set

$$F_1 := \{x \in B_{r/2}(x_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)} \lambda^2\}.$$

We have

$$\begin{aligned}
&\mu\big(\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)} \lambda^2\}\big) \\
&\quad = \mu(F_1 \times B_{r/2}(y_0)) \\
&\quad = \int_{F_1} \int_{B_{r/2}(y_0)} \frac{dy\,dx}{|x-y|^{n-2\theta}} \\
&\quad \leq \operatorname{dist}(B_{r/2}(x_0), B_{r/2}(y_0))^{-(n-2\theta)} |F_1|\,|B_{r/2}(y_0)| \\
&\quad = \operatorname{dist}(B_{r/2}(x_0), B_{r/2}(y_0))^{-(n-2\theta)} |F_1|\,|B_{r/2}(x_0)| \\
&\quad \leq \operatorname{dist}(B_{r/2}(x_0), B_{r/2}(y_0))^{-(n-2\theta)} r^{n-2\theta} \int_{F_1} \int_{B_{r/2}(x_0)} \frac{dy\,dx}{|x-y|^{n-2\theta}} \\
&\quad = \tilde{\phi}(r, x_0, y_0)^{n-2\theta} \mu(F_1 \times B_{r/2}(x_0)) \\
&\quad = \tilde{\phi}(r, x_0, y_0)^{n-2\theta} \\
&\qquad \times \mu\big(\{(x,y) \in \mathcal{B}_{r/2}(x_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^q \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)} \lambda^2\}\big).
\end{aligned}$$

In order to further estimate the right-hand side, we claim that for all $x, y \in B_{r/2}(x_0)$ we have

$$\mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) \leq 3^{n+2\theta} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y). \tag{6.30}$$

Indeed, for any $\rho \geq r$ we have $\mathcal{B}_\rho(x) \subset \mathcal{B}_{2\rho}(x,y) \subset \mathcal{B}_{3\rho}(x)$ and therefore

$$\fint_{\mathcal{B}_{3\rho}(x)} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \leq \frac{\mu(\mathcal{B}_{3\rho}(x))}{\mu(B_\rho(x))} \fint_{\mathcal{B}_{2\rho}(x,y)} \chi_{\mathcal{B}_{4n}} U^2 \, d\mu \leq 3^{n+2\theta} \widetilde{\mathcal{M}}_{\mathcal{B}_{4n}}(U^2)(x,y),$$

which proves (6.30). By (6.30) and the display before that, we arrive at

$$\begin{aligned}
&\mu\big(\{(x,y) \in \mathcal{B}_{r/2}(x_0,y_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)}\}\big) \\
&\qquad \leq \tilde{\phi}(r, x_0, y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\{(x,y) \in \mathcal{B}_{r/2}(x_0) \mid \mathcal{M}_{\geq r, \mathcal{B}_{4n}}(U^2)(x,x) > 3^{n+2\theta} N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)}\}\big) \\
&\qquad \leq \tilde{\phi}(r, x_0, y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\{(x,y) \in \mathcal{B}_{r/2}(x_0) \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \tilde{\phi}(r, x_0, y_0)^{-2(\theta+s)}\}\big) \\
&\qquad \leq \phi(r, x_0, y_0)^{n-2\theta} \\
&\qquad\quad \times \mu\big(\{(x,y) \in \mathcal{Q}_r(x_0) \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(r, x_0, y_0)^{-2(\theta+s)}\}\big),
\end{aligned}$$

where we used the inequality (6.28) and the inclusion $\mathcal{B}_{r/2}(x_0) \subset \mathcal{Q}_r(x_0)$ in order to obtain the last inequality. As mentioned, by observing that the last three terms on the right-hand side in the second display of the proof can be estimated similarly, the proof is finished. ∎

7. A covering argument

In order to proceed, we split the unit cube Q_1 in $\mathbb{R}^n$ into dyadic cubes as follows. First, we split Q_1 into 2^n cubes of sidelength $\frac{1}{2}$. Next we split each of the resulting cubes into 2^n cubes of sidelength $\frac{1}{2^2} = \frac{1}{4}$ and iterate this process. The resulting family of cubes are called dyadic cubes. By using the same procedure with n replaced by $2n$, we can also split the unit cube $\mathcal{Q}_1 = Q_1 \times Q_1$ in $\mathbb{R}^{2n}$ into a family of dyadic cubes. We observe that any dyadic cube $\mathcal{K} \subset \mathcal{Q}_1$ of the resulting dyadic cubes can be written as $\mathcal{K} = K_1 \times K_2$, where K_1 and K_2 are n-dimensional dyadic cubes contained in Q_1. By construction, any such dyadic cube $\mathcal{K}$ has sidelength $2^{-k(\mathcal{K})}$ for some nonnegative integer $k(\mathcal{K}) \geq 0$. Moreover, for $k \geq 1$ any such dyadic cube $\mathcal{K}$ with sidelength $2^{-k(\mathcal{K})}$ is contained in exactly one dyadic cube $\widetilde{\mathcal{K}}$ with sidelength $2^{-k(\mathcal{K})+1}$; we call $\widetilde{\mathcal{K}}$ the predecessor of $\mathcal{K}$. We need the following version of the Calderón–Zygmund decomposition, which roughly speaking shows that a subset of $\mathcal{Q}_1$ with small enough density can be covered by a sequence of dyadic cubes with density properties that are desirable for our purposes. For a proof we refer to [8, Lemma 1.1], where the result is proved with respect to the Lebesgue measure instead of μ. However, the proof also works for the doubling measure μ by taking into account that the Lebesgue differentiation theorem (see Proposition 3.4) also holds with respect to μ.

Lemma 7.1. *Let $E \subset \mathcal{Q}_1$ be a measurable set satisfying*

$$0 < \mu(E) < \varepsilon\mu(\mathcal{Q}_1) \quad \text{for some } \varepsilon \in (0,1). \tag{7.1}$$

Then there exists a countable family $\mathcal{U}_\lambda$ of dyadic cubes obtained from $\mathcal{Q}_1$ such that

$$\mu\Big(E \setminus \bigcup_{\mathcal{K} \in \mathcal{U}_\lambda} \mathcal{K}\Big) = 0 \tag{7.2}$$

and such that for any $\mathcal{K} \in \mathcal{U}_\lambda$ we have

$$\mu(E \cap \mathcal{K}) \geq \varepsilon\mu(\mathcal{K}) \tag{7.3}$$

and

$$\mu(E \cap \widetilde{\mathcal{K}}) < \varepsilon\mu(\widetilde{\mathcal{K}}), \tag{7.4}$$

where $\widetilde{\mathcal{K}}$ denotes the predecessor of $\mathcal{K}$.

Next, fix some $\lambda \geq \lambda_0$ and consider the level set

$$E := \big\{(x,y) \in \mathcal{Q}_1 \,\big|\, \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2\lambda^2\big\}. \tag{7.5}$$

Provided that (7.1) is satisfied with respect to this E and some $\varepsilon \in (0,1)$ which we will choose later, by Lemma 7.1 there exists a countable family $\mathcal{U}_\lambda$ of dyadic cubes obtained from $\mathcal{Q}_1$, such that (7.2), (7.3) and (7.4) are satisfied with respect to E.

In order to treat the cubes of the family $\mathcal{U}_\lambda$ which are in some sense close enough to the diagonal, we also construct an auxiliary diagonal cover consisting of balls as follows. For $x_0 \in Q_1$ and $r > 0$, consider the quantity

$$\begin{aligned}\Psi_\lambda(x_0, r) := \mu\big(\big\{(x,y) \in \mathcal{B}_r(x_0) \cap \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big) \\ + \mu\big(\big\{(x,y) \in \mathcal{B}_r(x_0) \cap \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big),\end{aligned}$$

where q_0 is given by (5.8). Observe that since $N_d \leq N_{\varepsilon,q}$, for any $x_0 \in Q_1$ and any $r > 0$ we have

$$\mu(E \cap B_r(x_0)) \leq \Psi_\lambda(x_0, r). \tag{7.6}$$

Now fix $\kappa \in (0,1)$ to be chosen later and consider the following subset of the diagonal in $\mathcal{Q}_1$:

$$D_{\kappa\varepsilon} := \Big\{(x,x) \in \mathcal{Q}_1 \Bigm| \sup_{0<r<\sqrt{n}/2} \Psi_\lambda(x,r) \geq \kappa\varepsilon\mu(\mathcal{B}_r(x))\Big\}.$$

By the weak 1–1 estimate for the Hardy–Littlewood maximal function (see Proposition 3.3) as well as using that $\lambda \geq \lambda_0$, we obtain

$$\begin{aligned}&\mu\big(\big\{(x,y) \in \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big) \\ &\qquad + \mu\big(\big\{(x,y) \in \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big) \\ &\quad \leq \frac{C_1}{N_d^2\lambda_0^2}\int_{\mathcal{B}_{4n}} U^2\,d\mu + \frac{C_1}{\lambda_0^{q_0}}\int_{\mathcal{B}_{4n}} G^{q_0}\,d\mu < \kappa\varepsilon\mu(\mathcal{Q}_1),\end{aligned} \tag{7.7}$$

where $C_1 = C_1(n,s,\theta) > 0$ and the last inequality is obtained by choosing M_0 large enough in (5.10). Since $N_d \leq N_{\varepsilon,q}$, (7.7) in particular implies condition (7.1) with respect to the set E defined in (7.5), so that the family $\mathcal{U}_\lambda$ of dyadic cubes as stated above indeed exists.

Since by (7.7) for any $x \in Q_1$ and any $r \geq \sqrt{n}/2$ we have

$$\begin{aligned}\Psi_\lambda(x,r) &\leq \mu\big(\big\{(x,y) \in \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big) \\ &\qquad + \mu\big(\big\{(x,y) \in \mathcal{Q}_1 \bigm| \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big) \\ &< \kappa\varepsilon\mu(\mathcal{Q}_1) \leq \kappa\varepsilon\mu(\mathcal{B}_r) = \kappa\varepsilon\mu(\mathcal{B}_r(x)),\end{aligned} \tag{7.8}$$

and by definition of $D_{\kappa\varepsilon}$ for any $(x,x) \in D_{\kappa\varepsilon}$ there exists some $0 < r < \sqrt{n}/2$ such that

$$\Psi_\lambda(x,r) \geq \kappa\varepsilon\mu(\mathcal{B}_r(x)),$$

we see that for any $(x,x) \in D_{\kappa\varepsilon}$ there exists some exit radius $r_x \in (0, \sqrt{n}/2)$ such that

$$\Psi_\lambda(x, r_x) \geq \kappa\varepsilon\mu(\mathcal{B}_{r_x}(x)) \tag{7.9}$$

and

$$\Psi_\lambda(x,r) \leq \kappa\varepsilon\mu(\mathcal{B}_r(x)) \quad \text{for all } r > r_x. \tag{7.10}$$

Now we consider the diagonal covering $\{\mathcal{B}_{r_x}(x) \mid (x,x) \in D_{\kappa\varepsilon}\}$. Since $(\mathbb{R}^{2n}, \|\cdot\|)$ is a separable metric space, by the Vitali covering lemma there exists a countable subset J_D of $D_{\kappa\varepsilon}$, such that the family of balls $\{B_{r_x}(x)\}_{(x,x)\in J_D}$ is disjoint and we have

$$\bigcup_{(x,x)\in D_{\kappa\varepsilon}} \mathcal{B}_{r_x}(x) \subset \bigcup_{(x,x)\in J_D} \mathcal{B}_{5r_x}(x). \tag{7.11}$$

Next we classify the cubes from the family $\mathcal{U}_\lambda$ as follows. Let $\mathcal{U}_\lambda^d$ be the collection of all cubes from $\mathcal{U}_\lambda$ that can be sucked up by the diagonal cover, that is,

$$\mathcal{U}_\lambda^d := \left\{\mathcal{K} \in \mathcal{U}_\lambda \;\middle|\; \mathcal{K} \subset \textstyle\bigcup_{(x,x)\in D_{\kappa\varepsilon}} \mathcal{B}_{r_x}(x)\right\}. \tag{7.12}$$

Moreover, we define the family $\mathcal{U}_\lambda^{nd} := \mathcal{U}_\lambda \setminus \mathcal{U}_\lambda^d$, so that $\mathcal{U}_\lambda$ is the disjoint union of $\mathcal{U}_\lambda^d$ and $\mathcal{U}_\lambda^{nd}$.

The following lemma reduces the problem of estimating the measure of E with respect to μ to estimating the measures of the diagonal balls in the family J_D and the measures of the off-diagonal cubes in the family $\mathcal{U}_\lambda^{nd}$.

Lemma 7.2. *Let E be given by (7.5) and let $\varepsilon \in (0,1)$. Then we have*

$$\mu(E) \le C\varepsilon\bigg(\kappa \sum_{(x,x)\in J_D} \mu(\mathcal{B}_{r_x}(x) \cap \mathcal{Q}_1) + \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \mu(\mathcal{K})\bigg), \tag{7.13}$$

where $C = C(n,\theta) > 0$.

Proof. Any $\mathcal{K} \in \mathcal{U}_\lambda^{nd}$ can be written as $\mathcal{K} = \mathcal{Q}_r(x_0,y_0)$ for some $r > 0$ and some $x_0, y_0 \in \mathbb{R}^n$. Since $\widetilde{\mathcal{K}} \subset \mathcal{Q}_{3r}(x_0,y_0)$, we have

$$\begin{aligned} \mu(\widetilde{\mathcal{K}}) \le \mu(\mathcal{Q}_{3r}(x_0,y_0)) \le \mu(\mathcal{B}_{\frac{3\sqrt{n}}{2}r}(x_0,y_0)) \le (3\sqrt{n})^{n+2\theta}\mu(\mathcal{B}_{\frac{r}{2}}(x_0,y_0)) \\ \le (3\sqrt{n})^{n+2\theta}\mu(\mathcal{K}). \end{aligned} \tag{7.14}$$

By (7.2), (7.11), (7.6), (7.10), (7.4), (7.14) and Lemma 3.2, we have

$$\begin{aligned} \mu(E) &\le \mu\bigg(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda} (\mathcal{K}\cap E)\bigg) \\ &= \mu\bigg(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^d} (\mathcal{K}\cap E)\bigg) + \mu\bigg(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} (\mathcal{K}\cap E)\bigg) \\ &\le \mu\bigg(\bigcup_{(x,x)\in D_{\kappa\varepsilon}} (\mathcal{B}_{r_x}(x)\cap E)\bigg) + \mu\bigg(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} (\mathcal{K}\cap E)\bigg) \\ &\le \mu\bigg(\bigcup_{(x,x)\in J_D} (\mathcal{B}_{5r_x}(x)\cap E)\bigg) + \mu\bigg(\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} (\widetilde{\mathcal{K}}\cap E)\bigg) \end{aligned}$$

$$\leq \sum_{(x,x)\in J_D} \mu(\mathcal{B}_{5r_x}(x)\cap E) + \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \mu(\widetilde{\mathcal{K}}\cap E)$$
$$\leq \sum_{(x,x)\in J_D} \Psi_\lambda(x,5r_x) + \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \mu(\widetilde{\mathcal{K}}\cap E)$$
$$\leq \sum_{(x,x)\in J_D} \kappa\varepsilon\mu(\mathcal{B}_{5r_x}(x)) + \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \varepsilon\mu(\widetilde{\mathcal{K}})$$
$$= \varepsilon\Big(\kappa 5^{n+2\theta} \sum_{(x,x)\in J_D} \mu(\mathcal{B}_{r_x}(x)) + \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \mu(\widetilde{\mathcal{K}})\Big)$$
$$\leq \varepsilon\Big(\kappa(5\sqrt{n})^{n+2\theta}C_1 \sum_{(x,x)\in J_D} \mu(\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1) + (3\sqrt{n})^{n+2\theta} \sum_{\mathcal{K}\in\mathcal{U}_\lambda^{nd}} \mu(\mathcal{K})\Big),$$

where $C_1 = C_1(n,\theta) \geq 1$. Thus, (7.13) holds with $C = (5\sqrt{n})^{n+2\theta}C_1$. ∎

7.1. Diagonal estimates

Our next goal is to use the results from Section 6 in order to further estimate the right-hand side of (7.13). We start by estimating the first sum on the right-hand side of (7.13), i.e. we are first dealing with the diagonal case, so that Lemma 6.1 is the crucial tool.

Lemma 7.3. *Let $\delta = \delta(\varepsilon,\kappa,n,s,\theta,\Lambda) \in (0,1)$ be given by Lemma 6.1. Then we have*

$$\sum_{(x,x)\in J_D} \mu(\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1) \leq \mu\big(\{(x,y)\in\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big) + 2\kappa^{-1}\varepsilon^{-1}\mu\big(\{(x,y)\in\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \delta^{q_0}\lambda^{q_0}\}\big). \tag{7.15}$$

Proof. Fix some $(x,x)\in J_D$, so that by (7.9) for the corresponding exit radius r_x at least one of the following two inequalities must hold:

$$\mu\big(\{(x,y)\in\mathcal{B}_r(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\}\big) \geq \frac{\kappa\varepsilon}{2}\mu(\mathcal{B}_r(x)), \tag{7.16}$$
$$\mu\big(\{(x,y)\in\mathcal{B}_r(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\}\big) \geq \frac{\kappa\varepsilon}{2}\mu(\mathcal{B}_r(x)). \tag{7.17}$$

If (7.16) is satisfied, then Lemma 6.1 with κ replaced by $\kappa/2$ implies

$$B_{r_x}(x)\cap\mathcal{Q}_1 \subset \{(x,y)\in B_{r_x}(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\} \cup \{(x,y)\in B_{r_x}(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \delta^{q_0}\lambda^{q_0}\}. \tag{7.18}$$

If on the other hand (7.17) is satisfied, then we directly obtain that

$$\begin{aligned}\mu(\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1) &\leq \mu(\mathcal{B}_{r_x}(x))\\ &\leq 2\kappa^{-1}\varepsilon^{-1}\mu\big(\{(x,y)\in\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\}\big)\\ &\leq 2\kappa^{-1}\varepsilon^{-1}\mu\big(\{(x,y)\in\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \delta^{q_0}\lambda^{q_0}\}\big).\end{aligned} \tag{7.19}$$

Therefore, in view of (7.18) and (7.19), for any $(x,x) \in J_D$ we have

$$\begin{aligned}
&\mu(\mathcal{B}_{r_x}(x) \cap \mathcal{Q}_1) \\
&\quad \le \mu\big(\{(x,y) \in \mathcal{B}_{r_x}(x) \cap \mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big) \\
&\qquad + 2\kappa^{-1}\varepsilon^{-1}\mu\big(\{(x,y) \in \mathcal{B}_{r_x}(x) \cap \mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \delta^{q_0}\lambda^{q_0}\}\big).
\end{aligned}$$

Since the family of balls $\{B_{r_x}(x)\}_{(x,x)\in J_D}$ is disjoint, assertion (7.15) immediately follows from the last display. ■

7.2. Off-diagonal estimates

In order to estimate the measures of the off-diagonal cubes of class $\mathcal{U}_\lambda^{nd}$, we have ensure that as our terminology suggests, such cubes are indeed sufficiently far away from the diagonal in terms of their own sidelength.

Lemma 7.4. *There exists $\kappa = \kappa(n,\theta) > 0$ small enough, such that for any cube $\mathcal{K} = K_1 \times K_2 \in \mathcal{U}_\lambda^{nd}$ of sidelength $2^{-k(\mathcal{K})}$, we have*

$$\operatorname{dist}(K_1, K_2) \ge (3\sqrt{n}+1)2^{-k(\mathcal{K})}.$$

Proof. Let $\mathcal{K} = K_1 \times K_2 \in \mathcal{U}_\lambda^{nd}$ and assume that $\operatorname{dist}(K_1,K_2) < (3\sqrt{n}+1)2^{-k(\mathcal{K})}$. Let us show that in this case for κ small enough we have $\mathcal{K} \in \mathcal{U}_\lambda^{d}$, leading to a contradiction. Let x be the center of K_1, y be the center of K_2 and set $z := (x+y)/2$. Then for $r := (5\sqrt{n}+1)2^{-k(\mathcal{K})}$, we have $\mathcal{K} \subset \mathcal{B}_r(z)$. Moreover, we have

$$\begin{aligned}
\mu(\mathcal{B}_r(z)) &= c r^{n+2\theta} \\
&= c(5\sqrt{n}+1)^{n+2\theta}(2^{-k(\mathcal{K})})^{n+2\theta} \\
&= c(5\sqrt{n}+1)^{n+2\theta}(2^{-k(\mathcal{K})})^{-(n-2\theta)} \int_{K_1}\int_{K_2} dx\,dy \\
&\le c(5\sqrt{n}+1)^{n+2\theta}(5\sqrt{n}+1)^{n-2\theta} \int_{K_1}\int_{K_2} \frac{dx\,dy}{|x-y|^{n-2\theta}} \\
&= c(5\sqrt{n}+1)^{2n}\mu(\mathcal{K}),
\end{aligned}$$

where $c = c(n,\theta) > 0$. Now we assume that

$$\kappa \le c^{-1}(5\sqrt{n}+1)^{-2n}. \tag{7.20}$$

Together with (7.3) applied to the set E defined in (7.5), we obtain

$$\Psi_\lambda(z,r) \ge \mu(E \cap \mathcal{K}) \ge \varepsilon\mu(\mathcal{K}) \ge \kappa\varepsilon\mu(\mathcal{B}_r(z)). \tag{7.21}$$

In particular, (7.21) implies that $r < \sqrt{n}/2$, since otherwise we get a contradiction to (7.8). Therefore, we have

$$\sup_{0<r<\sqrt{n}/2} \Psi_\lambda(z,r) \ge \kappa\varepsilon\mu(\mathcal{B}_r(z)),$$

so that by definition of $\mathcal{D}_{\kappa\varepsilon}$ we obtain that $(z,z) \in \mathcal{D}_{\kappa\varepsilon}$. Moreover, in view of (7.10) we deduce that $r \leq r_z$, where r_z is the exit radius at the point z determined in (7.9) and (7.10). We therefore have

$$\mathcal{K} \subset \mathcal{B}_r(z) \subset \mathcal{B}_{r_z}(z) \subset \bigcup_{(x,x)\in D_{\kappa\varepsilon}} \mathcal{B}_{r_x}(x),$$

so that $\mathcal{K} \in \mathcal{U}^d_\lambda$, which contradicts the assumption that $\mathcal{K} \in \mathcal{U}^{nd}_\lambda$. ∎

In what follows, we set

$$\kappa := \min\{c^{-1}(5\sqrt{n}+1)^{-2n}, (6\sqrt{n})^{-(n+2\theta)}\}, \tag{7.22}$$

where c is given as in (7.20), so that in particular, Lemma 7.4 is at our disposal.

For any cube $\mathcal{K} = K_1 \times K_2 \in \mathcal{U}^{nd}_\lambda$, we write $P_1(\mathcal{K}) := K_1 \times K_1$ and $P_2(\mathcal{K}) := K_2 \times K_2$. Furthermore, we write

$$\phi(\mathcal{K}) := \frac{2^{-k(\mathcal{K})}}{\operatorname{dist}(K_1, K_2)},$$

which matches the function $\phi(r, x_0, y_0)$ introduced in (6.27).

In view of Corollary 6.4 and Lemma 7.4, for any cube $\mathcal{K} \in \mathcal{U}^{nd}_\lambda$ we have

$$\begin{aligned}
\mu(\mathcal{K}) &= \mu(\mathcal{Q}_r(x_0,y_0)) \\
&\leq (\sqrt{n})^{n+2\theta}\Big(\mu\big(\{(x,y)\in\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big) \\
&\quad + \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y)\in P_1\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big) \\
&\quad + \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y)\in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big) \\
&\quad + \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y)\in P_1\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\}\big) \\
&\quad + \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y)\in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\}\big)\Big).
\end{aligned} \tag{7.23}$$

For $h = 1, 2$, for simplicity of notation we define

$$Z^{U,h}_\lambda(\mathcal{K}) := \{(x,y)\in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\}$$

and

$$Z^{G,h}_\lambda(\mathcal{K}) := \{(x,y)\in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\}.$$

In order to handle the diagonal level sets on the right-hand side of (7.23), for $h = 1, 2$ we also define the subfamilies

$$\mathcal{G}^h_\lambda := \{\mathcal{K} \in \mathcal{U}^{nd}_\lambda \mid \mu(Z^{U,h}_\lambda(\mathcal{K})) + \mu(Z^{G,h}_\lambda(\mathcal{K})) \leq \varepsilon\mu(P_h\mathcal{K})\}$$

and

$$\mathcal{N}^h_\lambda := \{\mathcal{K} \in \mathcal{U}^{nd}_\lambda \mid \mu(Z^{U,h}_\lambda(\mathcal{K})) + \mu(Z^{G,h}_\lambda(\mathcal{K})) > \varepsilon\mu(P_h\mathcal{K})\}.$$

Moreover, set

$$\mathcal{G}_\lambda := \mathcal{G}_\lambda^1 \cap \mathcal{G}_\lambda^2 \quad \text{and} \quad \mathcal{N}_\lambda := \mathcal{N}_\lambda^1 \cup \mathcal{N}_\lambda^2,$$

so that we clearly have $\mathcal{U}_\lambda^{nd} = \mathcal{G}_\lambda \cup \mathcal{N}_\lambda$, where the union is disjoint.

The following lemma shows that if $\mathcal{K} \in \mathcal{G}_\lambda$, then the diagonal terms on the right-hand side of (7.23) can be treated in a fairly straightforward manner.

Lemma 7.5. *For ε small enough, we have*

$$\sum_{\mathcal{K} \in \mathcal{G}_\lambda} \mu(\mathcal{K}) \leq C\mu\big(\big\{(x,y) \in \mathcal{Q}_1 \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big), \tag{7.24}$$

where $C = C(n,\theta) > 0$.

Proof. Since in view of Lemma 7.4 we have $\phi(\mathcal{K}) \leq 1$, together with (7.23), using that $\mu(P_1\mathcal{K}) = \mu(P_2\mathcal{K})$, for any $\mathcal{K} \in \mathcal{G}_\lambda$ we obtain

$$\begin{aligned}
\mu(\mathcal{K}) \leq (\sqrt{n})^{n+2\theta} \Big(& \mu\big(\big\{(x,y) \in \mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) \\
& + \phi(\mathcal{K})^{n-2\theta}\big(\mu(Z_\lambda^{U,1}(\mathcal{K})) + \mu(Z_\lambda^{U,2}(\mathcal{K})) \\
& \qquad + \mu(Z_\lambda^{G,1}(\mathcal{K})) + \mu(Z_\lambda^{G,2}(\mathcal{K}))\big)\Big) \\
\leq (\sqrt{n})^{n+2\theta} \big(& \mu\big(\big\{(x,y) \in \mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) + 2\phi(\mathcal{K})^{n-2\theta}\varepsilon\mu(P_1\mathcal{K})\big).
\end{aligned}$$

Note that we have $\mathcal{K} = \mathcal{Q}_r(x_0,y_0)$ for $r = 2^{-k(\mathcal{K})}$ and $x_0, y_0 \in Q_1$. In view of Lemma 7.4 we can argue in the same way as in (6.9) and (6.10) to obtain

$$\begin{aligned}
\frac{2^{-2nk(\mathcal{K})}}{\operatorname{dist}(K_1,K_2)^{n-2\theta}} &\leq 2^{2n+1} \frac{(r/2)^{2n}}{\operatorname{dist}(B_{\frac{r}{2}}(x_0), B_{\frac{r}{2}}(y_0))^{n-2\theta}} \\
&\leq 2^{2n+1}\mu(\mathcal{B}_{\frac{r}{2}}(x_0,y_0)) \leq 2^{2n+1}\mu(\mathcal{K}).
\end{aligned}$$

Since also

$$\mu(P_1\mathcal{K}) \leq \mu(\mathcal{B}_{\frac{\sqrt{n}}{2}r}(x_0)) = C_1 r^{n+2\theta} = C_1 2^{-(n+2\theta)k(\mathcal{K})}$$

for some $C_1 = C_1(n,\theta) \geq 1$, we deduce that

$$\phi(\mathcal{K})^{n-2\theta} \leq C_2 \frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}, \tag{7.25}$$

where $C_2 = C_2(n,\theta) \geq 1$. By connecting the previous display with the first one in the proof and assuming that

$$\varepsilon \leq \frac{1}{4(\sqrt{n})^{n+2\theta}C_2}, \tag{7.26}$$

we arrive at

$$\begin{aligned}
\mu(\mathcal{K}) &\leq (\sqrt{n})^{n+2\theta}\Big(\mu\big(\big\{(x,y) \in \mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) + 2C_2\varepsilon\mu(\mathcal{K})\Big) \\
&\leq (\sqrt{n})^{n+2\theta}\mu\big(\big\{(x,y) \in \mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big) + \frac{\mu(\mathcal{K})}{2},
\end{aligned}$$

which by reabsorbing the last term on the right-hand side into the left-hand side implies

$$\mu(\mathcal{K}) \leq 2(\sqrt{n})^{n+2\theta} \mu\big(\big\{(x,y) \in \mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\big\}\big).$$

Summing over $\mathcal{K} \in \mathcal{G}_\lambda$ and using that all cubes in the family $\mathcal{U}_\lambda$ and therefore also all cubes in $\mathcal{G}_\lambda$ are disjoint and contained in $\mathcal{Q}_1$, we see that (7.24) holds with $C = 2(\sqrt{n})^{n+2\theta}$. This finishes the proof. ∎

From now on, we will always assume that the restriction (7.26) on ε holds, so that estimate (7.24) from Lemma 7.5 holds.

It remains to also control the last four terms on the right-hand side of (7.23) in the more involved case when $\mathcal{K} \in \mathcal{N}_\lambda$, which requires a delicate combinatorial argument inspired by [28]. In order to accomplish this, for $h = 1, 2$ we define the families

$$P_h \mathcal{N}_\lambda := \big\{P_h \mathcal{K} \;\big|\; \mathcal{K} \in \mathcal{N}_\lambda^h\big\}.$$

Since $P_1 \mathcal{N}_\lambda \cup P_2 \mathcal{N}_\lambda$ is a family of dyadic cubes, clearly there is a disjoint subfamily $P\mathcal{N}_\lambda$ of $P_1 \mathcal{N}_\lambda \cup P_2 \mathcal{N}_\lambda$ such that

$$\bigcup_{\mathcal{H} \in P\mathcal{N}_\lambda} \mathcal{H} = \bigcup_{\mathcal{K} \in P_1 \mathcal{N}_\lambda \cup P_2 \mathcal{N}_\lambda} \mathcal{K}. \tag{7.27}$$

In other words, any cube in $P_1 \mathcal{N}_\lambda \cup P_2 \mathcal{N}_\lambda$ is contained in exactly one cube $\mathcal{H} \in P\mathcal{N}_\lambda$. The following lemma plays a crucial role in the mentioned combinatorial argument. It shows that a cube of class $\mathcal{N}_\lambda$ is not only far away from the diagonal in terms of its own sidelength as shown in Lemma 7.4, but also in terms of the sidelength of the larger cube $\mathcal{H} \in P\mathcal{N}_\lambda$ in which its projection onto the diagonal is contained.

Lemma 7.6. *Let $\mathcal{K} = K_1 \times K_2 \in \mathcal{N}_\lambda$, so that for some $h \in \{1, 2\}$, $P_h \mathcal{K}$ belongs to $P_h \mathcal{N}_\lambda$. Moreover, let $\mathcal{H} = H \times H$ be the unique cube that belongs to $P\mathcal{N}_\lambda$ and contains $P_h \mathcal{K}$. Then we have* $\operatorname{dist}(K_1, K_2) \geq 2^{-k(\mathcal{H})}$.

Proof. Without loss of generality we can assume that $h = 1$, since the case when $h = 2$ can be treated in the same way. We prove by contradiction. Assume that

$$\operatorname{dist}(K_1, K_2) < 2^{-k(\mathcal{H})} \tag{7.28}$$

and denote by $x_{\mathcal{H}}$ the center of the cube H. Choose points $x_1 \in \overline{H}$ and $y_1 \in \overline{K}_2$ such that $\operatorname{dist}(\overline{H}, \overline{K}_2) = |x_1 - y_1|$ and denote by y_0 the center of K_2. Then for any $y \in K_2$, we have

$$\begin{aligned} |x_{\mathcal{H}} - y| &\leq |x_1 - y_1| + |x_{\mathcal{H}} - x| + |y_1 - y_0| + |y_0 - y| \\ &< 2^{-k(\mathcal{H})} + \frac{\sqrt{n}}{2} 2^{-k(\mathcal{H})} + \sqrt{n} 2^{-k(\mathcal{K})} \leq 3\sqrt{n} 2^{-k(\mathcal{H})}, \end{aligned}$$

so that $K_2 \subset B_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}})$. Since by assumption $K_1 \subset H$, we arrive at

$$\mathcal{K} \subset \mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}}). \tag{7.29}$$

Since $\mathcal{H}$ belongs to $P\mathcal{N}_\lambda$ and thus to $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$, we have

$$\mu\big(\big\{(x,y)\in\mathcal{H} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big) \\ + \mu\big(\big\{(x,y)\in\mathcal{H} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big) > \varepsilon\mu(\mathcal{H}). \tag{7.30}$$

Inequality (7.30) implies

$$\begin{aligned}\Psi_\lambda(x_{\mathcal{H}}, 3\sqrt{n}2^{-k(\mathcal{H})}) &= \mu\big(\big\{(x,y)\in\mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}}) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big)\\ &\quad + \mu\big(\big\{(x,y)\in\mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}}) \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big)\\ &\geq \mu\big(\big\{(x,y)\in\mathcal{H} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\lambda^2\big\}\big)\\ &\quad + \mu\big(\big\{(x,y)\in\mathcal{H} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \lambda^{q_0}\big\}\big)\\ &> \varepsilon\mu(\mathcal{H}) \geq \varepsilon\mu(\mathcal{B}_{2^{-(k(\mathcal{H})+1)}}(x_{\mathcal{H}}))\\ &\geq \frac{1}{(6\sqrt{n})^{n+2\theta}}\varepsilon\mu(\mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}})) \geq \kappa\varepsilon\mu(\mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}})),\end{aligned}$$

where we also used (7.22) in order to obtain the last inequality. Therefore, in view of (7.9) and (7.10) we have $3\sqrt{n}2^{-k(\mathcal{H})} \leq r_{x_{\mathcal{H}}}$, where $r_{x_{\mathcal{H}}}$ is the exit radius at the point $x_{\mathcal{H}}$. In particular, we have

$$\mathcal{B}_{3\sqrt{n}2^{-k(\mathcal{H})}}(x_{\mathcal{H}}) \subset \mathcal{B}_{r_{x_{\mathcal{H}}}}(x_{\mathcal{H}}),$$

which together with (7.29) implies

$$\mathcal{K} \subset \mathcal{B}_{r_{x_{\mathcal{H}}}}(x_{\mathcal{H}}).$$

But then by definition of $\mathcal{U}_\lambda^d$ (see (7.12)), we obtain that $\mathcal{K}\in\mathcal{U}_\lambda^d$, which is a contradiction since $\mathcal{K}\in\mathcal{N}_\lambda\subset\mathcal{U}_\lambda^{nd}$ and $\mathcal{U}_\lambda^d\cap\mathcal{U}_\lambda^{nd}=\emptyset$. Thus, the proof is finished. ■

We now estimate the measures of the cubes of class $\mathcal{N}_\lambda$. The main tools are (7.23) and the above Lemma 7.6, which allow the projected diagonal cubes to be classified in a way that enables us to control the diagonal terms in (7.23).

Lemma 7.7. *We have*

$$\sum_{\mathcal{K}\in\mathcal{N}_\lambda}\mu(\mathcal{K}) \leq \frac{C}{\lambda^2}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\ + \frac{C}{\lambda^{q_0}}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu, \tag{7.31}$$

where $C=C(n,s,\theta)>0$.

Proof. Step 1: A combinatorial argument. For any $\mathcal{H}\in P\mathcal{N}_\lambda$ and $h\in\{1,2\}$, we define the family of cubes

$$\mathcal{N}_\lambda^h(\mathcal{H}) := \big\{\mathcal{K}\in\mathcal{N}_\lambda \;\big|\; P_h\mathcal{K}\subset\mathcal{H}\big\}.$$

Since the family $P\mathcal{N}_\lambda$ is a disjoint covering of the family $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$, we can decompose $\mathcal{N}_\lambda^h$ into mutually disjoint subfamilies as

$$\mathcal{N}_\lambda^h = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \mathcal{N}_\lambda^h(\mathcal{H}), \tag{7.32}$$

in the sense that we have $\mathcal{N}_\lambda^h(\mathcal{H}_1) \cap \mathcal{N}_\lambda^h(\mathcal{H}_2) = \emptyset$ whenever $\mathcal{H}_1 \neq \mathcal{H}_2$. Since for any $\mathcal{H} \in P\mathcal{N}_\lambda$ and any $\mathcal{K} \in \mathcal{N}_\lambda^h(\mathcal{H})$ we have $k(\mathcal{K}) = k(P_1\mathcal{K}) \geq k(\mathcal{H})$, we can decompose $\mathcal{N}_\lambda^h(\mathcal{H})$ into the classes

$$[\mathcal{N}_\lambda^h(\mathcal{H})]_i := \big\{\mathcal{K} \in \mathcal{N}_\lambda^h(\mathcal{H}) \;\big|\; k(\mathcal{K}) = i + k(\mathcal{H})\big\}, \quad i \in \mathbb{N}_0,\ h \in \{1,2\}.$$

More precisely, we have the decomposition into mutually disjoint subfamilies

$$\mathcal{N}_\lambda^h(\mathcal{H}) = \bigcup_{i\geq 0} [\mathcal{N}_\lambda^h(\mathcal{H})]_i, \tag{7.33}$$

in the sense that $[\mathcal{N}_\lambda^h(\mathcal{H})]_i \cap [\mathcal{N}_\lambda^h(\mathcal{H})]_j = \emptyset$ whenever $i \neq j$.

Next, for $h \in \{1,2\}$ and $i, j \in \mathbb{N}_0$ we define further subfamilies by

$$[\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j} := \big\{\mathcal{K} = K_1 \times K_2 \in [\mathcal{N}_\lambda^h(\mathcal{H})]_i \;\big|\; 2^{j-k(\mathcal{H})} \leq \operatorname{dist}(K_1, K_2) < 2^{j+1-k(\mathcal{H})}\big\}.$$

Since by Lemma 7.6 for any $\mathcal{K} = K_1 \times K_2 \in \mathcal{N}_\lambda^h(\mathcal{H})$ we have $2^{-k(\mathcal{H})} \leq \operatorname{dist}(K_1, K_2)$, we have the disjoint decomposition

$$\mathcal{N}_\lambda^h(\mathcal{H}) = \bigcup_{i,j\geq 0} [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j}, \tag{7.34}$$

in the sense that $[\mathcal{N}_\lambda^h(\mathcal{H})]_{i_1,j_1} \cap [\mathcal{N}_\lambda^h(\mathcal{H})]_{i_2,j_2} = \emptyset$ whenever $(i_1, j_1) \neq (i_2, j_2)$.

Therefore, by combining (7.32) and (7.34), we arrive at the following decomposition of mutually disjoint subfamilies:

$$\mathcal{N}_\lambda^h = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \bigcup_{i,j\geq 0} [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j}. \tag{7.35}$$

Now fix some $\mathcal{H} = H \times H \in P\mathcal{N}_\lambda$. Our next goal is to prove that for $h \in \{1,2\}$ the following inequality holds:

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{N}_\lambda^h(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta} \mu\Big(\big\{(x,y) \in P_h\mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\big\}\Big) \\
&\quad + \sum_{\mathcal{K}\in\mathcal{N}_\lambda^h(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta} \mu\Big(\big\{(x,y) \in P_h\mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\big\}\Big) \\
&\leq C_0\Bigg(\frac{1}{\lambda^2}\int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\qquad + \frac{1}{\lambda^{q_0}}\int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\Bigg),
\end{aligned} \tag{7.36}$$

where $C_0 = C_0(n,s) > 0$. By definition of the class $[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}$, for any $\mathcal{K} = K_1 \times K_2 \in [\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}$ we have

$$\phi(\mathcal{K}) = \frac{2^{-k(\mathcal{K})}}{\operatorname{dist}(K_1,K_2)} = \frac{1}{2^i}\frac{2^{-k(\mathcal{H})}}{\operatorname{dist}(K_1,K_2)} \leq \frac{1}{2^{i+j}}.$$

By using Chebychev's inequality, (7.34) and then the last display, we obtain

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{N}^h_\lambda(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta}\mu\big(\big\{(x,y)\in P_h\mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\big\}\big)\\
&\leq \frac{1}{N_d^2\lambda^2}\sum_{\mathcal{K}\in\mathcal{N}^h_\lambda(\mathcal{H})}\phi(\mathcal{K})^{n+2s}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>N_d^2\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu\\
&\leq \frac{1}{\lambda^2}\sum_{i,j=0}^{\infty}\sum_{\mathcal{K}\in[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}}\phi(\mathcal{K})^{n+2s}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu\\
&\leq \frac{1}{\lambda^2}\sum_{i,j=0}^{\infty}\Big(\frac{1}{2^{i+j}}\Big)^{n+2s}\sum_{\mathcal{K}\in[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \qquad (7.37)
\end{aligned}$$

and similarly by additionally using that $n + (q_0-2)\theta + q_0 s \geq n + 2s$,

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{N}^h_\lambda(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta}\mu\big(\big\{(x,y)\in P_h\mathcal{K} \;\big|\; \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\big\}\big)\\
&\leq \frac{1}{\lambda^{q_0}}\sum_{\mathcal{K}\in\mathcal{N}^h_\lambda(\mathcal{H})}\phi(\mathcal{K})^{n+(q_0-2)\theta+q_0 s}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\\
&\leq \frac{1}{\lambda^{q_0}}\sum_{i,j=0}^{\infty}\sum_{\mathcal{K}\in[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}}\phi(\mathcal{K})^{n+(q_0-2)\theta+q_0 s}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\\
&\leq \frac{1}{\lambda^{q_0}}\sum_{i,j=0}^{\infty}\Big(\frac{1}{2^{i+j}}\Big)^{n+2s}\sum_{\mathcal{K}\in[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}}\int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu. \qquad (7.38)
\end{aligned}$$

In order to proceed, we need to further decompose the families $[\mathcal{N}^h_\lambda(\mathcal{H})]_{i,j}$. For any $i \geq 0$, $\mathcal{H}$ contains exactly 2^{ni} disjoint diagonal cubes $\mathcal{H}_i^m = H_i^m \times H_i^m$, $m \in \{1,\dots,2^{ni}\}$, with sidelength $2^{-i-k(\mathcal{H})}$. In particular, the disjointness of these cubes implies

$$\sum_{m=1}^{2^{ni}}\int_{\mathcal{H}_i^m\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \leq \int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}}\mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \qquad (7.39)$$

and

$$\sum_{m=1}^{2^{ni}}\int_{\mathcal{H}_i^m\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu \leq \int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}}\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu. \qquad (7.40)$$

For $h \in \{1,2\}$, $i, j \geq 0$ and $m \in \{1,\dots,2^{ni}\}$, define the families

$$[\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j,m} := \left\{\mathcal{K} \in [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j} \;\middle|\; P_h\mathcal{K} = \mathcal{H}_i^m\right\}$$

and observe that we have the disjoint decomposition

$$[\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j} = \bigcup_{m=1}^{2^{ni}} [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j,m}. \tag{7.41}$$

For a moment, let us focus on the case when $h = 1$. Note that since $\mathcal{N}_\lambda^1$ is a family of dyadic cubes, we have $P_2\mathcal{K}_1 \cap P_2\mathcal{K}_2 = \emptyset$ for all cubes $\mathcal{K}_1, \mathcal{K}_2 \in [\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j,m}$ with $\mathcal{K}_1 \neq \mathcal{K}_2$, since otherwise $\mathcal{K}_1$ and $\mathcal{K}_2$ would coincide. Therefore, we observe that the cubes in $[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j,m}$ are all contained in the family $\mathcal{F}_{i,j,m}^1(\mathcal{H})$ consisting of all distinct dyadic cubes of the form $\mathcal{K} = H_i^m \times K$ with $K \subset \mathcal{Q}_1$ and sidelength $2^{-i-k(\mathcal{H})}$ that additionally satisfy

$$2^{j-k(\mathcal{H})} \leq \operatorname{dist}(H_i^m, K) < 2^{j+1-k(\mathcal{H})}. \tag{7.42}$$

Then in view of a combinatorial consideration, we have the estimate

$$\#[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j,m} \leq \#\mathcal{F}_{i,j,m}^1(\mathcal{H}) \leq C_1 2^{n(i+j)}, \tag{7.43}$$

where $C_1 = C_1(n) > 0$. Thus, in view of (7.41), (7.43), (7.39) and (7.40), we obtain

$$\begin{aligned}
&\sum_{\mathcal{K}\in[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j}} \int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\quad = \sum_{m=1}^{2^{ni}} \sum_{\mathcal{K}\in[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j}} \int_{\mathcal{H}_i^m\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\quad \leq C_1 2^{n(i+j)} \sum_{m=1}^{2^{ni}} \int_{\mathcal{H}_i^m\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\quad \leq C_1 2^{n(i+j)} \int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu
\end{aligned}$$

and by the same reasoning,

$$\begin{aligned}
&\sum_{\mathcal{K}\in[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j}} \int_{P_h\mathcal{K}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu \\
&\quad \leq C_1 2^{n(i+j)} \int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu.
\end{aligned}$$

In addition, by arguing similarly, the last two displays clearly also hold for $[\mathcal{N}_\lambda^1(\mathcal{H})]_{i,j}$ replaced by $[\mathcal{N}_\lambda^2(\mathcal{H})]_{i,j}$. Therefore, for $h \in \{1,2\}$ we deduce

$$\begin{aligned}
&\sum_{i,j=0}^{\infty} \Big(\frac{1}{2^{i+j}}\Big)^{n+2s} \sum_{\mathcal{K} \in [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j}} \int_{P_h \mathcal{K} \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > \lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\, d\mu \\
&\quad \leq C_1 \sum_{i,j=0}^{\infty} \Big(\frac{1}{2^{i+j}}\Big)^{2s} \int_{\mathcal{H} \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > \lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\, d\mu \\
&\quad \leq C_0 \int_{\mathcal{H} \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > \lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\, d\mu,
\end{aligned}$$

where $C_0 = C_1(\frac{1}{1-2^{-2s}})^2 < \infty$. Similarly, we also have

$$\begin{aligned}
&\sum_{i,j=0}^{\infty} \Big(\frac{1}{2^{i+j}}\Big)^{n+2s} \sum_{\mathcal{K} \in [\mathcal{N}_\lambda^h(\mathcal{H})]_{i,j}} \int_{P_h \mathcal{K} \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}) > \lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\, d\mu \\
&\quad \leq C_0 \int_{\mathcal{H} \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}) > \lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\, d\mu.
\end{aligned}$$

By combining the last two displays with (7.37) and (7.38), we finally arrive at estimate (7.36) with respect to C_0.

Step 2: Summation. For any $\mathcal{K} \in \mathcal{N}_\lambda$, we either have $\mathcal{K} \in \mathcal{M}_\lambda^1 \cap \mathcal{N}_\lambda^2$, $\mathcal{K} \in \mathcal{M}_\lambda^2 \cap \mathcal{N}_\lambda^1$ or $\mathcal{K} \in \mathcal{N}_\lambda^1 \cap \mathcal{N}_\lambda^2$. If $\mathcal{K} \in \mathcal{M}_\lambda^1 \cap \mathcal{N}_\lambda^2$, then in a similar way to the proof of Lemma 7.5, by using (7.23) and taking into account (7.25), we have

$$\begin{aligned}
\mu(\mathcal{K}) \leq{} & (\sqrt{n})^{n+2\theta} \\
& \times \Big(\mu\big(\{(x,y) \in \mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big) + C_2 \varepsilon \mu(\mathcal{K}) \\
& \quad + \phi(\mathcal{K})^{n-2\theta} \mu\big(\{(x,y) \in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(\mathcal{K})^{-2(\theta+s)} \lambda^2\}\big) \\
& \quad + \phi(\mathcal{K})^{n-2\theta} \mu\big(\{(x,y) \in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)} \lambda^{q_0}\}\big)\Big),
\end{aligned}$$

so that in view of the restriction (7.26) imposed on ε, reabsorbing the second term on the right-hand side into the left-hand side of the previous display yields

$$\begin{aligned}
\mu(\mathcal{K}) \leq C_3 \Big(& \mu\big(\{(x,y) \in \mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big) \\
& + \phi(\mathcal{K})^{n-2\theta} \mu\big(\{(x,y) \in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2 \phi(\mathcal{K})^{-2(\theta+s)} \lambda^2\}\big) \\
& + \phi(\mathcal{K})^{n-2\theta} \mu\big(\{(x,y) \in P_2\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)} \lambda^{q_0}\}\big)\Big),
\end{aligned}$$

where $C_3 = C_3(n,\theta) > 0$. By a similar argument, we also obtain that for any $\mathcal{K} \in \mathcal{M}_\lambda^2 \cap \mathcal{N}_\lambda^1$, we have

$$\begin{aligned}\mu(\mathcal{K}) \leq C_3\Big(&\mu\big(\{(x,y) \in \mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big)\\ &+ \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y) \in P_1\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big)\\ &+ \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y) \in P_1\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\}\big)\Big).\end{aligned}$$

By combining the last two displays with the fact that for any $\mathcal{K} \in \mathcal{N}_\lambda^1 \cap \mathcal{N}_\lambda^2$ we have estimate (7.23), we arrive at

$$\begin{aligned}&\sum_{\mathcal{K}\in\mathcal{N}_\lambda} \mu(\mathcal{K})\\ &\leq C_4\Bigg(\sum_{\mathcal{K}\in\mathcal{N}_\lambda} \mu\big(\{(x,y) \in \mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big)\\ &\qquad + \sum_{h=1}^{2}\Bigg(\sum_{\mathcal{K}\in\mathcal{N}_\lambda^h} \phi(\mathcal{K})^{n-2\theta}\\ &\qquad\qquad \times \mu\big(\{(x,y) \in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big)\Bigg)\\ &\qquad + \sum_{h=1}^{2}\Bigg(\sum_{\mathcal{K}\in\mathcal{N}_\lambda^h} \phi(\mathcal{K})^{n-2\theta}\\ &\qquad\qquad \times \mu\big(\{(x,y) \in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\}\big)\Bigg)\Bigg),\end{aligned}$$

where $C_4 = C_4(n,s,\theta) > 0$. Using the disjointness of the cubes $\mathcal{K} \in \mathcal{N}_\lambda$ and then Chebychev's inequality, for the first term on the right-hand side of the previous display, we deduce

$$\begin{aligned}\sum_{\mathcal{K}\in\mathcal{N}_\lambda} &\mu\big(\{(x,y) \in \mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big)\\ &\leq \mu\big(\{(x,y) \in \mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > \lambda^2\}\big)\\ &\leq \frac{1}{\lambda^2}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu.\end{aligned}$$

Moreover, in view of (7.32), (7.36) and the disjointness of the cubes $\mathcal{H} \in P\mathcal{N}_\lambda$ for $h \in \{1,2\}$, we obtain that

$$\begin{aligned}&\sum_{\mathcal{K}\in\mathcal{N}_\lambda^h} \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y) \in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big)\\ &\quad + \sum_{\mathcal{K}\in\mathcal{N}_\lambda^h} \phi(\mathcal{K})^{n-2\theta}\mu\big(\{(x,y) \in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^2\}\big)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\mathcal{H}\in P\mathcal{N}_\lambda} \sum_{\mathcal{K}\in\mathcal{N}_\lambda^h(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta} \\
&\qquad\qquad \times \mu\big(\{(x,y)\in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_d^2\phi(\mathcal{K})^{-2(\theta+s)}\lambda^2\}\big) \\
&\quad + \sum_{\mathcal{H}\in P\mathcal{N}_\lambda} \sum_{\mathcal{K}\in\mathcal{N}_\lambda^h(\mathcal{H})} \phi(\mathcal{K})^{n-2\theta} \\
&\qquad\qquad \times \mu\big(\{(x,y)\in P_h\mathcal{K} \mid \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})(x,y) > \phi(\mathcal{K})^{-q_0(\theta+s)}\lambda^{q_0}\}\big) \\
&\le C_0 \sum_{\mathcal{H}\in P\mathcal{N}_\lambda} \bigg(\frac{1}{\lambda^2}\int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\qquad\qquad + \frac{1}{\lambda^{q_0}}\int_{\mathcal{H}\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\bigg) \\
&= C_0\bigg(\frac{1}{\lambda^2}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\qquad + \frac{1}{\lambda^{q_0}}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\bigg).
\end{aligned}
$$

The estimate (7.31) now follows directly by combining the last three displays. ■

7.3. Level set estimate

By combining the above results, we are finally able to estimate the measure of the level set of $\mathcal{M}_{\mathcal{B}_{4n}}(U^2)$ in the whole cube $\mathcal{Q}_1$.

Corollary 7.8. *Under all the assumptions made above, for any $\lambda \ge \lambda_0$ we have*

$$
\begin{aligned}
&\mu\big(\{(x,y)\in\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2\lambda^2\}\big) \\
&\quad \le C\bigg(\frac{\varepsilon}{\lambda^2}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\qquad\qquad + \frac{1}{\delta^{q_0}\lambda^{q_0}}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\delta^{q_0}\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\bigg),
\end{aligned}
$$

where $C = C(n,s,\theta) > 0$ and $\delta = \delta(\varepsilon,n,s,\theta,\Lambda) \in (0,1)$ is given by Lemma 6.1.

Proof. In view of Lemmas 7.2, 7.3, 7.5, 7.7 and Chebychev's inequality, we obtain

$$
\begin{aligned}
&\mu\big(\{(x,y)\in\mathcal{Q}_1 \mid \mathcal{M}_{\mathcal{B}_{4n}}(U^2)(x,y) > N_{\varepsilon,q}^2\lambda^2\}\big) \\
&\quad \le C_1\varepsilon\bigg(\kappa\sum_{(x,x)\in J_D}\mu(\mathcal{B}_{r_x}(x)\cap\mathcal{Q}_1) + \sum_{\mathcal{K}\in\mathcal{G}_\lambda}\mu(\mathcal{K}) + \sum_{\mathcal{K}\in\mathcal{N}_\lambda}\mu(\mathcal{K})\bigg) \\
&\quad \le C\bigg(\frac{\varepsilon}{\lambda^2}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2)>\lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\,d\mu \\
&\qquad\qquad + \frac{1}{\delta^{q_0}\lambda^{q_0}}\int_{\mathcal{Q}_1\cap\{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})>\delta^{q_0}\lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\,d\mu\bigg),
\end{aligned}
$$

where all constants depend only on n, s and θ. This finishes the proof. ■

8. L^p estimates for U

We first prove the estimate we are interested in on a fixed scale in the form of an a priori estimate and under the additional assumption that U satisfies estimate (6.8). In order to do this, we use the following standard alternative characterization of the L^p norm which follows from Fubini's theorem in a straightforward way.

Lemma 8.1. *Let ν be a σ-finite measure on $\mathbb{R}^n$ and let $h\colon\Omega \to [0,+\infty]$ be a ν-measurable function in a domain $\Omega \subset \mathbb{R}^n$. Then for any $0 < \beta < \infty$ we have*

$$\int_\Omega h^\beta \, d\nu = \beta \int_0^\infty \lambda^{\beta-1} \nu\big(\{x \in \Omega \mid h(x) > \lambda\}\big)\, d\lambda.$$

Proposition 8.2. *Let $q \in [2,p)$ and $\tilde{q} \in (q_0, q^\star)$, where q_0 is given by (5.8). Then there exists some small enough $\delta = \delta(n,s,\theta,\Lambda,q,\tilde{q}) > 0$ such that if $A \in \mathcal{L}_0(\Lambda)$ is δ-vanishing in $\mathcal{B}_{4n}$ and $g \in W^{s,2}(\mathbb{R}^n)$ satisfies $G \in L^{\tilde{q}}(\mathcal{B}_{4n},\mu)$, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation $L_A^\Phi u = (-\Delta)^s g$ in B_{4n} that satisfies $U \in L^{\tilde{q}}(\mathcal{B}_{4n},\mu)$ and estimate (6.8) in any ball contained in B_{4n} with respect to q, we have*

$$\begin{aligned}\Big(\fint_{\mathcal{B}_{1/2}} U^{\tilde{q}}\, d\mu\Big)^{\frac{1}{\tilde{q}}} \leq C\bigg(&\sum_{k=1}^\infty 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k 4n}} U^2\, d\mu\Big)^{\frac{1}{2}}\\ &+\Big(\fint_{\mathcal{B}_{4n}} G^{\tilde{q}}\, d\mu\Big)^{\frac{1}{\tilde{q}}} + \sum_{k=1}^\infty 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k 4n}} G^2\, d\mu\Big)^{\frac{1}{2}}\bigg),\end{aligned}$$

where $C = C(n,s,\theta,\Lambda,q,\tilde{q},p) > 0$.

Proof. Let ε be chosen small enough and consider the corresponding $\delta = \delta(\varepsilon,n,s,\theta,\Lambda) > 0$ given by Lemma 6.1. Then by using Lemma 8.1 multiple times, first with

$$\beta = \tilde{q}, \quad h = \mathcal{M}_{\mathcal{B}_{4n}}(U^2)^{\frac{1}{2}}, \quad d\nu = d\mu,$$

then with

$$\beta = \tilde{q}-2, \quad h = \mathcal{M}_{\mathcal{B}_{4n}}(U^2)^{\frac{1}{2}}, \quad d\nu = \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\, d\mu,$$

and also with

$$\beta = \tilde{q}-q_0, \quad h = \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})^{\frac{1}{q_0}}, \quad d\nu = \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\, d\mu,$$

a change of variables, Corollary 7.8 and the definition of $N_{\varepsilon,q}$ from (6.11), we obtain

$$\begin{aligned}&\int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(U^2))^{\frac{\tilde{q}}{2}}\, d\mu\\ &\quad= \tilde{q}\int_0^\infty \lambda^{\tilde{q}-1}\mu\big(\mathcal{Q}_1 \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > \lambda^2\}\big)\, d\lambda\end{aligned}$$

$$
\begin{aligned}
&= \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \int_0^\infty \lambda^{\tilde{q}-1} \mu\big(\mathcal{Q}_1 \cap \big\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > N_{\varepsilon,q}^2 \lambda^2\big\}\big)\, d\lambda \\
&= \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \int_0^{\lambda_0} \lambda^{\tilde{q}-1} \mu\big(\mathcal{Q}_1 \cap \big\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > N_{\varepsilon,q}^2 \lambda^2\big\}\big)\, d\lambda \\
&\quad + \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \int_{\lambda_0}^\infty \lambda^{\tilde{q}-1} \mu\big(\mathcal{Q}_1 \cap \big\{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > N_{\varepsilon,q}^2 \lambda^2\big\}\big)\, d\lambda \\
&\leq \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \mu(\mathcal{Q}_1) \lambda_0^{\tilde{q}} + C_1 \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \varepsilon \int_0^\infty \lambda^{\tilde{q}-3} \int_{\mathcal{Q}_1 \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(U^2) > \lambda^2\}} \mathcal{M}_{\mathcal{B}_{4n}}(U^2)\, d\mu\, d\lambda \\
&\quad + C_1 \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \delta^{-q_0} \int_0^\infty \lambda^{\tilde{q}-q_0-1} \int_{\mathcal{Q}_1 \cap \{\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}) > \delta^{q_0} \lambda^{q_0}\}} \mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0})\, d\mu\, d\lambda \\
&= \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \mu(\mathcal{Q}_1) \lambda_0^{\tilde{q}} + C_1 \tilde{q} C_{nd} C_{s,\theta} N_d 10^{10n} \varepsilon^{1-\tilde{q}/q^\star} \int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(U^2))^{\frac{\tilde{q}}{2}}\, d\mu \\
&\quad + C_1 \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \delta^{-q_0} \int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}))^{\frac{\tilde{q}}{q_0}}\, d\mu,
\end{aligned}
$$

where $C_1 = C_1(n, s, \theta) \geq 1$. Now we set

$$
\varepsilon := \min\big\{(4(\sqrt{n})^{n+2\theta} C_2)^{-1}, (2C_1 \tilde{q} C_{nd} C_{s,\theta} N_d 10^{10n})^{-\frac{q^\star}{q^\star - \tilde{q}}}\big\},
$$

so that ε satisfies the restriction (7.26) and, moreover, we have

$$
C_1 \tilde{q} C_{nd} C_{s,\theta} N_d 10^{10n} \varepsilon^{1-\tilde{q}/q^\star} \leq \tfrac{1}{2}.
$$

Since, in addition, by assumption we have $U \in L^{\tilde{q}}(\mathcal{B}_{4n}, \mu)$, by Proposition 3.3 we have

$$
\int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(U^2))^{\frac{\tilde{q}}{2}}\, d\mu < \infty,
$$

so that we can reabsorb the second-to-last term on the right-hand side of the first display of the proof in the left-hand side, which yields

$$
\int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(U^2))^{\frac{\tilde{q}}{2}}\, d\mu \leq 2\tilde{q} N_{\varepsilon,q}^{\tilde{q}} \mu(\mathcal{Q}_1) \lambda_0^{\tilde{q}} + 2C_1 \tilde{q} N_{\varepsilon,q}^{\tilde{q}} \delta^{-q_0} \int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}))^{\frac{\tilde{q}}{q_0}}\, d\mu.
$$

Now, in view of Corollary 3.5 and Proposition 3.3, taking into account the definition of λ_0 from (5.10) and using Hölder's inequality, we obtain

$$
\begin{aligned}
\fint_{\mathcal{B}_{1/2}} U^{\tilde{q}}\, d\mu &\leq \frac{1}{\mu(\mathcal{B}_{1/2})} \int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(U^2))^{\frac{\tilde{q}}{2}}\, d\mu \\
&\leq C_2 \Big(\lambda_0^{\tilde{q}} + \int_{\mathcal{Q}_1} (\mathcal{M}_{\mathcal{B}_{4n}}(G^{q_0}))^{\frac{\tilde{q}}{q_0}}\, d\mu \Big) \\
&\leq C_3 \Big(\sum_{k=1}^\infty 2^{-k(s-\theta)} \Big(\fint_{\mathcal{B}_{2^k 4n}} U^2\, d\mu \Big)^{\frac{1}{2}} + \sum_{k=1}^\infty 2^{-k(s-\theta)} \Big(\fint_{\mathcal{B}_{2^k 4n}} G^2\, d\mu \Big)^{\frac{1}{2}} \\
&\qquad + \Big(\fint_{\mathcal{B}_{4n}} G^{q_0}\, d\mu \Big)^{\frac{1}{q_0}} \Big)^{\tilde{q}} + C_3 \int_{\mathcal{B}_{4n}} G^{\tilde{q}}\, d\mu
\end{aligned}
$$

$$\leq C_4\Big(\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k 4n}} U^2\,d\mu\Big)^{\frac{1}{2}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k 4n}} G^2\,d\mu\Big)^{\frac{1}{2}}\Big)^{\tilde{q}} + C_4 \fint_{\mathcal{B}_{4n}} G^{\tilde{q}}\,d\mu,$$

where all constants depend only on $n, s, \theta, \Lambda, q, \tilde{q}$ and p. This proves the desired estimate with $C = C_4^{1/\tilde{q}}$. ■

Corollary 8.3. *Consider some $q \in [2, p)$ and some $\tilde{q} \in (q_0, q^\star)$. Then there exists some small enough $\delta = \delta(n, s, \theta, \Lambda, q, \tilde{q}) > 0$ such that if $A \in \mathcal{L}_0(\Lambda)$ is δ-vanishing in B_1 and $g \in W^{s,2}(\mathbb{R}^n)$ satisfies $G \in L^{\tilde{q}}(\mathcal{B}_1, \mu)$, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation $L_A^\Phi u = (-\Delta)^s g$ in B_1 that satisfies $U \in L^{\tilde{q}}(\mathcal{B}_1, \mu)$ and estimate (6.8) in any ball contained in B_1 with respect to q, we have the estimate*

$$\Big(\fint_{\mathcal{B}_{1/2}} U^{\tilde{q}}\,d\mu\Big)^{\frac{1}{\tilde{q}}} \leq C\Big(\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k}} U^2\,d\mu\Big)^{\frac{1}{2}} + \Big(\fint_{\mathcal{B}_1} G^{\tilde{q}}\,d\mu\Big)^{\frac{1}{\tilde{q}}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)}\Big(\fint_{\mathcal{B}_{2^k}} G^2\,d\mu\Big)^{\frac{1}{2}}\Big),$$

where $C = C(n, s, \theta, \Lambda, q, \tilde{q}, p) > 0$.

Proof. There exists some small enough radius $r_1 \in (0, 1)$ such that

$$B_{4nr_1}(z) \Subset B_1 \tag{8.1}$$

for any $z \in B_{1/2}$. Now fix some $z \in B_{1/2}$ and consider the scaled functions $u_z, g_z \in W^{s,2}(\mathbb{R}^n)$ given by

$$u_z(x) := u(r_1 x + z), \quad g_z(x) := g(r_1 x + z), \quad A_z(x, y) := A(r_1 x + z, r_1 y + z).$$

Since A is δ-vanishing in $\mathcal{B}_1$, we see that A_z clearly is δ-vanishing in $B_{\frac{1}{4nr_1}}(-z) \supset B_{4n}$. Furthermore, in view of (8.1), u_z is a weak solution of $L_{A_z}^\Phi u_z = (-\Delta)^s g_z$ in $B_{\frac{1}{4nr_1}}(-z) \supset B_{4n}$. Now fix some $r > 0$ and some $x_0 \in \mathbb{R}^n$ such that $B_r(x_0) \subset B_{4n}$. Then again in view of (8.1), we clearly have

$$B_{r_1 r}(r_1 x_0 + z) \subset B_1,$$

so that by the assumption that estimate (6.8) holds for any ball contained in B_1, estimate (6.8) holds with respect to the ball $B_{r_1 r}(r_1 x_0 + z)$. Together with changes of variables, a simple computation now shows that the functions

$$U_z(x, y) := \frac{|u_z(x) - u_z(y)|}{|x - y|^{s+\theta}}, \quad G_z(x, y) := \frac{|g_z(x) - g_z(y)|}{|x - y|^{s+\theta}}$$

satisfy

$$\left(\fint_{\mathcal{B}_{r/2}(x_0)} U_z^q \, d\mu\right)^{\frac{1}{q}} \leq C_1 \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} U_z^2 \, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_r(x_0)} G_z^{q_0} \, d\mu\right)^{\frac{1}{q_0}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(x_0)} G_z^2 \, d\mu\right)^{\frac{1}{2}} \right), \tag{8.2}$$

where $C_1 = C_1(q, n, s, \theta, \Lambda) > 0$. Therefore, we see that U_z and G_z satisfy estimate (6.8) in any ball that is contained in B_{4n}. Since, in addition, the assumption that $U \in L^{\tilde{q}}(\mathcal{B}_1, \mu)$ clearly implies that $U_z \in L^{\tilde{q}}(\mathcal{B}_{\frac{1}{4nr_1}(-z)}, \mu) \subset L^{\tilde{q}}(\mathcal{B}_{4n}, \mu)$, by Proposition 8.2 we obtain

$$\left(\fint_{\mathcal{B}_{1/2}} U_z^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} \leq C_2 \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} U_z^2 \, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{4n}} G_z^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4n}} G_z^2 \, d\mu\right)^{\frac{1}{2}} \right),$$

where $C_2 = C_2(n, s, \theta, \Lambda, q, \tilde{q}, p) > 0$. By combining the last display with changes of variables, it is now straightforward to deduce that

$$\left(\fint_{\mathcal{B}_{r_1/2}(z)} U^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} \leq C_3 \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4nr_1}(z)} U^2 \, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{4nr_1}(z)} G^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4nr_1}(z)} G^2 \, d\mu\right)^{\frac{1}{2}} \right), \tag{8.3}$$

where $C_3 = C_3(q, \tilde{q}, p, n, s, \theta, \Lambda) > 0$. Since $\{B_{r_1/2}(z)\}_{z \in B_{1/2}}$ is an open covering of $\overline{B}_{1/2}$ and $\overline{B}_{1/2}$ is compact, there is a finite subcover $\{B_{r_1/2}(z_j)\}_{j=1}^m$ of $\overline{B}_{1/2}$ and hence also of $B_{1/2}$. In particular, $\{\mathcal{B}_{r/2}(z_j)\}_{j=1}^m$ is a finite subcover of $\mathcal{B}_{1/2}$. Therefore, by summing the above estimates, we obtain

$$\begin{aligned}\left(\fint_{\mathcal{B}_{1/2}} U^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} &\leq \sum_{j=1}^{m} \left(\fint_{\mathcal{B}_{r_1/2}(z_j)} U^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} \\ &\leq C_3 \sum_{j=1}^{m} \left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4nr_1}(z_j)} U^2 \, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{4nr_1}(z_j)} G^{\tilde{q}} \, d\mu\right)^{\frac{1}{\tilde{q}}} \right. \\ &\qquad \left. + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k 4nr_1}(z_j)} G^2 \, d\mu\right)^{\frac{1}{2}} \right)\end{aligned}$$

$$\begin{aligned}
&\leq C_4 \sum_{j=1}^{m} \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\frac{\mu(\mathcal{B}_{2^k 4n})}{\mu(\mathcal{B}_{2^k 4n r_1})} \fint_{\mathcal{B}_{2^k}} U^2 \, d\mu \Bigg)^{\frac{1}{2}} \\
&\qquad + \Bigg(\frac{\mu(\mathcal{B}_{4n})}{\mu(\mathcal{B}_{4n r_1})} \fint_{\mathcal{B}_1} G^{\tilde{q}} \, d\mu \Bigg)^{\frac{1}{\tilde{q}}} \\
&\qquad + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\frac{\mu(\mathcal{B}_{2^k})}{\mu(\mathcal{B}_{2^k 4n r_1})} \fint_{\mathcal{B}_{2^k 4n}} G^2 \, d\mu \Bigg)^{\frac{1}{2}} \Bigg) \\
&\leq C_4 m r_1^{-n/2-\theta} \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\fint_{\mathcal{B}_{2^k}} U^2 \, d\mu \Bigg)^{\frac{1}{2}} \\
&\qquad + \Bigg(\fint_{\mathcal{B}_1} G^{\tilde{q}} \, d\mu \Bigg)^{\frac{1}{\tilde{q}}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\fint_{\mathcal{B}_{2^k}} G^2 \, d\mu \Bigg)^{\frac{1}{2}} \Bigg),
\end{aligned}$$

where $C_4 = C_4(n, s, \theta, \Lambda, q, \tilde{q}, p) > 0$. Since in addition m and r_1 depend only on $B_{1/2}$ and thus only on n, the proof is finished. ∎

Corollary 8.4. *Let $r > 0$ and $z \in \mathbb{R}^n$ and consider some $q \in [2, p)$ and some $\tilde{q} \in (q_0, q^\star)$. Then there exists some small enough $\delta = \delta(n, s, \theta, \Lambda, q, \tilde{q}) > 0$ such that if $A \in \mathcal{L}_0(\Lambda)$ is δ-vanishing in $\mathcal{B}_r(z)$ and $g \in W^{s,2}(\mathbb{R}^n)$ satisfies $G \in L^{\tilde{q}}(\mathcal{B}_r(z), \mu)$, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation $L_A^\Phi u = (-\Delta)^s g$ in $B_r(z)$ that satisfies $U \in L^{\tilde{q}}(\mathcal{B}_r(z), \mu)$ and estimate (6.8) in any ball contained in $B_r(z)$ with respect to q, we have the estimate*

$$\begin{aligned}
\Bigg(\fint_{\mathcal{B}_{r/2}(z)} U^{\tilde{q}} \, d\mu \Bigg)^{\frac{1}{\tilde{q}}} &\leq C \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\fint_{\mathcal{B}_{2^k r}(z)} U^2 \, d\mu \Bigg)^{\frac{1}{2}} + \Bigg(\fint_{\mathcal{B}_r(z)} G^{\tilde{q}} \, d\mu \Bigg)^{\frac{1}{\tilde{q}}} \\
&\qquad + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \Bigg(\fint_{\mathcal{B}_{2^k r}(z)} G^2 \, d\mu \Bigg)^{\frac{1}{2}} \Bigg),
\end{aligned} \tag{8.4}$$

where $C = C(n, s, \theta, \Lambda, q, \tilde{q}, p) > 0$.

Remark 8.5. It is essential that the constant C in (8.4) does not depend on r and z.

Proof of Corollary 8.4. Consider the scaled functions $u_1, g_1 \in W^{s,2}(\mathbb{R}^n)$ given by

$$u_1(x) := u(rx + z), \quad g_1(x) := g(rx + z)$$

and also

$$A_1(x, y) := A(rx + z, ry + z).$$

Since A is δ-vanishing in $\mathcal{B}_r(z)$, A_1 clearly is δ-vanishing in B_1. Also, in view of a change of variables, for $U_1(x, y) := \frac{u_1(x) - u_1(y)}{|x-y|^{s+\theta}}$ we clearly have $U_1 \in L^{\tilde{q}}(\mathcal{B}_1, \mu)$. Moreover, since u and g satisfy estimate (6.8) in any ball contained in $B_r(z)$, by a straightforward scaling argument we deduce that u_1 and g_1 satisfy estimate (6.8) in any ball contained in B_1. In addition, u_1 is a weak solution of $L^{\Phi}_{A_1} u_1 = (-\Delta)^s g_1$ in B_1. Therefore, u_1 and A_1 satisfy all assumptions from Corollary 8.3, so that the estimate from Corollary 8.3 is satisfied by u_1 and g_1. The desired estimate (8.4) now follows by rescaling. ∎

Proposition 8.6. *Let $r > 0$, $z \in \mathbb{R}^n$, $s \in (0,1)$ and $p \in (2, \infty)$. Then there exists some small enough $\delta = \delta(n,s,\theta,\Lambda,p) > 0$ such that if $A \in \mathcal{L}_0(\Lambda)$ is δ-vanishing in $\mathcal{B}_{r_1}(z)$ and $g \in W^{s,2}(\mathbb{R}^n)$ satisfies $G \in L^p(\mathcal{B}_r(z), \mu)$, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation $L^{\Phi}_A u = (-\Delta)^s g$ in $B_r(z)$ that satisfies $U \in L^p(\mathcal{B}_r(z), \mu)$, we have the estimate*

$$\left(\fint_{\mathcal{B}_{r/2}(z)} U^p \, d\mu \right)^{\frac{1}{p}} \leq C \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(z)} U^2 \, d\mu \right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_r(z)} G^p \, d\mu \right)^{\frac{1}{p}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(z)} G^2 \, d\mu \right)^{\frac{1}{2}} \Bigg), \tag{8.5}$$

where $C = C(n,s,\theta,\Lambda,p) > 0$.

Proof. Define iteratively a sequence $\{q_i\}_{i=1}^{\infty}$ of real numbers by

$$q_1 := 2, \quad q_{i+1} := \min\big\{(q_i + (q_i)^\star)/2, p\big\},$$

where, as in (5.7), we let

$$(q_i)^\star = \begin{cases} \dfrac{nq_i}{n - sq_i} & \text{if } n > sq_i, \\ 2p & \text{if } n \leq sq_i. \end{cases}$$

Since for any i with $n > sq_{i+1}$ we have

$$\left(q_i + \frac{nq_i}{n - sq_i} \right)/2 - q_i = \frac{nq_i}{2(n - sq_i)} - \frac{q_i}{2} \geq \frac{4s}{2(n-s)} > 0,$$

there clearly exists some $i_p \in \mathbb{N}$ such that $q_{i_p} = p$. Since estimate (6.8) is trivially satisfied for $q = q_1 = 2$, and in view of the additional assumption that $U \in L^p(\mathcal{B}_r(z), \mu)$ we in particular have $U \in L^{q_2}(\mathcal{B}_r(z), \mu)$, if we choose δ small enough such that Corollary 8.4 is applicable with $q = 2$ and $\tilde{q} = q_2$, then all assumptions of Corollary 8.4 are satisfied with respect to $q = q_1 = 2$ and $\tilde{q} = q_2 \in (q_1, (q_1)^\star)$, so that we obtain

$$\left(\fint_{\mathcal{B}_{r/2}(z)} U^{q_2} \, d\mu \right)^{\frac{1}{q_2}} \leq C_1 \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(z)} U^2 \, d\mu \right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_r(z)} G^{q_2} \, d\mu \right)^{\frac{1}{q_2}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r}(z)} G^2 \, d\mu \right)^{\frac{1}{2}} \Bigg), \tag{8.6}$$

where $C_1 = C_1(n,s,\theta,\Lambda,p) > 0$. If $i_p = 2$, then $q_2 = p$ and the proof is finished. Otherwise, we note that since r and z are arbitrary, estimate (8.6) also holds in any ball that is contained in $B_r(z)$, which means that estimate (6.8) is satisfied with respect to $q = q_2$ in any ball contained in $B_r(z)$. Since also $U \in L^p(\mathcal{B}_r(z), \mu) \subset L^{q_3}(\mathcal{B}_r(z), \mu)$, if we choose δ smaller if necessary such that Corollary 8.4 is applicable with $q = q_2$ and $\tilde{q} = q_3$, then all assumptions of Corollary 8.4 are satisfied with respect to $q = q_2$ and $\tilde{q} = q_3 = (q_i + (q_i)^\star)/2 \in (q_2, (q_2)^\star)$, so that we obtain the estimate

$$\left(\fint_{\mathcal{B}_{r/2}(z)} U^{q_3}\, d\mu\right)^{\frac{1}{q_3}} \leq C_2\left(\sum_{k=1}^{\infty} 2^{-k(s-\theta)}\left(\fint_{\mathcal{B}_{2^k r(z)}} U^2\, d\mu\right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_r(z)} G^{q_3}\, d\mu\right)^{\frac{1}{q_3}} + \sum_{k=1}^{\infty} 2^{-k(s-\theta)}\left(\fint_{\mathcal{B}_{2^k r}(z)} G^2\, d\mu\right)^{\frac{1}{2}}\right),$$

where $C_2 = C_2(n, s, \theta, \Lambda, p) > 0$. By iterating this procedure $i_p - 1$ times and using that $q_{i_p} = p$, we finally arrive at estimate (8.5). ∎

9. Proofs of the main results

We are now in the position to prove our main results.

Theorem 9.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $s \in (0, 1)$, $\Lambda \geq 1$, $R > 0$ and $p \in (2, \infty)$. Moreover, fix some t such that*

$$s < t < \min\left\{2s\left(1 - \frac{1}{p}\right), 1 - \frac{2 - 2s}{p}\right\}. \tag{9.1}$$

Then there exists some small enough $\delta = \delta(p, n, s, t, \Lambda) > 0$, such that if $A \in \mathcal{L}_0(\Lambda)$ is (δ, R)-BMO in Ω and if Φ satisfies conditions (1.5) and (1.6) with respect to Λ, then for any weak solution $u \in W^{s,2}(\mathbb{R}^n)$ of the equation

$$L_A^\Phi u = f \quad \text{in } \Omega,$$

we have the implication

$$f \in L_{\mathrm{loc}}^{\frac{np}{n+(2s-t)p}}(\Omega) \Rightarrow u \in W_{\mathrm{loc}}^{t,p}(\Omega).$$

Moreover, for all relatively compact bounded open sets $\Omega' \Subset \Omega'' \Subset \Omega$, we have the estimate

$$[u]_{W^{t,p}(\Omega')} \leq C\left([u]_{W^{s,2}(\mathbb{R}^n)} + \|f\|_{L^{\frac{np}{n+(2s-t)p}}(\Omega'')}\right), \tag{9.2}$$

where $C = C(n, s, t, \Lambda, R, p, \Omega', \Omega'') > 0$.

Remark 9.2. Note that in view of Proposition 2.6, the conclusion that $u \in W_{\mathrm{loc}}^{t,p}(\Omega)$ for some $p \in (2, \infty)$ and for any t in the range (9.1) also implies that $u \in W_{\mathrm{loc}}^{t,p_0}(\Omega)$ for any $p_0 \in (1, p]$.

Proof of Theorem 9.1. Fix relatively compact bounded open sets $\Omega' \Subset \Omega'' \Subset \Omega$. Let $\delta = \delta(p,n,s,\theta,\Lambda) > 0$ be given by Proposition 8.6. There exists some small enough $r_1 \in (0,1)$ such that $2r_1 \leq R$ and $B_{2r_1}(z) \Subset \Omega''$ for any $z \in \Omega'$. Now fix some $z \in \Omega''$. Since A is (δ, R)-BMO in Ω, we obtain that A is δ-vanishing in $B_{2r_1}(z)$. Let $\{\psi_m\}_{m=1}^{\infty}$ be a sequence of standard mollifiers in $\mathbb{R}^n$ with the properties

$$\psi_m \in C_0^\infty(B_{1/m}), \quad \psi_m \geq 0, \quad \int_{\mathbb{R}^n} \psi_m(x)\,dx = 1 \quad \text{for all } m \in \mathbb{N}. \tag{9.3}$$

In addition, for any $m \in \mathbb{N}$ we define

$$A_m(x,y) := \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} A(x-x', y-y')\psi_m(x')\psi_m(y')\,dy'\,dx'.$$

Clearly, A_m is symmetric and belongs to $\mathcal{L}_0(\Lambda)$ for any $m \in \mathbb{N}$. In addition, there exists some large enough $m_0 \in \mathbb{N}$ such that

$$\frac{1}{m_0} < \min\{r_1, \operatorname{dist}(B_{2r_1}(z), \Omega)\}.$$

Fix $r > 0$ and $x_0, y_0 \in B_{r_1}(z)$ with $B_r(x_0) \subset B_{r_1}(z)$, $B_r(y_0) \subset B_{r_1}(z)$. Then for any $m \geq m_0$ and all $x', y' \in B_{1/m}$, we have $B_r(x_0 - x') \subset B_{2r_1}(z)$ and $B_r(y_0 - y') \subset B_{2r_1}(z)$. Therefore, since A is δ-vanishing in $B_{2r_1}(z)$, for any $m \geq m_0$ and $x', y' \in B_{1/m}$, we have

$$\fint_{B_r(x_0-x')} \fint_{B_r(y_0-y')} |A(x,y) - \bar{A}_{r,x_0-x',y_0-y'}|\,dy\,dx \leq \delta.$$

Therefore, together with changes of variables, Fubini's theorem and (9.3), we obtain

$$\begin{aligned}
&\fint_{B_r(x_0)} \fint_{B_r(y_0)} |A_m(x,y) - \overline{(A_m)}_{r,x_0,y_0}|\,dy\,dx \\
&\quad \leq \int_{B_{\frac{1}{m}}} \int_{B_{\frac{1}{m}}} \fint_{B_r(x_0)} \fint_{B_r(y_0)} \Bigg| A(x-x', y-y') \\
&\qquad\qquad - \fint_{B_r(x_0)} \fint_{B_r(y_0)} A(x_1 - x', y_1 - y')\,dy_1\,dx_1 \Bigg|\,dy\,dx \\
&\qquad\qquad \times \psi_m(x')\psi_m(y')\,dy'\,dx' \\
&\quad = \int_{B_{\frac{1}{m}}} \int_{B_{\frac{1}{m}}} \left(\fint_{B_r(x_0-x')} \fint_{B_r(y_0-y')} |A(x,y) - \bar{A}_{r,x_0-x',y_0-y'}|\,dy\,dx \right) \\
&\qquad\qquad \times \psi_m(x')\psi_m(y')\,dy'\,dx' \\
&\quad \leq \delta \int_{B_{\frac{1}{m}}} \int_{B_{\frac{1}{m}}} \psi_m(x')\psi_m(y')\,dy'\,dx' = \delta,
\end{aligned}$$

so that we conclude that A_m is δ-vanishing in $B_{r_1}(z)$ for any $m \geq m_0$. Now define

$$\tilde{A}_m(x,y) := \begin{cases} A_m(x,y) & \text{if } (x,y) \in \mathcal{B}_{2r_1}(z), \\ A(x,y) & \text{if } (x,y) \notin \mathcal{B}_{2r_1}(z). \end{cases}$$

Since $A \in L^\infty(\mathbb{R}^n) \subset L^1_{\mathrm{loc}}(\mathbb{R}^n)$, by standard properties of mollifiers we have

$$\tilde{A}_m \xrightarrow{m\to\infty} A \quad \text{in } L^1(\mathcal{B}_{2r_1}(z)) \tag{9.4}$$

and also $\tilde{A}_m(x,y) \in C^\infty(\mathcal{B}_{2r_1}(z))$. In particular, $\tilde{A}_m$ is continuous inside $B_{2r_1}(z) \times B_{2r_1}(z)$.

Next, for any $m \in \mathbb{N}$ and $x \in \Omega_m := \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) > 1/m\}$, define

$$f_m(x) := \int_\Omega f(y)\psi_m(x-y)\,dy.$$

Since $f \in L_{\mathrm{loc}}^{\frac{np}{n+(2s-t)p}}(\Omega)$ and $B_{2r_1}(z) \Subset \Omega$, again by standard properties of mollifiers we have

$$f_m \xrightarrow{m\to\infty} f \quad \text{in } L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z)) \tag{9.5}$$

and $f_m \in C^\infty(B_{2r_1}(z)) \subset L^\infty(B_{\frac{3}{2}r_1}(z))$. Next, for any $m \geq m_0$ we let $u_m \in W^{s,2}(\mathbb{R}^n)$ be the unique weak solution of the Dirichlet problem

$$\begin{cases} L^\Phi_{\tilde{A}_m} u_m = f_m & \text{in } B_{2r_1}(z), \\ u_m = u & \text{a.e. in } \mathbb{R}^n \setminus B_{2r_1}(z). \end{cases} \tag{9.6}$$

Since $2_\star = \frac{2n}{n+2s} < \frac{np}{n+(2s-t)p}$, we can choose the number $\sigma_0 > 0$ from Theorem 4.3 small enough such that

$$2_\star + \sigma_0 \leq \frac{np}{n+(2s-t)p}.$$

Then by Proposition 5.2, (6.12), Hölder's inequality, (9.4) and (9.5), for $w_m := u - u_m$ we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{(w_m(x)-w_m(y))^2}{|x-y|^{n+2s}}\,dy\,dx \\
&\leq C_1 \omega(A-\tilde{A}_m, 2r_1, x_0)^{\frac{\gamma}{n-\gamma}} \mu(\mathcal{B}_{r_1}(z)) \Bigg(\sum_{k=1}^\infty 2^{-k(s-\theta)} \bigg(\fint_{\mathcal{B}_{2^k r_1}(z)} U^2\,d\mu \bigg)^{\frac{1}{2}} \Bigg)^2 \\
&\quad + C_1 \omega(A-\tilde{A}_m, 2r_1, z)^{\frac{\gamma}{n-\gamma}} r_1^{2(s-\theta)} \mu(\mathcal{B}_{r_1}(z)) \bigg(\fint_{B_{2r_1}(z)} |f|^{2_\star+\sigma_0}\,dx \bigg)^{\frac{2}{2_\star+\sigma_0}} \\
&\quad + C_1 r_1^{2(s-\theta)} \mu(\mathcal{B}_{r_1}(z)) \bigg(\fint_{B_{2r_1}(z)} |f-f_m|^{2_\star}\,dx \bigg)^{\frac{2}{2_\star}} \\
&\leq C_2 \Big(\|A-\tilde{A}_m\|^{\frac{\gamma}{n-\gamma}}_{L^1(\mathcal{B}_{2r_1}(z))} [u]^2_{W^{s,2}(\mathbb{R}^n)} + \|A-\tilde{A}_m\|^{\frac{\gamma}{n-\gamma}}_{L^1(\mathcal{B}_{2r_1}(z))} \|f\|^2_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))} \\
&\qquad + \|f-f_m\|^2_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))} \Big) \xrightarrow{k\to\infty} 0,
\end{aligned}$$

where $C_1 = C_1(n, s, \theta, \Lambda, \sigma_0) > 0$ and $C_2 = C_2(n, s, \theta, \Lambda, \sigma_0, r_1) > 0$. Thus, we deduce

$$\lim_{m\to\infty} [u_m]_{W^{s,2}(\mathbb{R}^n)} = [u]_{W^{s,2}(\mathbb{R}^n)}. \tag{9.7}$$

Next, for any $m \in \mathbb{N}$, let $g_m \in W^{s,2}(\mathbb{R}^n)$ be the unique weak solution of the Dirichlet problem

$$\begin{cases} (-\Delta)^s g_m = f_m & \text{in } B_{2r_1}(z), \\ g_m = 0 & \text{a.e. in } \mathbb{R}^n \setminus B_{2r_1}(z). \end{cases} \tag{9.8}$$

Then by Proposition 5.2, we have the estimate

$$[g_m]_{W^{s,2}(\mathbb{R}^n)} \leq C_3 \| f_m \|_{L^{2_\star}(B_{2r_1}(z))} \leq C_4 \| f_m \|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))}, \tag{9.9}$$

where $C_3 = C_3(n, s, r_1) > 0$ and $C_4 = C_4(n, s, t, p, r_1) > 0$. In addition, by Theorem 4.4 we have the estimate

$$\| g_m \|_{H^{2s, \frac{np}{n+(2s-t)p}}(B_{r_1}(z))} \leq C_5 \| f_m \|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))}, \tag{9.10}$$

where $C_5 = C_5(n, s, t, p) > 0$. Also, by Proposition 2.5, we have

$$[g_m]_{W^{t,p}(B_{r_1}(z))} \leq C_6 \| g_m \|_{H^{2s, \frac{np}{n+(2s-t)p}}(B_{r_1}(z))}, \tag{9.11}$$

where $C_6 = C_6(n, s, t, p) > 0$. In view of (9.6) and (9.8), u_m is a weak solution of the equation

$$L^{\Phi}_{\tilde{A}_m} u_m = (-\Delta)^s g_m \quad \text{in } B_{2r_1}(z).$$

Define

$$U_m(x, y) := \frac{|u_m(x) - u_m(y)|}{|x - y|^{s+\theta}}, \quad G_m(x, y) := \frac{|g_m(x) - g_m(y)|}{|x - y|^{s+\theta}}.$$

Since $\tilde{A}_m$ is continuous in $B_{2r_1}(z) \times B_{2r_1}(z)$ and $f_m \in L^\infty(B_{\frac{3}{2}r_1}(z))$, by Theorem 4.1 we have $u_m \in C^{s+\theta}(B_{r_1}(z))$ and therefore $U_m \in L^\infty(B_{r_1}(z), \mu) \subset L^p(B_{r_1}(x_0), \mu)$. Therefore, by Proposition 8.6, (6.12), (9.9), (9.11) and (9.10), we have

$$\begin{aligned}
\left(\fint_{\mathcal{B}_{r_1/2}(z)} U_m^p \, d\mu \right)^{\frac{1}{p}} &\leq C_7 \Bigg(\sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r_1}(z)} U_m^2 \, d\mu \right)^{\frac{1}{2}} + \left(\fint_{\mathcal{B}_{r_1}(z)} G_m^p \, d\mu \right)^{\frac{1}{p}} \\
&\quad + \sum_{k=1}^{\infty} 2^{-k(s-\theta)} \left(\fint_{\mathcal{B}_{2^k r_1}(z)} G_m^2 \, d\mu \right)^{\frac{1}{2}} \Bigg) \\
&\leq C_8 \left([u_m]_{W^{s,2}(\mathbb{R}^n)} + [g_m]_{W^{t,p}(B_{r_1}(z))} + [g_m]_{W^{s,2}(\mathbb{R}^n)} \right) \\
&\leq C_9 \left([u_m]_{W^{s,2}(\mathbb{R}^n)} + \| f_m \|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))} \right),
\end{aligned}$$

where all constants depend only on n, s, t, θ, Λ, p and r_1. Combining the previous display with Fatou's lemma (which is applicable after passing to a subsequence if necessary), (9.7) and (9.5), we conclude that

$$\begin{aligned}\Big(\fint_{\mathcal{B}_{r_1/2}(z)} U^p \, d\mu\Big)^{\frac{1}{p}} &\leq \liminf_{m\to\infty}\Big(\fint_{\mathcal{B}_{r_1/2}(z)} U_m^p \, d\mu\Big)^{\frac{1}{p}} \\ &\leq C_{10} \lim_{m\to\infty}\big([u_m]_{W^{s,2}(\mathbb{R}^n)} + \|f_m\|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))}\big) \\ &= C_{10}\big([u]_{W^{s,2}(\mathbb{R}^n)} + \|f\|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z))}\big), \end{aligned} \tag{9.12}$$

where $C_{10} = C_{10}(n, s, t, \theta, \Lambda, p, r_1) > 0$.

Since $\{B_{r_1/2}(z)\}_{z\in\Omega'}$ is an open covering of $\overline{\Omega'}$ and $\overline{\Omega'}$ is compact, there exists a finite subcover $\{B_{r_1/2}(z_i)\}_{i=1}^N$ of $\overline{\Omega'}$ and hence of Ω'. Now summing over $i = 1, \dots, N$ and using estimate (9.12) for any i, we arrive at

$$\begin{aligned}\Big(\int_{\Omega'\times\Omega'} U^p \, d\mu\Big)^{\frac{1}{p}} &\leq \sum_{i=1}^N \Big(\int_{\mathcal{B}_{r_1/2}(z_i)} U^p \, d\mu\Big)^{\frac{1}{p}} \\ &\leq \sum_{i=1}^N C_{11}\big([u]_{W^{s,2}(\mathbb{R}^n)} + \|f\|_{L^{\frac{np}{n+(2s-t)p}}(B_{2r_1}(z_i))}\big) \\ &\leq C_{11} N\big([u]_{W^{s,2}(\mathbb{R}^n)} + \|f\|_{L^{\frac{np}{n+(2s-t)p}}(\Omega'')}\big), \end{aligned} \tag{9.13}$$

where $C_{11} = C_{11}(n, s, t, \theta, \Lambda, p, r_1) > 0$. Clearly, for any t in the range (9.1), there exists some $0 < \theta < \min\{s, 1-s\}$ such that $t = s + \theta(1 - \frac{2}{p})$, so that by choosing this θ in our definition of μ, we arrive at

$$[u]_{W^{t,p}(\Omega')} = \Big(\int_{\Omega'\times\Omega'} U^p \, d\mu\Big)^{\frac{1}{p}} \leq C\big([u]_{W^{s,2}(\mathbb{R}^n)} + \|f\|_{L^{\frac{np}{n+(2s-t)p}}(\Omega'')}\big),$$

where $C = C(n, s, t, \Lambda, p, \Omega', \Omega'') > 0$. Here we also used that r_1 depends only on R, Ω' and Ω'' and that θ depends only on s, p and t. This proves estimate (9.2).

Next, let us prove that $u \in L^p_{\mathrm{loc}}(\Omega)$. For any $q \in (2, p]$, we fix some

$$s < t_q < \min\Big\{2s\Big(1 - \frac{1}{q}\Big), 1 - \frac{2-2s}{q}\Big\}$$

and define

$$q^\star := \begin{cases} \min\Big\{\dfrac{nq}{n - t_q q}, p\Big\} & \text{if } t_q q < n, \\ p & \text{if } t_q q \geq n. \end{cases}$$

Since $u \in W^{s,2}(\mathbb{R}^n)$, by the Sobolev embedding (Proposition 2.3) we have $u \in L^{2^\star}_{\mathrm{loc}}(\Omega)$, where $2^\star := \min\{\frac{2n}{n-2s}, p\}$. If $p = 2^\star$, the proof is finished. Otherwise, together with estimate (9.2) with p replaced by $2^\star$, we conclude that $u \in W^{t_{2^\star},2^\star}_{\mathrm{loc}}(\Omega)$. Again by Proposition 2.3, we then obtain that $u \in L^{2^{\star\star}}_{\mathrm{loc}}(\Omega)$. If $p = 2^{\star\star}$, the proof is finished. Otherwise, iterating this procedure also leads to the conclusion that $u \in L^p_{\mathrm{loc}}(\Omega)$ at some point, so that $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$. This finishes the proof. ∎

Proof of Theorem 1.1. Let us first handle the case when $t > s$. Since A is assumed to be VMO in Ω, for any $\delta > 0$, there exists some $R > 0$ such that A is (δ, R)-BMO in Ω. Therefore, in this case Theorem 1.1 follows directly from Theorem 9.1. This finishes the proof in the case when $t > s$.

In the case when $t = s$, fix some small enough $\varepsilon > 0$ such that $\tilde{s} := s + \varepsilon$ belongs to the range (9.1) and $\tilde{p} := \frac{np}{n+\varepsilon p} > 2$. Then by assumption and an elementary computation, we have

$$f \in L^{\frac{np}{n+sp}}_{\mathrm{loc}}(\Omega) = L^{\frac{n\tilde{p}}{n+(2s-\tilde{s})\tilde{p}}}_{\mathrm{loc}}(\Omega).$$

By applying the previous case when $t > s$ with $t = \tilde{s}$ and with p replaced by $\tilde{p}$, we obtain that $u \in W^{\tilde{s},\tilde{p}}_{\mathrm{loc}}(\Omega)$, which by Proposition 2.5 leads to $u \in W^{s,p}_{\mathrm{loc}}(\Omega)$. Thus, the proof is finished. ∎

Proof of Theorem 1.3. Fix some t such that $s \leq t < 1$. First, we assume that t satisfies (1.11). Then we have $n > (2s-t)q$ and set $p := \frac{nq}{n-(2s-t)q}$, so that we have $q = \frac{np}{n+(2s-t)p}$ and thus $f \in L^{\frac{np}{n+(2s-t)p}}_{\mathrm{loc}}(\Omega)$. Then in view of (1.11) and elementary computations, we obtain that $p > 2$ and

$$s \leq t < \min\Big\{2s\Big(1 - \frac{1}{p}\Big), 1 - \frac{2-2s}{p}\Big\}$$

so that by Theorem 1.1 we obtain $u \in W^{t,p}_{\mathrm{loc}}(\Omega) = W^{t,\frac{nq}{n-(2s-t)q}}_{\mathrm{loc}}(\Omega)$.

Next, suppose that $t = 2s - \frac{n}{q}$. Since $t < 1$, in this case we have $2s - \frac{n}{q} < 1$. Using the latter inequality, a direct computation shows that there exists some $t' \geq s$ such that

$$2s - \frac{n}{q} < t' < \min\Big\{2s\Big(1 - \frac{n}{(n+2s)q}\Big), 1 - \frac{(2-2s)(n+q-2sq)}{(n+2-2s)q}\Big\}. \tag{9.14}$$

Then by the previous case, we obtain that $u \in W^{t',\frac{nq}{n-(2s-t')q}}_{\mathrm{loc}}(\Omega)$, which by Propositions 2.5 and 2.6 implies that $u \in W^{2s-\frac{n}{q},p}_{\mathrm{loc}}(\Omega) = W^{t,p}_{\mathrm{loc}}(\Omega)$ for any $p \in (1,\infty)$.

Finally, if we have $2s - \frac{n}{q} > t$, then there exists some $\varepsilon > 0$ such that $2s - \frac{n}{q} > t + \varepsilon$. Then for any $p > 1$, we have $q \geq \frac{np}{n+(2s-t-\varepsilon)p}$ and therefore $f \in L^{\frac{np}{n+(2s-t-\varepsilon)p}}_{\mathrm{loc}}(\Omega)$. Furthermore, for $p > \max\{\frac{2s}{2s-t}, \frac{2-2s}{1-t}\}$, we see that

$$s \leq t + \varepsilon < \min\Big\{2s\Big(1 - \frac{1}{p}\Big), 1 - \frac{2-2s}{p}\Big\},$$

so that by Theorem 1.1 we obtain that $u \in W^{t+\varepsilon,p}_{\mathrm{loc}}(\Omega)$ for any $p > \max\{\frac{2s}{2s-t}, \frac{2-2s}{1-t}\}$, which by Proposition 2.6 implies that $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$ for any $p \in (1,\infty)$, which finishes the proof. ∎

Proof of Theorem 1.4. First of all, we remark that by the assumption that $q > \frac{n}{2s}$, we always have $2s - \frac{n}{q} > 0$. Moreover, by a simple computation we have $q > \frac{2n}{n+2(2s-t)}$ for any $t < 1$. Now consider the case when $0 < 2s - \frac{n}{q} < 1$. Then, as in (9.14), there exists some $t \geq s$ such that

$$2s - \frac{n}{q} < t < \min\Big\{2s\Big(1 - \frac{n}{(n+2s)q}\Big), 1 - \frac{(2-2s)(n+q-2sq)}{(n+2-2s)q}\Big\},$$

so that by Theorem 1.3, for $p := \frac{nq}{n-(2s-t)q}$ we obtain that $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$. Since in addition we have

$$t - \frac{n}{p} = t - \frac{n(n-(2s-t)q)}{nq} = 2s - \frac{n}{q} > 0,$$

by the Sobolev embedding (Proposition 2.3) we conclude that $u \in C^{2s-\frac{n}{q}}_{\mathrm{loc}}(\Omega)$.

Next, consider the case when $2s - \frac{n}{q} \geq 1$. In this case, by Theorem 1.3 we obtain $u \in W^{t,p}_{\mathrm{loc}}(\Omega)$ for any $s \leq t < 1$ and any $p \in (1,\infty)$, which by the Sobolev embedding implies that $u \in C^{\alpha}_{\mathrm{loc}}(\Omega)$ for any $\alpha \in (0,1)$. This finishes the proof. ∎

Remark 9.3. Our main results remain valid for another large class of coefficients A that in general might not be VMO. Namely, the conclusions of Theorems 1.1, 1.3 and 1.4 remain true if we replace the assumption that A is VMO with the following assumption used, for example, in [35]: namely, our main results remain true if there exists some small $\varepsilon > 0$ such that

$$\lim_{h\to 0} \sup_{\substack{x,y\in K \\ |x-y|\leq\varepsilon}} |A(x+h,y+h) - A(x,y)| = 0 \quad \text{for any compact set } K \subset \Omega. \tag{9.15}$$

This is because by [35, Theorem 1.1], the Hölder estimate from Theorem 4.1 remains valid under assumption (9.15). Therefore, in contrast to the case when A is VMO, under assumption (9.15) the above proof can be executed without the need to freeze the coefficient A, so that the proof actually simplifies in this case. Condition (9.15) is for example satisfied in the case when A is translation invariant in Ω, that is, if there exists a measurable function $a: \mathbb{R}^n \to \mathbb{R}$ such that $A(x,y) = a(x-y)$ for all $x, y \in \Omega$. Since in this case A is otherwise not required to satisfy any additional smoothness assumption, A might not be VMO in Ω but still satisfies (9.15).

Funding. The author is supported by grant SFB 1283 of the German Research Foundation.

References

[1] E. Acerbi and G. Mingione, Gradient estimates for a class of parabolic systems. *Duke Math. J.* **136** (2007), no. 2, 285–320 Zbl 1113.35105 MR 2286632

[2] A. K. Balci, L. Diening, R. Giova, and A. Passarelli di Napoli, Elliptic equations with degenerate weights. *SIAM J. Math. Anal.* **54** (2022), no. 2, 2373–2412 MR 4410267

[3] U. Biccari, M. Warma, and E. Zuazua, Local elliptic regularity for the Dirichlet fractional Laplacian. *Adv. Nonlinear Stud.* **17** (2017), no. 2, 387–409 Zbl 1360.35033 MR 3641649

[4] L. Brasco and E. Lindgren, Higher Sobolev regularity for the fractional p-Laplace equation in the superquadratic case. *Adv. Math.* **304** (2017), 300–354 Zbl 1364.35055 MR 3558212

[5] L. Brasco, E. Lindgren, and A. Schikorra, Higher Hölder regularity for the fractional p-Laplacian in the superquadratic case. *Adv. Math.* **338** (2018), 782–846 Zbl 1400.35049 MR 3861716

[6] S.-S. Byun, Elliptic equations with BMO coefficients in Lipschitz domains. *Trans. Amer. Math. Soc.* **357** (2005), no. 3, 1025–1046 Zbl 1087.35027 MR 2110431

[7] L. Caffarelli and L. Silvestre, Regularity results for nonlocal equations by approximation. *Arch. Ration. Mech. Anal.* **200** (2011), no. 1, 59–88 Zbl 1231.35284 MR 2781586

[8] L. A. Caffarelli and I. Peral, On $W^{1,p}$ estimates for elliptic equations in divergence form. *Comm. Pure Appl. Math.* **51** (1998), no. 1, 1–21 Zbl 0906.35030 MR 1486629

[9] L. A. Caffarelli and P. R. Stinga, Fractional elliptic equations, Caccioppoli estimates and regularity. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **33** (2016), no. 3, 767–807 Zbl 1381.35211 MR 3489634

[10] J. Chaker and M. Kassmann, Nonlocal operators with singular anisotropic kernels. *Comm. Partial Differential Equations* **45** (2020), no. 1, 1–31 Zbl 07143800 MR 4037095

[11] J. Chaker and M. Kim, Regularity estimates for fractional orthotropic p-Laplacians of mixed order. *Adv. Nonlinear Anal.* **11** (2022), no. 1, 1307–1331 MR 4399827

[12] M. Cozzi, Interior regularity of solutions of non-local equations in Sobolev and Nikol'skii spaces. *Ann. Mat. Pura Appl. (4)* **196** (2017), no. 2, 555–578 Zbl 1371.35312 MR 3624965

[13] C. De Filippis and G. Palatucci, Hölder regularity for nonlocal double phase equations. *J. Differential Equations* **267** (2019), no. 1, 547–586 Zbl 1412.35041 MR 3944281

[14] A. Di Castro, T. Kuusi, and G. Palatucci, Local behavior of fractional p-minimizers. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **33** (2016), no. 5, 1279–1299 Zbl 1355.35192 MR 3542614

[15] G. Di Fazio, L^p estimates for divergence form elliptic equations with discontinuous coefficients. *Boll. Un. Mat. Ital. A (7)* **10** (1996), no. 2, 409–420 Zbl 0865.35048 MR 1405255

[16] E. Di Nezza, G. Palatucci, and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** (2012), no. 5, 521–573 Zbl 1252.46023 MR 2944369

[17] H. Dong and D. Kim, Elliptic equations in divergence form with partially BMO coefficients. *Arch. Ration. Mech. Anal.* **196** (2010), no. 1, 25–70 Zbl 1206.35249 MR 2601069

[18] H. Dong and D. Kim, On L_p-estimates for a class of non-local elliptic equations. *J. Funct. Anal.* **262** (2012), no. 3, 1166–1199 Zbl 1232.35182 MR 2863859

[19] M. M. Fall, Regularity estimates for nonlocal Schrödinger equations. *Discrete Contin. Dyn. Syst.* **39** (2019), no. 3, 1405–1456 Zbl 1407.35209 MR 3918224

[20] M. M. Fall, Regularity results for nonlocal equations and applications. *Calc. Var. Partial Differential Equations* **59** (2020), no. 5, Paper No. 181 Zbl 1450.35093 MR 4153907

[21] G. F. Foghem Gounoue, M. Kassmann, and P. Voigt, Mosco convergence of nonlocal to local quadratic forms. *Nonlinear Anal.* **193** (2020), 111504 Zbl 1437.49027 MR 4062977

[22] G. Grubb, Fractional Laplacians on domains, a development of Hörmander's theory of μ-transmission pseudodifferential operators. *Adv. Math.* **268** (2015), 478–528 Zbl 1318.47064 MR 3276603

[23] T. Iwaniec and C. Sbordone, Riesz transforms and elliptic PDEs with VMO coefficients. *J. Anal. Math.* **74** (1998), 183–212 Zbl 0909.35039 MR 1631658

[24] M. Kassmann, A priori estimates for integro-differential operators with measurable kernels. *Calc. Var. Partial Differential Equations* **34** (2009), no. 1, 1–21 Zbl 1158.35019 MR 2448308

[25] M. Kassmann, T. Mengesha, and J. Scott, Solvability of nonlocal systems related to peridynamics. *Commun. Pure Appl. Anal.* **18** (2019), no. 3, 1303–1332 Zbl 1470.74010 MR 3917708

[26] J. Kinnunen and S. Zhou, A local estimate for nonlinear equations with discontinuous coefficients. *Comm. Partial Differential Equations* **24** (1999), no. 11-12, 2043–2068 Zbl 0941.35026 MR 1720770

[27] T. Kuusi, G. Mingione, and Y. Sire, Nonlocal equations with measure data. *Comm. Math. Phys.* **337** (2015), no. 3, 1317–1368 Zbl 1323.45007 MR 3339179

[28] T. Kuusi, G. Mingione, and Y. Sire, Nonlocal self-improving properties. *Anal. PDE* **8** (2015), no. 1, 57–114 Zbl 1317.35284 MR 3336922

[29] T. Leonori, I. Peral, A. Primo, and F. Soria, Basic estimates for solutions of a class of nonlocal elliptic and parabolic equations. *Discrete Contin. Dyn. Syst.* **35** (2015), no. 12, 6031–6068 Zbl 1332.45009 MR 3393266

[30] A. Maugeri, D. K. Palagachev, and L. G. Softova, *Elliptic and parabolic equations with discontinuous coefficients*. Mathematical Research 109, Wiley-VCH, Berlin, 2000 Zbl 0958.35002 MR 2260015

[31] T. Mengesha and T. Phan, Weighted $W^{1,p}$ estimates for weak solutions of degenerate elliptic equations with coefficients degenerate in one variable. *Nonlinear Anal.* **179** (2019), 184–236 Zbl 1404.35195 MR 3886630

[32] T. Mengesha, A. Schikorra, and S. Yeepo, Calderon-Zygmund type estimates for nonlocal PDE with Hölder continuous kernel. *Adv. Math.* **383** (2021), Paper No. 107692 Zbl 1462.35117 MR 4233278

[33] G. Mingione, The singular set of solutions to non-differentiable elliptic systems. *Arch. Ration. Mech. Anal.* **166** (2003), no. 4, 287–301 Zbl 1142.35391 MR 1961442

[34] S. Nowak, $H^{s,p}$ regularity theory for a class of nonlocal elliptic equations. *Nonlinear Anal.* **195** (2020), 111730 MR 4046549

[35] S. Nowak, Higher Hölder regularity for nonlocal equations with irregular kernel. *Calc. Var. Partial Differential Equations* **60** (2021), no. 1, Paper No. 24 Zbl 07309168 MR 4201647

[36] S. Nowak, Higher integrability for nonlinear nonlocal equations with irregular kernel. In *Analysis and partial differential equations on manifolds, fractals and graphs*, pp. 459–492, Adv. Anal. Geom. 3, De Gruyter, Berlin, 2021 Zbl 1476.35068 MR 4320100

[37] X. Ros-Oton and J. Serra, The Dirichlet problem for the fractional Laplacian: regularity up to the boundary. *J. Math. Pures Appl. (9)* **101** (2014), no. 3, 275–302 Zbl 1285.35020 MR 3168912

[38] X. Ros-Oton and J. Serra, Regularity theory for general stable operators. *J. Differential Equations* **260** (2016), no. 12, 8675–8715 Zbl 1346.35220 MR 3482695

[39] D. Sarason, Functions of vanishing mean oscillation. *Trans. Amer. Math. Soc.* **207** (1975), 391–405 Zbl 0319.42006 MR 377518

[40] A. Schikorra, Nonlinear commutators for the fractional p-Laplacian and applications. *Math. Ann.* **366** (2016), no. 1-2, 695–720 Zbl 1351.35255 MR 3552254

[41] L. Silvestre, Hölder estimates for solutions of integro-differential equations like the fractional Laplace. *Indiana Univ. Math. J.* **55** (2006), no. 3, 1155–1174 Zbl 1101.45004 MR 2244602

[42] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Math. Ser. 43, Princeton University Press, Princeton, NJ, 1993 Zbl 0821.42001 MR 1232192

[43] H. Triebel, *Theory of function spaces. IV*. Monogr. Math. 107, Birkhäuser/Springer, Cham, 2020 Zbl 1445.46002 MR 4298338

[44] S. Yeepo, W. Lewkeeratiyutkul, S. Khomrutai, and A. Schikorra, On the Calderon-Zygmund property of Riesz-transform type operators arising in nonlocal equations. *Commun. Pure Appl. Anal.* **20** (2021), no. 9, 2915–2939 Zbl 1477.45007 MR 4315490

Received 14 June 2021; revised 19 October 2021; accepted 22 October 2021.

Simon Nowak
Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany;
simon.nowak@uni-bielefeld.de