A NOTE ON ESTIMATES OF LEVEL SETS AND THEIR ROLE IN DEMONSTRATING REGULARITY OF SOLUTIONS TO NONLOCAL DOUBLE PHASE EQUATIONS

TADELE MENGESHA AND JAMES M. SCOTT

ABSTRACT. In this note we prove an estimate on the level sets of a function with (p, q) growth that depends on the difference quotient of a bounded weak solution to a nonlocal double phase equation. This estimate is related to a self improving property of these solutions.

1. INTRODUCTION AND MAIN RESULTS

This note is a companion to the article [6], in which the authors study regularity properties of weak solutions u to

$$(1.1) \qquad \mathcal{L}u(x) = f(x),$$

where for measurable functions $u : \mathbb{R}^n \to \mathbb{R}$ and for $x \in \mathbb{R}^n$ the nonlocal *double phase* operator $\mathcal{L}$ is defined as

$$\mathcal{L}u(x) := \mathrm{P.V.} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^{p-2}}{|x - y|^{n+sp}} (u(x) - u(y)) + a(x, y) \frac{|u(x) - u(y)|^{q-2}}{|x - y|^{n+tq}} (u(x) - u(y)) \, \mathrm{d}y .$$

Throughout, we assume $n \geq 2$ and the integrability indices p, q belong to $(1, \infty)$ with $p \leq q$ and differentiability indices s, t belong to $(0, 1)$. The abbreviation P.V. stands for principal value.

The operator $\mathcal{L}$ is the archetype of a class of nonlocal double phase operators first introduced in [2], in which the Hölder continuity of bounded viscosity solutions to $\mathcal{L}u = f$ with bounded data f was obtained. In the work [6] the authors demonstrate regularity of solutions on a different scale; that under suitable assumptions on the data f, the modulating coefficient $a(\cdot, \cdot)$, and a certain ratio of integrability and differentiability exponents solutions u to $\mathcal{L}u = f$ exhibit a self-improvement property. Precisely, distributional solutions u belonging to the fractional Sobolev space $W^{s,p}(\mathbb{R}^n)$ in fact belong to a Sobolev space with higher exponents of integrability and differentiability.

For ease of reference, we summarize the relevant definitions, assumptions, properties and results found in [6]. We assume that the modulating coefficient a is measurable, and satisfies

$$(A1) \qquad a(x, y) \in L^\infty(\mathbb{R}^{2n}), \qquad 0 \leq a(x, y) \leq M, \qquad a(x, y) = a(y, x).$$

We also require that

$$(A2) \qquad p \leq q, \quad t \leq s, \qquad \frac{1}{p'} \leq \frac{tq}{sp} \leq 1,$$

where p' is the Hölder conjugate of p: $\frac{1}{p} + \frac{1}{p'} = 1$. Additionally we will restrict ourselves to the case

$$(A3) \qquad sp < n.$$

1

The higher differentiability and integrability described above and in [6] applies to bounded solutions $u \in W^{s,p}(\mathbb{R}^n)$ of a weak formulation of the equation (1.1), that is

$$(1.2) \qquad \mathcal{E}(u, \varphi) = \int_{\mathbb{R}^n} f(x)\varphi(x)\,\mathrm{d}x\,, \qquad \text{for any } \varphi \in C_c^\infty(\mathbb{R}^n)\,,$$

where the form $\mathcal{E}(u, \varphi)$ is defined as

$$
\begin{aligned}
(1.3) \quad \mathcal{E}(u, \varphi) &:= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^{p-2}}{|x - y|^{n+sp}}(u(x) - u(y))(\varphi(x) - \varphi(y)) \\
&\qquad + a(x, y)\frac{|u(x) - u(y)|^{q-2}}{|x - y|^{n+tq}}(u(x) - u(y))(\varphi(x) - \varphi(y))\,\mathrm{d}y\,\mathrm{d}x\,.
\end{aligned}
$$

We assume the data f belongs to a Lebesgue space with sufficiently high exponent. Precisely, for a given $\delta_0 > 0$

$$f \in L_{loc}^{p_{*s}+\delta_0}(\mathbb{R}^n)\,,$$

where we are using standard notation for Hölder and Sobolev exponents; that is, for any $r \in (1, \infty)$ and any $\sigma \in (0, 1)$ we write

$$r' = \frac{r}{r-1}\,, \qquad r^* = r^{*\sigma} = \frac{nr}{n - \sigma r}\,, \qquad r_* = r_{*\sigma} = \frac{nr'}{n + \sigma r'} = (r^*)'\,.$$

(The dependence of the embedding exponents on σ will be suppressed whenever it is clear from context.)

Weak solutions u are assumed to be *a priori* bounded, a point clarified by the following definition:

Definition 1.1. A function $u \in W^{s,p}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ is a *bounded weak solution* to (1.1) with data f if the nonlocal double phase energy $\mathcal{E}(u, u) < \infty$ and if u satisfies (1.2).

If we denote the integrand of $\mathcal{E}(u, u)$ by $P(x, y, u)$ so that

$$\mathcal{E}(u, u) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} P(x, y, u)\,\mathrm{d}y\,\mathrm{d}x\,,$$

then by definition of u as a bounded weak solution the function $P(\cdot, \cdot, u)$ belongs to $L^1(\mathbb{R}^{2n})$. The following theorem concerning P constitutes the main result of [6]:

Theorem 1.2. *Let p, $q \geq 2$ and s, $t \in (0, 1)$ satisfy (A2)-(A3) and let $a(x, y)$ satisfy (A1). Fix $\delta_0 > 0$, and let $f \in L_{loc}^{p_{*s}+\delta_0}(\mathbb{R}^n)$. Let $u \in W^{s,p}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ be any bounded weak solution to (1.1) with data f. Then there exists $\varepsilon_0 \in (0, 1)$ depending only on n, p, q, s, t, M, δ_0 and $\|u\|_{L^\infty(\mathbb{R}^n)}$ such that for every $\tau \in (0, \varepsilon_0)$*

$$P(\cdot, \cdot, u) \in L_{loc}^{1+\tau}(\mathbb{R}^{2n})\,.$$

In particular, there exist positive constants ε_1 and ε_2 such that $u \in W_{loc}^{s+\varepsilon_1, p+\varepsilon_2}(\mathbb{R}^n)$, and if $(s + \varepsilon_1)(p + \varepsilon_2) > n$ then u is locally Hölder continuous.

To prove Theorem 1.2 we use an argument developed by Kuusi, Mingione and Sire announced in [3] and presented in [4] that builds a nonlocal fractional Gehring lemma in order to prove a self-improvement result for solutions to a class of monotone operators with quadratic growth related to the fractional Laplacian. The arguments in [6] and in this note are heavily based on the work and presentation done for the case $p = 2$ in [4]. While it is apparent from a careful reading of that work that their methods apply to functionals with more general p-growth, the precise treatment of such classes of operators does not appear in the literature. Since we are further working with operators of mixed (p, q) growth, in [6] and in this note we have written the arguments of [4] for a

general exponent p so that the robustness of their technique and as well as results can be clearly seen as applicable in a wealth of contexts. One such instances is an extension of these arguments to vector-valued solutions of nonlocal systems. A specific example is the strongly coupled system of nonlinear equations studied in [5]. We also would like to mention a sequence of recent papers in [9, 10] where the same argument is applied to obtain higher Sobolev regularity estimates for nonlocal equations with VMO coefficients corresponding to the case $p = 2$.

The fractional Gehring lemma relies on a level set estimate of a quantity related to the solution, and its proof is the main contribution of this note. If a reader of [6] accepts the level set estimate as true then the rest of the proof in [6] of the Gehring lemma follows in a straightforward way. However, because the proof of the estimate itself is quite technical and very closely resembles the argument from [4], we have written it here instead of in [6].

Due to its technical nature we do not write a statement of the level set estimate precisely until the beginning of Section 4. To ensure there is no interruption in the thread of reasoning between this work and [6], the prerequisite results are given in Section 3 as a summarized version of their counterparts in [6].

Following the structure introduced in [4], we define *dual pairs* of measures and functions (U, ν). For small $\varepsilon \in (0, 1/p)$ we define the locally finite doubling Borel measure in $\mathbb{R}^{2n}$

$$(1.4) \qquad \nu(A) := \int_A \frac{\mathrm{d}x \, \mathrm{d}y}{|x - y|^{n - \varepsilon p}}, \qquad A \subset \mathbb{R}^{2n} \text{ measurable},$$

and we define the function

$$(1.5) \qquad U(x, y) := \frac{|u(x) - u(y)|}{|x - y|^{s + \varepsilon}}.$$

It is then clear that

$$u \in W^{s,p}(\mathbb{R}^n) \qquad \text{if and only if} \qquad U \in L^p(\mathbb{R}^{2n}; \nu).$$

The integrand $P(x, y, u)$ of the energy $\mathcal{E}(u, u)$ can be expressed in terms of U as

$$(1.6) \qquad [U^p + A(x, y)U^q]|x - y|^{-n + \epsilon p}, \quad \text{where } A(x, y) := a(x, y)|x - y|^{(s-t)q + \varepsilon(q - p)}.$$

We can therefore write the double phase energy $\mathcal{E}(u.u)$ in terms the dual pair as

$$(1.7) \qquad \mathcal{E}(u, u) = \int_{\mathbb{R}^{2n}} (U^p + A(x, y)U^q) \, \mathrm{d}\nu =: \int_{\mathbb{R}^{2n}} G(x, y, U) \, \mathrm{d}\nu$$

where the integrand $G(x, y, U) := U^p + A(x, y)U^q$. Then it now becomes clear that

$$P(\cdot, \cdot, u) \in L^1(\mathbb{R}^{2n}) \quad \text{if and only if} \quad G(\cdot, \cdot, U) \in L^1(\mathbb{R}^{2n}; \nu).$$

Theorem 1.3 (Higher Regularity Result)**.** *With all the assumptions of Theorem 1.2, there exists $\varepsilon_0 > 0$ depending only on* **data** *such that for every $\delta \in (0, \varepsilon_0)$ we have*

$$(1.8) \qquad G(x, y, U) \in L_{loc}^{1 + \delta}(\mathbb{R}^{2d}; \nu).$$

where **data** *represents $n, p, q, s, t, M,$ and $\|u\|_{L^\infty}$.*

Theorem 1.2 is a simple consequence of the above theorem. In [6] we show (1.8) directly by way of a fractional Gehring lemma applied to the dual pair of function and measure (G, ν). This fractional Gehring lemma in turn relies on a kind of fractional reverse Hölder inequality; for the exact statement see Theorem 3.3 below. This inequality holds only for diagonal sets of the type $B \times B \subset \mathbb{R}^{2n}$, and it is insufficient to apply tools traditionally used to prove Gehring's lemma such as the maximal function. Nevertheless, Kuusi, Mingione, and Sire in [4] used a novel localization technique to show that the fractional reverse Hölder-type inequality over diagonal balls is sufficient

to prove a special fractional version of Gehring's lemma that is applicable for dual pairs of the above type. A key ingredient of this localization technique is a level set estimate

$$\frac{1}{\lambda^2} \int_{\mathcal{B}(x_0,\beta)\cap\{U>\lambda\}} U^2 \, \mathrm{d}\nu \precsim \frac{1}{\lambda^r} \int_{\mathcal{B}(x_0,\alpha)\cap\{U>\lambda\}} U^r \, \mathrm{d}\nu + \text{ terms involving level sets of } f, g$$

for some fixed $r < 2$ and for any $\lambda \geq \lambda_0$, where λ_0 is a finite constant depending on the solution. Here $\mathcal{B} = B \times B$, and $B \subset \mathbb{R}^n$ is a ball. We adapt the statement and proof of this level set estimate to our setting; see Proposition 4.1. Key steps of the proof are described in Section 4, and we additionally refer to the original discussions and summaries of the technique in [3, 4].

We finally remark that the fractional Gehring lemma incorporates the level set estimate, and the proof of the fractional Gehring lemma itself is contained in [6].

This note is organized as follows: In the next section we identify notation and conventions. The reverse Hölder inequality is proved in [6], and for reference is stated in Section 3. Section 4 contains the proof of the level set estimate in its entirety.

2. Preliminaries

Throughout, we denote positive constants by c, C, etc., and they may change from line to line. We list the dependencies in parentheses after the constant when we wish to make them explicit, i.e. if a constant C depends only on n, p and s, we write $C = C(n,p,s)$. We will abbreviate the following set of parameters as

$$\texttt{data} \equiv (n, p, q, s, t, M, \|u\|_{L^\infty}) \, .$$

In $\mathbb{R}^n$, denote the open ball of radius R centered at x_0 by

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

We will sometimes denote the ball $B \equiv B_R \equiv B_R(x_0)$ whenever the center and/or radius is clear from context. If B is a ball centered at x_0 with radius R, then σB is the ball centered at x_0 with radius σR. Given any measure μ, denote the average of a μ-measurable function h over a set $\mathcal{A}$ by

$$(h)_{\mathcal{A}} := \fint_{\mathcal{A}} h \, \mathrm{d}\mu = \frac{1}{\mu(\mathcal{A})} \int_{\mathcal{A}} h(x) \, \mathrm{d}\mu \, .$$

In dealing with functions defined on $\mathbb{R}^{2n}$ such as U, we consider the norm on $\mathbb{R}^{2n}$ defined by

$$\|(x,y)\| := \max\{|x|, |y|\} \, ,$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. Denote the balls defined by this norm as

$$\mathcal{B}(x_0, y_0, R) := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \, : \, \|(x,y) - (x_0,y_0)\| < R\}$$
$$= B(x_0, R) \times B(y_0, R) \, .$$

If we denote

$$B_{\mathbb{R}^{2n}}(x_0, y_0, R) := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \, : \, \sqrt{|x - x_0|^2 + |y - y_0|^2} < R\} \, ,$$

then clearly

$$B_{\mathbb{R}^{2n}}(x_0, y_0, R) \subset \mathcal{B}(x_0, y_0, R) \subset B_{\mathbb{R}^{2n}}(x_0, y_0, 2R) \, .$$

Often we will need to consider balls in $\mathbb{R}^{2n}$ centered at a point on the "diagonal," that is, a point of the form (x_0, x_0) for $x_0 \in \mathbb{R}^n$. In this case we abbreviate $\mathcal{B}(x_0, x_0, R) \equiv \mathcal{B}(x_0, R)$. We will also use the abbreviations $\mathcal{B}(x_0, R) \equiv \mathcal{B}_R(x_0) \equiv \mathcal{B}_R \equiv \mathcal{B}$ whenever the center and/or radius is clear from context. Whenever there is no ambiguity we write $\mathcal{B}(x_0, \sigma R) = \sigma \mathcal{B}$. We also denote

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

We will use the elementary inequality

$$(2.1) \qquad 2^{kr} \sum_{j=k-1}^{\infty} 2^{-jr} \leq \frac{4^r}{r \ln(2)} \,, \qquad \text{for } k \geq 1 \quad \text{and} \quad r \in (0, \infty) \,.$$

The cardinality of a finite set $\mathcal{A}$ is denoted by $\#\mathcal{A}$. The set of nonnegative integers $\{0, 1, 2, \ldots\}$ is designated by $\mathbb{Z}_+$.

For any domain $\Omega \subset \mathbb{R}^n$, $0 < \sigma < 1$ and $r \in [1, \infty)$ the fractional Sobolev spaces are defined by the Gagliardo seminorm

$$W^{\sigma,r}(\Omega) := \left\{ u \in L^r(\Omega) \, : \, [u]_{W^{\sigma,r}(\Omega)} := \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^r}{|x - y|^{n+\sigma r}} \, \mathrm{d}y \, \mathrm{d}x < \infty \right\}$$

with norm $\|\cdot\|_{W^{\sigma,r}(\Omega)}^r := \|\cdot\|_{L^r(\Omega)}^r + [\cdot]_{W^{\sigma,r}(\Omega)}^r$.

We will also use the following fractional Poincaré-Sobolev-type inequalities throughout the note. A proof of the first can be found in several places; see for instance [1, 8]. The second can be found in [7, 11].

Theorem 2.1 (Fractional Poincaré-Sobolev Inequality). *Let $r \in [1, \infty)$, $0 < \sigma < 1$. Let $B = B_R(x_0)$ for some $R > 0$, $x_0 \in \mathbb{R}^n$. Then there exists $C = C(n, r, \sigma) > 0$ such that*

$$\left(\fint_B \left| \frac{v(x) - (v)_B}{R^\sigma} \right|^{r^*\sigma} \mathrm{d}x \right)^{1/r^*\sigma} \leq C \left(\int_B \fint_B \frac{|v(x) - v(y)|^r}{|x - y|^{n+\sigma r}} \, \mathrm{d}y \, \mathrm{d}x \right)^{1/r}$$

for every $v \in W^{\sigma,r}(B)$.

Theorem 2.2 (Fractional Poincaré Inequality). *Let $r \in [1, \infty)$, $0 < \sigma < 1$. Let $B = B_R(x_0)$ for some $R > 0$, $x_0 \in \mathbb{R}^n$. Then there exists $C = C(n, r) > 0$ such that*

$$\left(\fint_B \left| \frac{v(x) - (v)_B}{R^\sigma} \right|^r \mathrm{d}x \right)^{1/r} \leq C \left(\int_B \fint_B \frac{|v(x) - v(y)|^r}{|x - y|^{n+\sigma r}} \, \mathrm{d}y \, \mathrm{d}x \right)^{1/r}$$

for every $v \in W^{\sigma,r}(B)$.

3. Sobolev Inequality for Dual Pairs and Reverse Hölder Inequality

3.1. **The Dual Pair Measure.** We summarize some basic properties of the measure ν defined in (1.4). These properties are natural extensions of those established in [4, Proposition 4.1]; their proof is sketched in [6].

Theorem 3.1. *For any $\varepsilon \in (0, 1/p)$, the measure ν defined as*

$$\nu(\mathcal{A}) := \int_{\mathcal{A}} \frac{1}{|x - y|^{n-\varepsilon p}} \, \mathrm{d}y \, \mathrm{d}x \,, \qquad \mathcal{A} \subset \mathbb{R}^{2n} \,,$$

is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^{2n}$. Additionally,

- *For $\mathcal{B} = B_R(x_0) \times B_R(x_0)$,*

$$(3.1) \qquad \nu(\mathcal{B}) = \frac{c(n, p, \varepsilon) R^{n+\varepsilon p}}{\varepsilon} \,,$$

  where $c(n, p, \varepsilon)$ is a constant depending only on n, p and ε that satisfies $1/\widetilde{c}(n, p) \leq c(n, p, \varepsilon) \leq \widetilde{c}(n, p)$, where $\widetilde{c}$ is another constant depending only on n and p.

- *For every $x \in \mathbb{R}^n$ and for $R \geq r > 0$,*

$$(3.2) \qquad \frac{\nu(\mathcal{B}(x,R))}{\nu(\mathcal{B}(x,r))} = \left(\frac{R}{r}\right)^{n+\varepsilon p}.$$

- *For every $a \leq 1$, $R > 0$ and $x \in \mathbb{R}^n$, there exists a constant $C_d = C_d(n,p)$ such that*

$$(3.3) \qquad \frac{\nu(\mathcal{B}(x,R))}{\nu(K_1 \times K_2)} \leq \frac{C_d}{a^{2n}\varepsilon}$$

for any two cubes K_1, $K_2 \subset B_R(x)$ with sides parallel to the coordinate axes and such that $|K_1| = |K_2| = (aR)^n$.

3.2. **Reverse Hölder Inequality.** Recall that

$$(3.4) \qquad U(x,y) = \frac{|u(x) - u(y)|}{|x-y|^{s+\varepsilon}}, \qquad \text{and define } F(x,y) := |f(x)|.$$

Then $F \in L_{loc}^{p_*+\delta}(\mathbb{R}^{2n})$ for every $\delta \in (0, \delta_0)$, as a direct calculation using the properties of measure ν.

We now report the compatibility of the Sobolev-Poincaré inequality with the definition of U. Given $B = B_R(x_0)$, define $\tau \in (0,1)$, and $\eta \in (1, \infty)$ to be differentiability and integrability constants respectively that have yet to be fixed. Letting $\varepsilon \in (0, \min\{\frac{s}{p}, 1-s\})$ and using (3.1),

$$\fint_B \int_B \frac{|u(x) - u(y)|^\eta}{|x-y|^{n+\tau\eta}} \, \mathrm{d}y \, \mathrm{d}x = \frac{CR^{\varepsilon p}}{\varepsilon} \fint_{\mathcal{B}} U^\eta \, \mathrm{d}\nu$$

so long as

$$\tau + \frac{\varepsilon p}{\eta} = s + \varepsilon.$$

Since $\varepsilon \in (0, \frac{s}{p})$ and $\varepsilon < 1-s$ the exponent τ remains in $(0,1)$ for every $\eta \in (1, \infty)$. With this choice of τ, by the fractional Sobolev inequality, Theorem 2.1,

$$\left(\fint_B \left|\frac{u(x) - (u)_B}{R^\tau}\right|^m \, \mathrm{d}x\right)^{1/m} \leq C \left(\fint_B \int_B \frac{|u(x) - u(y)|^\eta}{|x-y|^{n+\tau\eta}} \, \mathrm{d}y \, \mathrm{d}x\right)^{1/\eta}$$

for every $m \in [1, \eta^{*\tau}]$ with $\eta \in (1, \infty)$. We choose η to satisfy the relation

$$(3.5) \qquad p = \eta^{*\tau} = \frac{n\eta}{n - \tau\eta} = \frac{n\eta}{n - \eta(s + \varepsilon - \frac{\varepsilon p}{\eta})} \quad \Longleftrightarrow \quad \eta = \frac{np + \varepsilon p^2}{n + sp + \varepsilon p}.$$

This choice of η is a valid Lebesgue exponent; note that $\eta < p$ for all $n \geq 1$ and for all $p \in (1, \infty)$, and that $\eta > 1$ so long as $p \geq 2$. Taking $m = \eta^{*\tau}$ we summarize this discussion in the following lemma:

Lemma 3.2. *Let $\varepsilon \in (0, s/p)$ with $\varepsilon < 1-s$ and $p \geq 2$. Define $\eta = \frac{np+\varepsilon p^2}{n+sp+\varepsilon p}$. Then*

$$\left(\fint_B |u(x) - (u)_B|^p \, \mathrm{d}x\right)^{1/p} \leq \frac{CR^{s+\varepsilon}}{\varepsilon^{1/\eta}} \left(\fint_{\mathcal{B}} U^\eta \, \mathrm{d}\nu\right)^{1/\eta},$$

where $C = C(n, s, p)$. The same inequality holds when the ball B is replaced by a cube Q with sides of length R and with $\mathcal{B}$ replaced by $Q \times Q$.

Recall that $G(x, y, U) = U^p + A(x,y)U^q$. We have the following L^1_{loc} estimate for G which will lead us to a scale-invariant reverse Hölder's inequality. The statement is precisely [6, Proposition 4.3] and its proof can be found in the same paper.

Proposition 3.3. *Let $p \in [2, \infty)$, and let $\varepsilon < 1 - s$ with $\varepsilon \in (0, \min\{s(\frac{tq}{sp} - \frac{1}{p'}), \frac{s}{p}\})$. (This choice is possible by Assumption A2). Let η be given by the formula in (3.5), Let $B = B_R(x_0)$ be a ball with $R \leq 1$. Then there exists a constant C depending only on **data** such that for any solution $u \in W^{s,p}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ to (1.1) and for any $\sigma \in (0, 1)$*

$$
(3.6) \quad
\begin{aligned}
\left(\fint_{\frac{1}{4}\mathcal{B}} G(x, y, U) \, d\nu \right)^{1/p} &\leq \frac{C}{\varepsilon^{1/\eta - 1/p}} \left[\frac{1}{\sigma} \left(\fint_{\mathcal{B}} U^\eta \, d\nu \right)^{1/\eta} \right. \\
&\quad + \sigma \sum_{k=0}^\infty \left(2^{-k(\frac{sp}{p-1} - s - \varepsilon)} + 2^{-k(\frac{tq}{p-1} - s - \varepsilon)} \right) \left. \left(\fint_{2^k \mathcal{B}} U^\eta \, d\nu \right)^{1/\eta} \right] \\
&\quad + \frac{C[\varepsilon \nu(\mathcal{B})]^{\frac{\theta}{p-1}}}{\varepsilon^{(1/p_* - 1/p')\frac{1}{p-1}}} \left[\left(\fint_{\mathcal{B}} F^{p_*} \, d\nu \right)^{1/p_*} \right]^{1/(p-1)},
\end{aligned}
$$

where

$$
\theta := \frac{s - \varepsilon(p-1)}{n + \varepsilon p} > 0.
$$

Remark 3.4. We make some remarks. The upper bound in (3.6) can be simplified down to just one series. Since $sp \geq tq$

$$
(3.7) \qquad 2^{-k(\frac{sp}{p-1} - s - \varepsilon)} \leq 2^{-k(\frac{tq}{p-1} - s - \varepsilon)}, \qquad k \in \mathbb{Z}_+,
$$

so we can replace the infinite series on the right-hand side of (3.6) with

$$
\frac{C\sigma}{\varepsilon^{1/\eta - 1/p}} \sum_{k=0}^\infty \alpha_k \left(\fint_{2^k \mathcal{B}} U^\eta \, d\nu \right)^{1/\eta},
$$

where

$$
(3.8) \qquad \alpha_k := 2^{-k(\frac{tq}{p-1} - s - \varepsilon)}.
$$

Moreover, in the case $a \equiv 0$ one simply takes $\alpha_k = 2^{-k(\frac{s}{p-1} - \varepsilon)}$. In any case, since $\varepsilon \leq \min\{s(\frac{tq}{sp} - \frac{1}{p'}), \frac{s}{p}\}$ the series $\sum_{k=0}^\infty \alpha_k < \infty$ and as a consequence

$$
\begin{aligned}
\sum_{k=0}^\infty \alpha_k \left(\fint_{2^k \mathcal{B}} U^\eta \, d\nu \right)^{1/\eta} &\leq \sum_{k=0}^\infty \alpha_k \left(\fint_{2^k \mathcal{B}} U^p \, d\nu \right)^{1/p} \\
&= C(\varepsilon, p, s) \sum_{k=0}^\infty \alpha_k \left(\int_{2^k B} \int_{2^k B} \frac{|u(y) - u(x)|^p}{|x - y|^{n+sp}} \, dx dy \right)^{1/p} \\
&\leq R^{-n/p - \epsilon} C(\varepsilon, p, s) \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(y) - u(x)|^p}{|x - y|^{n+sp}} \, dx dy \right)^{1/p} < \infty.
\end{aligned}
$$

The following corollary establishes a genuine scale-invariant reverse Hölder inequality for an appropriately scaled version of the integrand G. This quantity will satisfy a self-improving result.

Corollary 3.4.1. *Let $\varepsilon \in \left(0, \min\{s(\frac{tq}{sp} - \frac{1}{p'}), \frac{s}{p}\} \right)$. (This choice is possible by Assumption A2). Let $B = B_R(x_0)$ be a ball with $R \leq 1$. Define $H(x, y, U) := G(x, y, U)^{(p-1)/p}$. Then there exists a constant C depending only on **data** such that for any solution $u \in W^{s,p}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ to (1.1)*

and for any $\sigma \in (0,1)$

$$\left(\fint_{\frac{1}{4}\mathcal{B}} H(x,y,U)^{p'} \, \mathrm{d}\nu\right)^{1/p'} \leq \frac{C}{\sigma \varepsilon^{1/\gamma - 1/p'}} \left(\fint_{\mathcal{B}} H(x,y,U)^{\gamma} \, \mathrm{d}\nu\right)^{1/\gamma}$$

(3.9)
$$+ \frac{C\sigma}{\varepsilon^{1/\gamma - 1/p'}} \sum_{k=0}^{\infty} \alpha_k \left(\fint_{2^k\mathcal{B}} H(x,y,U)^{\gamma} \, \mathrm{d}\nu\right)^{1/\gamma}$$

$$+ \frac{C[\nu(\mathcal{B})]^{\theta}}{\varepsilon^{1/p_* - 1/p'}} \left(\fint_{\mathcal{B}} F^{p_*} \, \mathrm{d}\nu\right)^{1/p_*},$$

where $\gamma := \frac{\eta}{p-1} = p' \cdot \frac{n+\varepsilon p}{n+sp+\varepsilon p} < p'$ and $\theta := \frac{s-\varepsilon(p-1)}{n+\varepsilon p}$.

Remark 3.5. If $a \equiv 0$ one can see from careful inspection of the proofs they need not assume $u \in L^{\infty}(\mathbb{R}^n)$ in Proposition 3.3 and Corollary 3.4.1.

4. Proof of the Level Set Estimate

We are now ready to state and prove the level set estimate. It is stated precisely in Proposition 4.1 below, and this section is devoted to its proof. We first define the following: using the notation of Corollary 3.4.1, for any $x_0 \in \mathbb{R}^n$ and $R > 0$ set

(4.1) $$\Theta(x_0, R) := \Upsilon_0(x_0, R) + Tail(x_0, R) + \Psi_1(x_0, R),$$

where

$$\Upsilon_0(x_0, R) := \left(\fint_{\mathcal{B}(x_0,R)} F^{p_* + \delta_f} \, \mathrm{d}\nu\right)^{1/(p_* + \delta_f)}, \quad \text{with } \delta_f \in (0, \delta_0) \text{ to be determined,}$$

(4.2)
$$Tail(x_0, R) := \sum_{k=0}^{\infty} 2^{-k(\frac{tq}{p-1} - s - \varepsilon)} \left(\fint_{\mathcal{B}(x_0, 2^k R)} H^{\gamma} \, \mathrm{d}\nu\right)^{1/\gamma},$$

and, for any constant $M \geq 1$,

(4.3) $$\Psi_M(x_0, R) := \left(\fint_{\mathcal{B}(x_0,R)} H^{p'} \, \mathrm{d}\nu\right)^{1/p'} + M\frac{[\nu(\mathcal{B}(x_0,R))]^{\theta}}{\varepsilon^{1/p_* - 1/p'}} \left(\fint_{\mathcal{B}(x_0,R)} F^{p_*} \, \mathrm{d}\nu\right)^{1/p_*};$$

we write Ψ_M with $M = 1$ as Ψ_1.

Proposition 4.1. *Assume* (A1), (A2), *and* (A3). *Assume that $\varepsilon > 0$ satisfies*

(A4) $$\varepsilon \in (0, s/p), \qquad \varepsilon < s\left(\frac{tq}{sp} - \frac{1}{p'}\right), \qquad \varepsilon < 1 - s.$$

Let $u \in W^{s,p}(\mathbb{R}^n) \cap L^{\infty}(\mathbb{R}^n)$ be a bounded weak solution to (1.1), *and let $f \in L^{p_* + \delta_0}(\mathbb{R}^n)$ for given $\delta_0 > 0$. Let U and F be as in* (3.4). *Let $\mathcal{B}(x_0, \varrho_0) \subset \mathbb{R}^{2n}$ with $0 < \varrho_0 \leq 1$, and let α and β be such that $\varrho_0 < \beta < \alpha < \frac{3}{2}\varrho_0$ so that we have*

$$\mathcal{B}(x_0, \varrho_0) \subset \mathcal{B}(x_0, \beta) \subset \mathcal{B}(x_0, \alpha) \subset \mathcal{B}\left(x_0, \frac{3}{2}\varrho_0\right).$$

Then there exist constants $C_{\alpha} = C_{\alpha}(\mathtt{data}) > 0$, $C_f = C_f(\mathtt{data}, \varepsilon) \geq 1$ and $\kappa_f = \kappa_f(\mathtt{data}, \varepsilon) \in (0, 1)$, with positive constants

(4.4) $$\vartheta := \frac{3(p' - \gamma)}{\gamma}, \qquad \vartheta_f := (p_* + \delta_f)\left(\frac{p_*\theta}{1 - p_*\theta}\right), \qquad \widetilde{\vartheta}_f := \frac{p_*(1 + \theta\delta_f)}{1 - p_*\theta},$$

such that

$$(4.5) \quad \frac{1}{\lambda^{p'}} \int_{\mathcal{B}(x_0,\beta) \cap \{H>\lambda\}} H^{p'} \, d\nu \leq \frac{C_\alpha}{\varepsilon^\vartheta \lambda^\gamma} \int_{\mathcal{B}(x_0,\alpha) \cap \{H>\lambda\}} H^\gamma \, d\nu + \frac{C_f \lambda_0^{\vartheta_f}}{\lambda^{\widetilde{\vartheta}_f}} \int_{\mathcal{B}(x_0,\alpha) \cap \{F>\kappa_f \lambda\}} F^{p_*} \, d\nu$$

for every $\lambda \geq \lambda_0$, where λ_0 is defined as

$$(4.6) \qquad \qquad \lambda_0 := \frac{C_a}{\varepsilon} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{2n+p} \Theta(x_0, 2\varrho_0),$$

and where $C_a = C_a(\textbf{\textit{data}})$ and Θ has been defined in (4.1). (see also (4.2) and (4.3)).

For $p = 2$, this proposition is proved in Section 5 of [4]. Our proof is essentially the same as the proof found in [4]. However, we are writing the proof to make sure that the choice of the other parameters in (4.4) are correctly made and to emphasize the robustness of the arguments in [4] and how they can be used for more general nonlinear operators. As it has been explained in [4], the main difficulty in proving (4.5) is that the reverse Hölder inequality (3.9) only holds on diagonal balls of the type $\mathcal{B}(x_0, x_0, R)$. Thus maximal function arguments cannot be used, and we must resort to more direct arguments. We use a Calderón-Zygmund decomposition to decompose the level set $\{H > \lambda\}$ into dyadic cubes. These cubes are then sorted into cubes situated on or near the diagonal (called "diagonal" cubes) and cubes far from the diagonal (called "off-diagonal" cubes). What is meant by "far from" will be quantified below. The level set estimate for the diagonal cubes are handled using the reverse Hölder inequality (3.9). It turns out that Sobolev functions automatically satisfy a type of reverse Hölder inequality on off-diagonal cubes, and we use this to obtain the level set estimate for said cubes.

4.1. **Vitali Covering.**

Just as in [4], we begin with an exit-time argument. The goal is to cover the portion of the set $\{H > \lambda\}$ that lies on or near the diagonal $\{(x, x) : x \in \mathbb{R}^n\}$. Let $\kappa \in (0, 1]$ be a constant that will be chosen later, in (4.64); all arguments in the paper up until then are independent of the choice of κ. Define

$$(4.7) \qquad \qquad \lambda_1 := \frac{1}{\kappa} \sup_{\frac{\alpha - \beta}{40^n} \leq R \leq \frac{\varrho_0}{2}} \sup_{x \in B(x_0, \beta)} \{\Psi_M(x, R) + \Upsilon_0(x, R) + Tail(x, R)\}.$$

For the same κ and for $\lambda \geq \lambda_1$, define the "diagonal" level set of the functional Ψ_M by

$$(4.8) \qquad \qquad D_{\kappa\lambda} := \left\{ (x, x) \in \mathcal{B}(x_0, \beta) : \sup_{0 < R < \frac{\alpha - \beta}{40^n}} \Psi_M(x, R) > \kappa\lambda \right\}.$$

Then by definition of λ_1 we have

$$(4.9) \qquad \Psi_M(x, R) \leq \kappa\lambda_1 \leq \kappa\lambda, \qquad \text{for each } (x, x) \in \mathcal{B}(x_0, \beta), \quad R \in \left[\frac{\alpha - \beta}{40^n}, \frac{\varrho_0}{2} \right],$$

and so it follows that for every (x, x) in the diagonal level set $D_{\kappa\lambda}$ there exists an exit time $R(x) \in (0, \frac{\alpha - \beta}{40^n})$ such that

$$(4.10) \qquad \Psi_M(x, R(x)) \geq \kappa\lambda \qquad \text{while at the same time} \quad \sup_{R(x) < R < \frac{\alpha - \beta}{40^n}} \Psi_M(x, R) \leq \kappa\lambda.$$

Thus the collection $\{\mathcal{B}(x, 2R(x))\}$ forms a cover of $D_{\kappa\lambda}$, so by the Vitali covering theorem we can find a countable subcollection $\{\mathcal{B}(x_j, 2R(x_j))\}$ such that

$$(4.11) \qquad \bigcup_{(x,x) \in D_{\kappa\lambda}} \mathcal{B}(x, 2R(x)) \subset \bigcup_j \mathcal{B}(x_j, 10R(x_j)), \qquad \mathcal{B}(x_j, 2R(x_j)) \text{ are mutually disjoint}.$$

We hereafter use the abbreviations

$$(4.12) \qquad \mathcal{B}_j := \mathcal{B}(x_j, R(x_j)), \qquad \sigma\mathcal{B}_j := \mathcal{B}(x_j, \sigma R(x_j)), \quad \sigma > 0.$$

The quantity $\sum \nu(\mathcal{B}_j)$ is treated by the diagonal estimates in the next section. Note that since $x_j \in \mathcal{B}(x_0, \beta)$ and $R(x_j) \leq \frac{\alpha-\beta}{40^n}$ we have $10\mathcal{B}_j \subset \mathcal{B}(x_0, \alpha)$ for every j. By (4.10) and by the doubling property of the measure ν in (3.2) we also have

$$(4.13) \qquad \sum_j \int_{10\mathcal{B}_j} H^{p'} \, d\nu \leq \sum_j \nu(10\mathcal{B}_j)[\Psi_M(x_j, 10R(x_j))]^{p'} \leq 10^{n+\varepsilon p} \kappa^{p'} \lambda^{p'} \sum_j \nu(\mathcal{B}_j).$$

4.2. **Analysis On the Diagonal.** By (4.10) it follows that at least one of two inequalities hold: either

$$(4.14) \qquad \left(\fint_{\mathcal{B}_j} H^{p'} \, d\nu \right)^{1/p} \geq \frac{\kappa\lambda}{2}$$

or

$$(4.15) \qquad \frac{M[\nu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p_*-1/p'}} \left(\fint_{\mathcal{B}_j} F^{p_*} \, d\nu \right)^{1/p_*} \geq \frac{\kappa\lambda}{2}.$$

<u>Case 1:</u> If (4.14) occurs, then by the Reverse Hölder inequality (3.9)
$$(4.16)$$
$$\kappa\lambda \leq \frac{C}{\sigma\varepsilon^{1/\gamma-1/p'}} \left(\fint_{4\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} + \frac{\sigma}{\varepsilon^{1/\gamma-1/p'}} \sum_{k=0}^\infty \alpha_k \left(\fint_{2^{k+2}\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} + \frac{C[\nu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p_*-1/p'}} \left(\fint_{4\mathcal{B}_j} F^{p_*} \, d\nu \right)^{1/p_*},$$

where $\sigma \in (0,1]$ has yet to be chosen and $C = C(\mathtt{data})$. Choose the unique $m \in \mathbb{Z}_+$ such that $2^{-m}\varrho_0 \leq R(x_j) < 2^{-m+1}\varrho_0$. Since $R(x_j) < \frac{\alpha-\beta}{40^n}$ and $0 < \alpha - \beta < \varrho_0/2$, we have $m \geq 3$. Further, $\frac{\alpha-\beta}{40^n} \leq \frac{\varrho_0}{2\cdot 40^n} \leq 2^{m-1}R(x_j)$, so by (4.7)

$$(4.17) \qquad \Upsilon_0(x_j, 2^{m-1}R(x_j)) + Tail(x_j, 2^{m-1}R(x_j)) \leq \kappa\lambda_1.$$

This allows us to estimate $Tail$. The first $m-2$ terms can be handled by the exit-time condition (4.10); that is,

$$(4.18) \qquad \left(\fint_{2^k\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} \leq \kappa\lambda \quad \text{if } 1 \leq k \leq m-2.$$

Then by (4.17) and (4.18), and recalling that $\alpha_k = 2^{-k(\frac{tq}{p-1}-s-\varepsilon)}$,

(4.19)
$$
\begin{aligned}
\sum_{k=0}^{\infty} \alpha_k \left(\fint_{2^{k+2}\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} &= \frac{1}{\alpha_2} \sum_{k=2}^{\infty} \alpha_k \left(\fint_{2^k\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} \\
&= \frac{1}{\alpha_2} \sum_{k=2}^{m-2} \alpha_k \left(\fint_{2^k\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} + \frac{1}{\alpha_2} \sum_{k=0}^{\infty} \alpha_{k+m-1} \left(\fint_{2^{k+m-1}\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} \\
&\leq \frac{1}{\alpha_2} \left[\kappa\lambda \sum_{k=2}^{m-2} \alpha_k + \alpha_{m-1} Tail(x_j, 2^{m-1}R(x_j)) \right] \\
&\leq \frac{1}{\alpha_2} \left[\kappa\lambda \sum_{k=2}^{m-2} \alpha_k + \alpha_{m-1}\kappa\lambda_1 \right] \\
&\leq \kappa\lambda \sum_{k=0}^{\infty} \alpha_k \leq \frac{4^{\frac{tq}{p-1}-s-\varepsilon}\kappa\lambda}{(\frac{tq}{p-1}-s-\varepsilon)\ln(2)} \leq \frac{4^q\kappa\lambda}{s(p'-1)(\frac{tq}{sp}-\frac{1}{p'})\ln(2)} := C_1\kappa\lambda,
\end{aligned}
$$

where in the last line we used (2.1) and the bound $\varepsilon < s(\frac{tq}{sp}-\frac{1}{p'})$ in (A4). The constant C_1 depends only on data. Now using the fact that $m \geq 3$ we gain that $2R(x_j) \leq \frac{1}{2}\varrho_0$, so by the exit-time condition (4.10)

(4.20)
$$
\frac{C[\nu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p_*-1/p'}} \left(\fint_{4\mathcal{B}_j} F^{p_*} \, d\nu \right)^{1/p_*} \leq C\frac{\Psi_M(x_j, 2R(x_j))}{M} \leq \frac{C_2\kappa\lambda}{M}, \quad C_2 = C_2(\texttt{data}).
$$

Combining (4.19) and (4.20) in (4.16) gives

(4.21)
$$
\kappa\lambda \leq \frac{C}{\sigma\varepsilon^{1/\gamma-1/p'}} \left(\fint_{4\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} + \frac{C_1\sigma\kappa\lambda}{\varepsilon^{1/\gamma-1/p'}} + \frac{C_2\kappa\lambda}{M}.
$$

Now we choose $\sigma \in (0,1)$ and $M \geq 1$, and absorb the last two terms. We set

(4.22)
$$
\sigma := \frac{\varepsilon^{1/\gamma-1/p'}}{4C_1}, \qquad M := 4C_2,
$$

and so we have for $C = C(\texttt{data})$

(4.23)
$$
\kappa\lambda \leq \frac{C}{\varepsilon^{2/\gamma-2/p'}} \left(\fint_{4\mathcal{B}_j} H^\gamma \, d\nu \right)^{1/\gamma} \quad \Rightarrow \quad \nu(\mathcal{B}_j) \leq \frac{C}{\varepsilon^{2-(2\gamma)/p'}(\kappa\lambda)^\gamma} \int_{4\mathcal{B}_j} H^\gamma \, d\nu.
$$

Now, let $\widetilde{\kappa} > 0$ be a constant, to be fixed in a moment. Using the doubling property (3.2),

(4.24)
$$
\begin{aligned}
\frac{C}{\varepsilon^{2-(2\gamma)/p'}(\kappa\lambda)^\gamma} \int_{4\mathcal{B}_j} H^\gamma \, d\nu &\leq \frac{C}{\varepsilon^{2-(2\gamma)/p'}(\kappa\lambda)^\gamma} \int_{4\mathcal{B}_j \cap \{H \leq \widetilde{\kappa}\kappa\lambda\}} H^\gamma \, d\nu + \frac{C}{\varepsilon^{2-(2\gamma)/p'}(\kappa\lambda)^\gamma} \int_{4\mathcal{B}_j \cap \{H > \widetilde{\kappa}\kappa\lambda\}} H^\gamma \, d\nu \\
&\leq \frac{C_3\nu(\mathcal{B}_j)\widetilde{\kappa}^\gamma}{\varepsilon^{2-(2\gamma)/p'}} + \frac{C_3}{\varepsilon^{2-(2\gamma)/p'}(\kappa\lambda)^\gamma} \int_{4\mathcal{B}_j \cap \{H > \widetilde{\kappa}\kappa\lambda\}} H^\gamma \, d\nu,
\end{aligned}
$$

where $C_3 = C_3(\texttt{data})$. Choose

(4.25)
$$
\widetilde{\kappa} = \frac{\varepsilon^{2/\gamma-2/p'}}{(2C_3)^{1/\gamma}},
$$

and then substituting (4.24) into (4.23)$_2$ and absorbing the term gives

$$(4.26) \qquad \nu(\mathcal{B}_j) \leq \frac{C_4}{\varepsilon^{2-2\gamma/p'}\kappa^\gamma\lambda^\gamma} \int_{4\mathcal{B}_j \cap \{\widetilde{\kappa}\kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,, \qquad C_4 = C_4(\mathtt{data}) \,.$$

<u>Case 2:</u> If (4.15) occurs then

$$\left(\frac{\kappa\lambda}{2}\right)^{p_*} \leq \frac{M^{p_*}[\nu(\mathcal{B}_j)]^{p_*\theta-1}}{\varepsilon^{1-p_*/p'}} \int_{\mathcal{B}_j} F^{p_*} \, \mathrm{d}\nu \,,$$

which implies

$$(4.27) \qquad \nu(\mathcal{B}_j) \leq \left(\frac{2M}{\varepsilon^{1/p_*-1/p'}\kappa\lambda}\right)^{p_*/(1-p_*\theta)} \left(\int_{\mathcal{B}_j} F^{p_*} \, \mathrm{d}\nu\right)^{1/(1-p_*\theta)}$$

where we used the inequality $p_*\theta < 1$ which follows from the definition of θ that $p_*\theta \leq \frac{nsp}{(n+sp)(n+\varepsilon p)} < \frac{p}{p+1}$. Moreover, since $p \geq 2$, we have that $3 \leq \frac{1}{1-p_*\theta} = 1+\frac{p_*\theta}{1-p_*\theta}$. This decomposition of the exponent will allow us to incorporate the level set of F in the integrand, and then remove the exponent on the integral. To that end, with a constant $\widehat{\kappa} \in (0,1)$ to be determined, we split the integral as

$$(4.28) \qquad \left(\int_{\mathcal{B}_j} F^{p_*} \, \mathrm{d}\nu\right)^{\frac{1}{1-p_*\theta}} \leq \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \, \mathrm{d}\nu + (\widehat{\kappa}\kappa\lambda)^{p_*}\nu(\mathcal{B}_j)\right)^{\frac{1}{1-p_*\theta}} \,.$$

Now, since $\varrho_0 \leq 1$,

$$\nu(\mathcal{B}(x_0, 2\varrho_0)) \leq \frac{C(n)2^{n+\varepsilon p}}{\varepsilon} := L = L(n,p,\varepsilon) \,.$$

Then by noting that $\mathcal{B}_j \subset \mathcal{B}(x_0, 2\varrho_0)$ we can estimate

$$[\nu(\mathcal{B}_j)]^{\frac{1}{1-p_*\theta}} \leq [\nu(\mathcal{B}(x_0, 2\varrho_0))]^{\frac{p_*\theta}{1-p_*\theta}}\nu(\mathcal{B}_j) \leq L^{\frac{p_*\theta}{1-p_*\theta}}\nu(\mathcal{B}_j) \,.$$

Then using the elementary inequality $(a+b)^r \leq 2^{r-1}(a^r+b^r)$ for any $r \geq 1$ in (4.28) and using the estimate above for $[\nu(\mathcal{B}_j)]^{1/(1-p_*\theta)}$ we have

$$(4.29) \qquad \begin{aligned} \left(\int_{\mathcal{B}_j} F^{p_*} \, \mathrm{d}\nu\right)^{\frac{1}{1-p_*\theta}} &\leq 2^{p_*\theta/(1-p_*\theta)} \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \, \mathrm{d}\nu\right)^{\frac{1}{1-p_*\theta}} \\ &\quad + (2(L+1))^{\frac{p_*\theta}{1-p_*\theta}}(\widehat{\kappa}\kappa\lambda)^{\frac{p_*\theta}{1-p_*\theta}}\nu(\mathcal{B}_j) \,. \end{aligned}$$

Combining (4.27) and (4.29), and using that $\theta \leq 1$,

$$\nu(\mathcal{B}_j) \leq \left(\frac{2^2 M}{\varepsilon^{1/p_*-1/p'}\kappa\lambda}\right)^{p_*/(1-p_*\theta)} \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \, \mathrm{d}\nu\right)^{1/(1-p_*\theta)} + \left(\frac{2^2 M(L+1)\widehat{\kappa}}{\varepsilon^{1/p_*-1/p'}}\right)^{p_*/(1-\theta p_*)}\nu(\mathcal{B}_j) \,.$$

Now set $\widehat{\kappa} \in (0,1)$ to satisfy

$$(4.30) \qquad \left(\frac{4M(L+1)\widehat{\kappa}}{\varepsilon^{1/p_*-1/p'}}\right)^{p_*/(1-p_*\theta)} \leq \frac{1}{2} \quad \Rightarrow \quad \widehat{\kappa} \leq \left(\frac{1}{2}\right)^{(1-p_*\theta)/p_*} \frac{\varepsilon^{1/p_*-1/p'}}{4M(L+1)} \,.$$

Then

$$\nu(\mathcal{B}_j) \leq 2\left(\frac{4M}{\varepsilon^{1/p_*-1/p'}\kappa\lambda}\right)^{\frac{p_*}{1-p_*\theta}} \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \, \mathrm{d}\nu\right)^{\frac{1}{1-p_*\theta}} \,.$$

Writing $\left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu\right)^{\frac{1}{1-p_*\theta}} = \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu\right) \left(\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu\right)^{\frac{p_*\theta}{1-p_*\theta}}$ let us transfer some of the "decay" from the integrand to the cutoff λ_0, and in so doing remove the exponent from the integral.

$$\int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu \leq (\widehat{\kappa}\kappa\lambda)^{p_*} \int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} \left(\frac{F}{\widehat{\kappa}\kappa\lambda}\right)^{p_*+\delta_f} \,\mathrm{d}\nu$$

$$\leq \frac{\nu(2^{m-1}\mathcal{B}_j)}{(\widehat{\kappa}\kappa\lambda)^{\delta_f}} \fint_{2^{m-1}\mathcal{B}_j} F^{p_*+\delta_f} \,\mathrm{d}\nu$$

$$\leq \frac{\nu(\mathcal{B}(x_0, 2\varrho_0))}{(\widehat{\kappa}\kappa\lambda)^{\delta_f}} [\Upsilon_0(x_j, 2^{m-1}R(x_j))]^{p_*+\delta_f} \leq \frac{L\lambda_1^{p_*+\delta_f}}{(\widehat{\kappa}\kappa\lambda)^{\delta_f}} \,.$$

Therefore, powering the above inequality by $\frac{p_*\theta}{1-p_*\theta}$ on both sides we obtain that

$$(4.31) \qquad \nu(\mathcal{B}_j) \leq \frac{C_5 \lambda_1^{(p_*+\delta_f)\theta p_*/(1-p_*\theta)}}{(\widehat{\kappa}\kappa\lambda)^{(1+\theta\delta_f)p_*/(1-p_*\theta)}} \int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu \,,$$

where

$$(4.32) \qquad C_5 := 2 \left(\frac{4M(L+1)}{\varepsilon^{1/p_*-1/p'}}\right)^{p_*/(1-p_*\theta)} .$$

Combining (4.26) and (4.31),

$$\nu(\mathcal{B}_j) \leq \frac{C_4}{\varepsilon^{2-2\gamma/p'}\kappa^\gamma\lambda^\gamma} \int_{4\mathcal{B}_j \cap \{H > \widetilde{\kappa}\kappa\lambda\}} H^\gamma \,\mathrm{d}\nu + \frac{C_5\lambda_1^{\vartheta_f}}{(\widehat{\kappa}\kappa\lambda)^{\widetilde{\vartheta}_f}} \int_{\mathcal{B}_j \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu \,,$$

and since $\{\mathcal{B}_j\}$ is a disjoint collection whose members are all contained in $\mathcal{B}(x_0, \alpha)$ we have

$$(4.33) \qquad \sum_j \nu(\mathcal{B}_j) \leq \frac{C_4}{\varepsilon^{2-2\gamma/p'}\kappa^\gamma\lambda^\gamma} \int_{\mathcal{B}(x_0,\alpha) \cap \{H > \widetilde{\kappa}\kappa\lambda\}} H^\gamma \,\mathrm{d}\nu + \frac{C_5\lambda_1^{\vartheta_f}}{(\widehat{\kappa}\kappa\lambda)^{\widetilde{\vartheta}_f}} \int_{\mathcal{B}(x_0,\alpha) \cap \{F > \widehat{\kappa}\kappa\lambda\}} F^{p_*} \,\mathrm{d}\nu \,,$$

where the constants $C_4 = C_4(\texttt{data})$, $C_5(\texttt{data}, \varepsilon)$, $\widetilde{\kappa}$ and $\widehat{\kappa}$ have been chosen to satisfy (4.26), (4.32), (4.25) and (4.30) respectively. The constant $\kappa \in (0, 1]$ defined in (4.64) will be fixed in the course of the off-diagonal estimates below.

4.3. Analysis Off the Diagonal.

The analysis far from the diagonal is much more technical. We begin with defining the collections of dyadic cubes and summarizing their properties that will be used repeatedly. We then define two dimensional constants that will be used throughout the analysis of cubes far from the diagonal. In Section 4.3.3 we recall the classical Calderón-Zygmund decomposition and adapt it to decompose the level set $\{H > \lambda\}$ into the aforementioned dyadic cubes. In Section 4.3.4 we analyze the cubes near the diagonal and show they can be covered by the collection $\{\mathcal{B}_j\}$. In the remaining sections we treat the cubes far from the diagonal using the "almost-reverse" Hölder inequality in Lemma 4.6, a careful choice of the constant κ, and combinatorial information about the cubes coming from their size and distance from the diagonal.

4.3.1. *Dyadic Cubes.* The following contains information regarding cubes arising from a Calderón-Zygmund decomposition of Euclidean space. The dyadic cubes considered here are centered at x_0 and their size has been changed so as to be compatible with the starting ball $\mathcal{B}(x_0, \beta)$. We will consider cubes with side length 2^{-k} for integers $k \geq k_0$, where

$$(4.34) \qquad k_0 := \left\lfloor -\log_2\left(\frac{\alpha - \beta}{n40^{n+1}}\right)\right\rfloor + 1 \,.$$

Here, $\lfloor \cdot \rfloor$ denotes the floor function. For each $k \geq k_0$, let $\mathcal{C}_k$ be the disjoint collection of half-open cubes in $\mathbb{R}^n$ centered at x_0 with side length 2^{-k} whose closures intersect $\overline{B}(x_0, \frac{1}{2}(\alpha - \beta))$. That is,

$$\mathcal{C}_k := \{x_0 + 2^{-k}z + [0, 2^{-k})^n \; : \; z \in \mathbb{Z}^n, \; (x_0 + 2^{-k}z + [0, 2^{-k})^n) \cap \overline{B}(x_0, \frac{1}{2}(\alpha - \beta)) \neq \emptyset\}.$$

Then by (4.34) we have

$$(4.35) \qquad\qquad B(x_0, \beta) \subset \bigcup_{K \in \mathcal{C}_k} K \subset B(x_0, \alpha).$$

Note that every cube $K \in \mathcal{C}_{k+1}$ has a unique "predecessor" $\widetilde{K} \in \mathcal{C}_k$ such that $K \subset \widetilde{K}$. With these cubes, we can define dyadic cubes on $\mathbb{R}^{2n}$, denoted by

$$(4.36) \qquad \Delta_k := \{\mathcal{K} := K_1 \times K_2 \; : \; K_1, K_2 \in \mathcal{C}_k\}, \qquad \Delta := \bigcup_{k \geq k_0} \Delta_k.$$

Denote the diagonal cubes $\Delta_k^d := \{K \times K \; : \; K \in \Delta_k\}$. Then

$$\mathcal{B}(x_0, \beta) \subset \bigcup_{\mathcal{K} \in \Delta_k} \mathcal{K} \subset \mathcal{B}(x_0, \alpha).$$

Note that these product cubes in $\mathbb{R}^{2n}$ satisfy all of the same properties as the cubes. Numerous times a cube $\mathcal{K} \in \Delta$ will be given; there exists $K_1, K_2 \in \mathcal{C}_k$ such that $\mathcal{K} = K_1 \times K_2$, and we denote

$$k(\mathcal{K}) = k.$$

We denote the cube projections for a cube $\mathcal{K} = K_1 \times K_2$ by

$$\pi_1 \mathcal{K} := K_1 \times K_1, \qquad \text{and} \qquad \pi_2 \mathcal{K} := K_2 \times K_2.$$

Proposition 4.2. *Let $\mathcal{K} = K_1 \times K_2 \in \Delta$. The following hold:*

- *$\pi_1 \mathcal{K}, \pi_2 \mathcal{K} \in \Delta$.*
- *$\nu(\pi_1 \mathcal{K}) = \nu(\pi_2 \mathcal{K})$ and $k(\mathcal{K}) = k(\pi_1 \mathcal{K}) = k(\pi_2 \mathcal{K})$.*
- *If $\mathcal{H} \in \Delta$ and $\mathcal{H} \subset \mathcal{K}$, then $k(\mathcal{K}) \leq k(\mathcal{H})$.*
- *If $\widetilde{\mathcal{K}} = \widetilde{K}_1 \times \widetilde{K}_2$ is the predecessor of $\mathcal{K}$, then*

$$(4.37) \qquad\qquad dist(\widetilde{K}_1, \widetilde{K}_2) \leq dist(K_1, K_2).$$

- *The following hold:*

$$(4.38) \qquad\qquad \mathrm{dist}(K_1, K_2) = \frac{1}{\sqrt{2}} \, \mathrm{dist}(\pi_1 \mathcal{K}, \pi_2 \mathcal{K}).$$

$$(4.39) \qquad \mathrm{dist}(\mathcal{K}, \mathrm{diag}) = \frac{1}{\sqrt{2}} \, \mathrm{dist}(K_1, K_2) = \frac{1}{2} \, \mathrm{dist}(\pi_1 \mathcal{K}, \pi_2 \mathcal{K}).$$

> *Note that for two sets $\mathcal{A}_1$ and $\mathcal{A}_2 \subset \mathbb{R}^{2n}$ the definition of $\mathrm{dist}(\mathcal{A}_1, \mathcal{A}_2)$ uses the traditional Euclidean norm.*

4.3.2. *Two Dimensional Constants.*
In the analysis that follows we will often compare the sizes of the cubes with their distance to the diagonal. This leads us to define two dimensional constants independent of ε that will be used repeatedly. The existence of these constants follows from geometric arguments.

Lemma 4.3. *There exists a constant C_{dd} depending only on n such that for $h \in \{1,2\}$*
(4.40)
$$C_{dd} \geq \sup_{\mathcal{K} \in \Delta} \left\{ \frac{1}{\varepsilon} \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{n-\varepsilon p} \frac{\nu(\mathcal{K})}{\nu(\pi_h \mathcal{K})} \right\} + \sup_{\substack{\mathcal{K} \in \Delta \\ \operatorname{dist}(K_1, K_2) \geq 2^{-k}}} \left\{ \varepsilon \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{\varepsilon p - n} \frac{\nu(\pi_h \mathcal{K})}{\nu(\mathcal{K})} \right\},$$

where $k = k(\mathcal{K})$.

Proof. Without loss of generality assume $\operatorname{dist}(K_1, K_2) > 0$. By definition of ν
$$\nu(\pi_h \mathcal{K}) \geq \frac{C(n)(2^{-k})^{n+\varepsilon p}}{\varepsilon}, \qquad \text{while} \qquad \nu(\mathcal{K}) \leq \frac{2^{(-k)2n}}{\operatorname{dist}(K_1, K_2)^{n-\varepsilon p}}.$$

Thus the first quantity is bounded;
$$\frac{1}{\varepsilon} \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{n-\varepsilon p} \frac{\nu(\mathcal{K})}{\nu(\pi_h \mathcal{K})} \leq C(n).$$

On the other hand, if $x \in K_1$ and $y \in K_2$ then
$$\operatorname{dist}(K_1, K_2) \leq |x - y| \leq \operatorname{dist}(K_1, K_2) + 2\sqrt{n}(2^{-k})$$
$$\leq 2\sqrt{n}(\operatorname{dist}(K_1, K_2) + 2^{-k}).$$

Thus by definition of ν
$$\nu(\mathcal{K}) \geq \frac{(2^{-k})^{2n}}{(2\sqrt{n})^{n-\varepsilon p}(\operatorname{dist}(K_1, K_2) + 2^{-k})^{n-\varepsilon p}}.$$

Since
$$\nu(\pi_h \mathcal{K}) = \frac{C(n)}{\varepsilon}(2^{-k})^{n+\varepsilon p},$$

we have
$$\varepsilon \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{\varepsilon p - n} \frac{\nu(\pi_h \mathcal{K})}{\nu(\mathcal{K})} \leq c(n) \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{\varepsilon p - n} \frac{(\operatorname{dist}(K_1, K_2) + 2^{-k})^{n-\varepsilon p}}{(2^{-k})^{n-\varepsilon p}}$$
$$\leq c(n) \left(\frac{\operatorname{dist}(K_1, K_2)}{2^{-k}} \right)^{\varepsilon p' - n} \frac{(\operatorname{dist}(K_1, K_2))^{n-\varepsilon p}}{(2^{-k})^{n-\varepsilon p}} \leq C(n),$$

where in the second inequality we used that $\operatorname{dist}(K_1, K_2) \geq 2^{-k}$. Thus the second term is bounded by a dimensional constant as well, and so (4.40) holds for some constant C_{dd} depending only on n. $\qquad \square$

Lemma 4.4. *There exists a constant C_{ddd} depending only on n such that*

(4.41)
$$\sup \left\{ \frac{\nu(\widetilde{\mathcal{K}})}{\nu(\mathcal{K})} \ : \ \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K}, \, \operatorname{dist}(\widetilde{K_1}, \widetilde{K_2}) \geq 2^{-k(\mathcal{K})} \right\} \leq C_{ddd}.$$

Proof. The triangle inequality gives
(4.42)
$$|x - y| \leq 2\sqrt{n}2^{-k(\mathcal{K})+1} + \operatorname{dist}(\widetilde{K_1}, \widetilde{K_2}) \leq 8\sqrt{n}\operatorname{dist}(\widetilde{K_1}, \widetilde{K_2}),$$

whenever $x \in K_1$, $y \in K_2$ and $\operatorname{dist}(\widetilde{K_1}, \widetilde{K_2}) \geq 2^{-k(\mathcal{K})}$. Therefore,
$$\nu(\widetilde{\mathcal{K}}) \leq \operatorname{dist}(\widetilde{K_1}, \widetilde{K_2})^{-n-\varepsilon p}|\widetilde{K_1}||\widetilde{K_2}|$$
$$= 4^n \operatorname{dist}(\widetilde{K_1}, \widetilde{K_2})^{-n-\varepsilon p} \int_{K_1} \int_{K_2} \mathrm{d}x \, \mathrm{d}y$$
$$\overset{(4.42)}{\leq} 4^n 8\sqrt{n} \int_{K_1} \int_{K_2} \frac{1}{|x - y|^{n-\varepsilon p}} \mathrm{d}x \, \mathrm{d}y = C(n)\nu(\mathcal{K}).$$

$\square$

4.3.3. *Calderón-Zygmund Covering and Sorting of Dyadic Cubes.* Here we write down a version of the classical Calderón-Zygmund decomposition adapted for use in our context. The proof is virtually identical to the classical proof, taking into account that the measure is doubling and absolutely continuous with respect to Lebesgue measure; see [4, 12].

Theorem 4.5. *Let $Q_0 \subset \mathbb{R}^{2n}$ be a cube and let $V \geq 0$ be a function in $L^1(Q_0)$. Let $\widetilde{\lambda}$ be a number such that*

$$\fint_{Q_0} V \, d\nu \leq \widetilde{\lambda}.$$

Then there exists a collection of at most countable cubes $\{Q_i\}$ that are pairwise disjoint with sides parallel to those of Q_0 such that

$$\widetilde{\lambda} < \fint_{Q_i} V \, d\nu \qquad and \qquad \fint_{\widetilde{Q}_i} V \, d\nu \leq \widetilde{\lambda} \quad for \ every \ Q_i \,,$$

where $\widetilde{Q}_i$ is the predecessor of Q_i, and

$$V \leq \widetilde{\lambda} \quad almost \ everywhere \ in \ Q_0 \setminus \bigcup_i Q_i \,.$$

We will use this theorem to cover the level set of H. Define

$$(4.43) \qquad \lambda_2 := \max \left\{ \lambda_1, \sup_{\mathcal{K} \in \Delta_{k_0}} \left(\fint_{\mathcal{K}} H^{p'} \, d\nu \right)^{1/p'} \right\} \,.$$

Recall that λ_1 and k_0 have been defined in (4.7) and (4.34) respectively. Note that by (4.35) the cubes $\{\mathcal{K}\}_{\mathcal{K} \in \Delta_{k_0}}$ cover $\mathcal{B}(x_0, \beta)$. For $\lambda \geq \lambda_2$ apply Theorem 4.5 with $Q_0 = \mathcal{K}_0$ for each and every cube $\mathcal{K}_0 \in \Delta_{k_0}$. We obtain a pairwise disjoint family of cubes $Q_i(\mathcal{K}_0)$ such that

$$(4.44) \qquad \lambda^{p'} < \fint_{Q_i(\mathcal{K}_0)} H^{p'} \, d\nu \qquad and \qquad \fint_{\widetilde{Q}_i(\mathcal{K}_0)} H^{p'} \, d\nu \leq \widetilde{\lambda}^{p'} \quad for \ every \ Q_i(\mathcal{K}_0) \,,$$

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

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

Thus, we get a countable collection of disjoint dyadic cubes

$$\mathcal{H}_\lambda := \bigcup_{\mathcal{K}_0 \in \Delta_{k_0}} \{Q_i(\mathcal{K}_0)\} = \{\mathcal{K}\}$$

that satisfy

$$(4.45) \qquad \lambda^{p'} < \fint_{\mathcal{K}} H^{p'} \, d\nu \quad and \quad \fint_{\widetilde{\mathcal{K}}} H^{p'} \, d\nu \leq \lambda^{p'} \qquad for \ every \ \mathcal{K} \in \mathcal{H}_\lambda$$

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

$$(4.46) \qquad H \leq \lambda \quad \text{a.e. in } \mathcal{B}(x_0, \alpha) \setminus \bigcup_{\mathcal{K} \in \mathcal{H}_\lambda} \mathcal{K} \,.$$

4.3.4. *Nearly Diagonal Cubes.* It turns out that the cubes whose distance to the diagonal is smaller than their size can be covered by diagonal balls $\mathcal{B}_j$ chosen in the exit-time argument above, see (4.10)-(4.11). This leaves us to deal with the cubes that are "far" from the diagonal in the next section.

To begin, we define the family of "nearly diagonal" cubes

$$\mathcal{H}_\lambda^d := \{\mathcal{K} \in \mathcal{H}_\lambda \,:\, \mathrm{dist}(\widetilde{K}_1, \widetilde{K}_2) < 2^{-k(\mathcal{K})}, \quad \widetilde{\mathcal{K}} = \widetilde{K}_1 \times \widetilde{K}_2 \text{ is the predecessor of } \mathcal{K}\}.$$

With $\mathcal{K} \in \mathcal{H}_\lambda^d$, let $(\widetilde{x}, \widetilde{x}) \in \mathrm{Diag}$ such that $\mathrm{dist}((\widetilde{x}, \widetilde{x}), \widetilde{\mathcal{K}}) = \mathrm{dist}(\mathrm{Diag}, \widetilde{\mathcal{K}})$ and a diagonal ball $\mathcal{B}(\widetilde{x}, R)$ such that

$$R \geq \frac{5\sqrt{n}}{2} \mathrm{dist}(\widetilde{K}_1, \widetilde{K}_2) + 5\sqrt{n}2^{-k(\mathcal{K})+1}.$$

Using (4.39) for $\widetilde{\mathcal{K}}$, it follows that $\widetilde{\mathcal{K}} \subset \mathcal{B}(\widetilde{x}, R)$. Thus, the diagonal ball $\mathcal{B} \equiv \mathcal{B}(\widetilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$ satisfies $\mathcal{B} \subset \mathcal{K}$. Now note that by (3.3) in Theorem 3.1 there exists a constant $C_d = C_d(n, p)$ such that

$$1 \leq \frac{\nu(\mathcal{B})}{\nu(\mathcal{K})} \leq \frac{C_d}{\varepsilon}.$$

Thus if $\mathcal{K} \in \mathcal{H}_\lambda^d$ then the lower bound in (4.45) gives

$$\lambda^{p'} < \fint_\mathcal{K} H^{p'}\, \mathrm{d}\nu \leq \frac{\nu(\mathcal{B})}{\nu(\mathcal{K})} \fint_\mathcal{B} H^{p'}\, \mathrm{d}\nu \leq \frac{C_d}{\varepsilon} \fint_\mathcal{B} H^{p'}\, \mathrm{d}\nu.$$

By choosing the number $\kappa \in (0, 1]$ introduced in (4.7) to satisfy

$$(4.47) \qquad\qquad\qquad \kappa \in (0, \kappa_0], \qquad \kappa_0 := \frac{\varepsilon^{1/p'}}{(2C_d)^{1/p'}},$$

we have therefore proved that

For every $\mathcal{K} \in \mathcal{H}_\lambda^d$ there exists $\mathcal{B}^\mathcal{K} = B^\mathcal{K} \times B^\mathcal{K}$ such that $\kappa^{p'}\lambda^{p'} < \fint_{\mathcal{B}^\mathcal{K}} H^{p'}\, \mathrm{d}\nu$ with $\mathcal{K} \subset \mathcal{B}^\mathcal{K}$.

Denote the center of $\mathcal{B}^\mathcal{K}$ by $\widetilde{x}$. Then by the choice (4.34) for the lower bound on $k(\mathcal{K})$, it follows that the radius of the ball $\mathcal{B} \equiv \mathcal{B}(\widetilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$ is smaller than $\frac{\alpha-\beta}{40^n}$. Therefore we can apply the exit time condition (4.10) to obtain that $(\widetilde{x}, \widetilde{x}) \in C_{\kappa\lambda}$ and then $\mathcal{B}^\mathcal{K} \subset \mathcal{B}(\widetilde{x}, R(\widetilde{x}))$. By (4.11) it follows that

$$(4.48) \qquad\qquad\qquad \bigcup_{\mathcal{K} \in \mathcal{H}_\lambda^d} \mathcal{K} \subset \bigcup_j 10\mathcal{B}_j.$$

4.3.5. *Off-Diagonal Reverse Hölder Inequalities.* Since the nearly diagonal cubes $\mathcal{H}_\lambda^d$ have been covered by the diagonal cover, we need only consider the off-diagonal cubes

$$(4.49) \qquad \mathcal{H}_\lambda^{nd} := \{\mathcal{K} \in \mathcal{H}_\lambda \,:\, \mathrm{dist}(\widetilde{K}_1, \widetilde{K}_2) \geq 2^{-k(\mathcal{K})}, \quad \widetilde{\mathcal{K}} = \widetilde{K}_1 \times \widetilde{K}_2 \text{ is the predecessor of } \mathcal{K}\}.$$

By (4.37) we also have

$$\mathrm{dist}(K_1, K_2) \leq 2^{-k(\mathcal{K})} \qquad \text{for every } \mathcal{K} \in \mathcal{H}_\lambda^{nd}.$$

Our objective now is to categorize and estimate sums of the measures of cubes in $\mathcal{H}_\lambda^{nd}$. We will use the following lemma to do so. This lemma states that for off-diagonal cubes an "almost reverse Hölder inequality" holds automatically regardless of whether the function u solves an equation. However, diagonal cubes appear in the estimate, which must be treated by a combinatorial argument in subsequent sections.

Lemma 4.6. *Let $k \geq k_0$ and let $\mathcal{K} = K_1 \times K_2 \in \Delta_k$. Then there exists a constant $C_{nd} = C_{nd}(\texttt{data})$ independent of ε such that if $\mathrm{dist}(K_1, K_2) \geq 2^{-k}$ then*

(4.50)
$$
\left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \leq C_{nd} \left(\fint_{\mathcal{K}} H^{\gamma} \, \mathrm{d}\nu \right)^{1/\gamma}
$$
$$
+ \frac{C_{nd}}{\varepsilon^{1/\gamma}} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{(p-1)(s+\varepsilon)} \left[\left(\fint_{\pi_1 \mathcal{K}} H^{\gamma} \, \mathrm{d}\nu \right)^{1/\gamma} + \left(\fint_{\pi_2 \mathcal{K}} H^{\gamma} \, \mathrm{d}\nu \right)^{1/\gamma} \right],
$$

with γ defined as in Corollary 3.4.1. In particular, (4.50) holds whenever $\mathcal{K} \in \mathcal{H}_\lambda^{nd}$.

Proof. First, there exist $x_1 \in \overline{K_1}$ and $y_1 \in \overline{K_2}$ such that $\mathrm{dist}(K_1, K_2) = |x_1 - y_1|$. Then for any $(x, y) \in \mathcal{K}$,

$$
|x - y| \leq \mathrm{dist}(K_1, K_2) + |x_1 - x| + |y_1 - y|
$$
$$
\leq \mathrm{dist}(K_1, K_2) + 2\sqrt{n} 2^{-k}
$$
$$
\leq 3\sqrt{n} \, \mathrm{dist}(K_1, K_2),
$$

since $\mathrm{dist}(K_1, K_2) \geq 2^{-k}$. Therefore,

(4.51)
$$
1 \leq \frac{|x - y|}{\mathrm{dist}(K_1, K_2)} \leq 3\sqrt{n} \quad \text{for all } (x, y) \in \mathcal{K};
$$

the first inequality is a consequence of the definition of $\mathrm{dist}(K_1, K_2)$. Next, by definition of ν we have

(4.52)
$$
\frac{1}{C(n, p)} \frac{4^{-nk}}{\mathrm{dist}(K_1, K_2)^{n-\varepsilon p}} \leq \nu(\mathcal{K}) \leq C(n, p) \frac{4^{-nk}}{\mathrm{dist}(K_1, K_2)^{n-\varepsilon p}}.
$$

We therefore have, using (4.51) and (4.52),

$$
\left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} = \left(\frac{1}{\nu(\mathcal{K})} \int_{K_1} \int_{K_2} \frac{|u(x) - u(y)|^p}{|x - y|^{n+sp}} + a(x, y) \frac{|u(x) - u(y)|^q}{|x - y|^{n+tq}} \, \mathrm{d}y \, \mathrm{d}x \right)^{1/p'}
$$
$$
\leq C \left[\left(\frac{\mathrm{dist}(K_1, K_2)^{n-\varepsilon p - (n+sp)}}{4^{-nk}} \int_{K_1} \int_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x \right) \right.
$$

(4.53)
$$
\left. + \left(\frac{\mathrm{dist}(K_1, K_2)^{n-\varepsilon p - (n+tq)}}{4^{-nk}} \int_{K_1} \int_{K_2} |u(x) - u(y)|^q \, \mathrm{d}y \, \mathrm{d}x \right) \right]^{1/p'}
$$
$$
\leq C \, \mathrm{dist}(K_1, K_2)^{-(p-1)(s+\varepsilon)} \left[\fint_{K_1} \fint_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x \right.
$$
$$
\left. + \mathrm{dist}(K_1, K_2)^{sp-tq} \fint_{K_1} \fint_{K_2} |u(x) - u(y)|^q \, \mathrm{d}y \, \mathrm{d}x \right]^{1/p'},
$$

where $C = C(\texttt{data})$. Using (4.35) we can estimate

$$
\mathrm{dist}(K_1, K_2)^{sp-tq} \leq (2\alpha)^{sp-tq} \leq (3\varrho_0)^{sp-tq} \leq 3^{sp-tq} \equiv C(\texttt{data}),
$$

and we also have

$$
\fint_{K_1} \fint_{K_2} |u(x) - u(y)|^q \, \mathrm{d}y \, \mathrm{d}x \leq (2 \|u\|_{L^\infty(\mathbb{R}^n)})^{q-p} \fint_{K_1} \fint_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x
$$
$$
\leq C(\texttt{data}) \fint_{K_1} \fint_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x.
$$

Combining these two estimates with (4.53) yields

$$(4.54) \qquad \left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \leq C \operatorname{dist}(K_1, K_2)^{-(p-1)(s+\varepsilon)} \left(\fint_{K_1} \fint_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x \right)^{1/p'},$$

where $C = C(\texttt{data})$. Now, using Minkowski's inequality,

$$\left(\fint_{K_1} \fint_{K_2} |u(x) - u(y)|^p \, \mathrm{d}y \, \mathrm{d}x \right)^{1/p} \leq \left(\fint_{K_1} |u(x) - (u)_{K_1}|^p \, \mathrm{d}x \right)^{1/p}$$
$$+ \left(\fint_{K_2} |u(x) - (u)_{K_2}|^p \, \mathrm{d}x \right)^{1/p} + |(u)_{K_1} - (u)_{K_2}|.$$

Then using the Sobolev embedding theorem in Lemma 3.2 adapted for U and ν and applied to cubes, we have

$$\left(\fint_{K_h} |u(x) - (u)_{K_1}|^p \, \mathrm{d}x \right)^{1/p} \leq \frac{C 2^{-k(s+\varepsilon)}}{\varepsilon^{1/\eta}} \left(\fint_{\pi_h \mathcal{K}} U^\eta \, \mathrm{d}\nu \right)^{1/\eta}, \qquad h \in \{1, 2\}, C = C(n, s, p).$$

Next, using Hölder's inequality and repeatedly using (4.51) and (4.52),

$$|(u)_{K_1} - (u)_{K_2}| \leq \fint_{K_1} \fint_{K_2} |u(x) - u(y)| \, \mathrm{d}y \, \mathrm{d}x$$
$$\leq \left(\fint_{K_1} \fint_{K_2} |u(x) - u(y)|^\eta \, \mathrm{d}y \, \mathrm{d}x \right)^{1/\eta}$$
$$\leq C \left(\frac{1}{\operatorname{dist}(K_1, K_2)^{n - \varepsilon p} \nu(\mathcal{K})} \int_{K_1} \int_{K_2} |u(x) - u(y)|^\eta \, \mathrm{d}y \, \mathrm{d}x \right)^{1/\eta}$$
$$\leq C \left(\fint_{\mathcal{K}} |u(x) - u(y)|^\eta \, \mathrm{d}\nu \right)^{1/\eta} \leq C \operatorname{dist}(K_1, K_2)^{s+\varepsilon} \left(\fint_{\mathcal{K}} U^\eta \, \mathrm{d}\nu \right)^{1/\eta},$$

where $C = C(\texttt{data})$. Combining the three above inequalities with (4.54) and using the inequality $(a + b + c)^r \leq 3^{r-1}(a^r + b^r + c^r)$ valid for $r \geq 1$ and $a, b, c \geq 0$ gives

$$\left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \leq C_{nd} \left(\fint_{\mathcal{K}} U^\eta \, \mathrm{d}\nu \right)^{(p-1)/\eta}$$
$$+ \frac{C_{nd}}{\varepsilon^{(p-1)/\eta}} \left(\frac{2^{-k}}{\operatorname{dist}(K_1, K_2)} \right)^{(p-1)(s+\varepsilon)} \left[\left(\fint_{\pi_1 \mathcal{K}} U^\eta \, \mathrm{d}\nu \right)^{(p-1)/\eta} + \left(\fint_{\pi_2 \mathcal{K}} U^\eta \, \mathrm{d}\nu \right)^{(p-1)/\eta} \right].$$

To see that (4.50) follows, recall that $\gamma = \frac{\eta}{p-1}$ and that by definition $U^\eta \leq H^\gamma$ pointwise. $\qquad \square$

Note that this lemma holds for all functions $u \in W^{s,p}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and for every integer k, and in the case $a \equiv 0$ it holds for every $u \in W^{s,p}(\mathbb{R}^n)$. We now apply it in order to begin the level set estimate.

Corollary 4.6.1. *Let $k \geq k_0$ be an integer, and suppose that $\mathcal{K} \in \Delta_k$ satisfies $\operatorname{dist}(K_1, K_2) \geq 2^{-k}$. Assume that*

$$\left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \geq \lambda,$$

and that the constant κ introduced in (4.7) satisfies

$$(4.55) \qquad \kappa \in (0, \kappa_1], \qquad \kappa_1 := \frac{\varepsilon^{1/\gamma}}{2^{1/\gamma} 3 C_{nd}},$$

where $C_{nd} \equiv C_{nd}(\textbf{\textit{data}})$ has been defined in Lemma 4.6. Then

(4.56)
$$\nu(\mathcal{K}) \leq \frac{3^\gamma C_{nd}^\gamma}{\lambda^\gamma} \int_{\mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu$$
$$+ \frac{3^\gamma C_{nd}^\gamma}{\varepsilon \lambda^\gamma} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{\eta(s+\varepsilon)} \left[\frac{\nu(\mathcal{K})}{\nu(\pi_1 \mathcal{K})} \int_{\pi_1 \mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu + \frac{\nu(\mathcal{K})}{\nu(\pi_2 \mathcal{K})} \int_{\pi_2 \mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu \right] .$$

In particular, (4.56) holds whenever $\mathcal{K} \in \mathcal{H}_\lambda^{nd}$.

Proof. Applying the inequality $(a + b + c)^r \leq 3^{r-1}(a^r + b^r + c^r)$ to the conclusion (4.50) in Lemma 4.6 and recalling that $(p-1)\gamma = \eta$, we get

(4.57)
$$\frac{\lambda^\gamma}{3^{\gamma-1} C_{nd}^\gamma} \leq \fint_{\mathcal{K}} H^\gamma \, \mathrm{d}\nu + \frac{1}{\varepsilon} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{\eta(s+\varepsilon)} \left(\fint_{\pi_1 \mathcal{K}} H^\gamma \, \mathrm{d}\nu + \fint_{\pi_2 \mathcal{K}} H^\gamma \, \mathrm{d}\nu \right) .$$

To estimate the integrals on the right-hand side, we use (4.55) to get

$$\fint_E H^\gamma \, \mathrm{d}\nu \leq \kappa_1^\gamma \lambda^\gamma + \frac{1}{\nu(E)} \int_{E \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu$$

for $E \in \{\mathcal{K}, \pi_1 \mathcal{K}, \pi_2 \mathcal{K}\}$, and therefore

$$\frac{\lambda^\gamma}{3^{\gamma-1} C_{nd}^\gamma} \leq \frac{3\kappa_1^\gamma \lambda^\gamma}{\varepsilon} + \frac{1}{\nu(\mathcal{K})} \int_{\mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu$$
$$+ \frac{1}{\varepsilon} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{\eta(s+\varepsilon)} \left[\frac{1}{\nu(\pi_1 \mathcal{K})} \int_{\pi_1 \mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu + \frac{1}{\nu(\pi_2 \mathcal{K})} \int_{\pi_2 \mathcal{K} \cap \{H > \kappa \lambda\}} H^\gamma \, \mathrm{d}\nu \right] .$$

Now the estimate (4.56) follows by using (4.55) in the above estimate and absorbing terms. $\qquad \square$

4.3.6. *Collections of Off-Diagonal Cubes.* We now split the collection $\mathcal{H}_\lambda^{nd}$ into collections where the average of H^γ is large and where the average of H^γ is small. The splitting is chosen based on the first exit-time argument.

Consider

(4.58)
$$\mathscr{G}_\lambda^h := \left\{ \mathcal{K} \in \mathcal{H}_\lambda^{nd} \; : \; \fint_{\pi_h \mathcal{K}} H^\gamma \, \mathrm{d}\nu \leq (10n)^{n+p} \kappa^\gamma \lambda^\gamma \right\}$$

and

(4.59)
$$\mathscr{B}_\lambda^h := \left\{ \mathcal{K} \in \mathcal{H}_\lambda^{nd} \; : \; \fint_{\pi_h \mathcal{K}} H^\gamma \, \mathrm{d}\nu > (10n)^{n+p} \kappa^\gamma \lambda^\gamma \right\}$$

for $h \in \{1, 2\}$ and κ and γ introduced in (4.7) and Corollary 3.4.1 respectively. We further define
$$\mathscr{G}_\lambda := \mathscr{G}_\lambda^1 \cap \mathscr{G}_\lambda^2, \qquad \text{and} \qquad \mathscr{B}_\lambda := \mathscr{B}_\lambda^1 \cup \mathscr{B}_\lambda^2 .$$

We further split the set $\mathscr{B}_\lambda$ in order to remove cubes that are already covered by the diagonal balls in (4.10)-(4.11):

(4.60)
$$\mathscr{B}_{\lambda,d} := \left\{ \mathcal{K} \in \mathscr{B}_\lambda \; : \; \mathcal{K} \subset \bigcup_j 10\mathcal{B}_j \right\}, \qquad \mathscr{B}_{\lambda,nd} := \mathscr{B}_\lambda \setminus \mathscr{B}_{\lambda,d} .$$

Thus we have the decomposition into disjoint families
$$\mathcal{H}_\lambda^{nd} = \mathscr{G}_\lambda \cup \mathscr{B}_{\lambda,d} \cap \mathscr{B}_{\lambda,nd} .$$

It turns out that the family $\mathscr{G}_\lambda$ is "good" in the sense that the measures of cubes in this collection are estimated by the μ-measure of the level set $\{H > \kappa\lambda\}$, where $\mathrm{d}\mu := H^{p'}\mathrm{d}\nu$. The cubes belonging to $\mathscr{B}_{\lambda,nd}$ are "bad" because there is no available control of the size of $H^{p'}$ on diagonal cubes via the exit-time argument. They will instead be dealt with using combinatorial arguments, and the cutoff $(10n)^{n+p}\kappa^\gamma\lambda^\gamma$ is chosen for precisely this purpose.

Lemma 4.7 (First (easier) off-diagonal estimate)**.** *We have*

$$\sum_{\mathcal{K}\in\mathscr{G}_\lambda} \nu(\mathcal{K}) \leq \frac{6^\gamma C_{nd}^\gamma}{\lambda^\gamma} \int_{\mathcal{B}(x_0,\alpha)\cap\{H>\kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$

whenever the number $\kappa \in (0,1]$ satisfies

(4.61)
$$\kappa \in (0, \kappa_2], \qquad \kappa_2 := \frac{\varepsilon^{1/\gamma}}{8^{1/\gamma} 3 C_{nd} (10n)^{(n+p)/\gamma}} \, .$$

The constant C_{nd} has been defined in Lemma 4.6.

Proof. Since the cubes $\mathcal{K} \in \mathscr{G}_\lambda$ are disjoint and since (4.35) holds, it suffices to show that

(4.62)
$$\nu(\mathcal{K}) \leq \frac{6^\gamma C_{nd}^\gamma}{\lambda^\gamma} \int_{\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$

for every $\mathcal{K} \in \mathscr{G}_\lambda$. Since $\mathcal{K} \in \mathscr{G}_\lambda$, using (4.61) we have

(4.63)
$$\frac{3^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \frac{\nu(\mathcal{K})}{\nu(\pi_h\mathcal{K})} \int_{\pi_h\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$
$$\leq \frac{3^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \nu(\mathcal{K}) \fint_{\pi_h\mathcal{K}} H^\gamma \, \mathrm{d}\nu \leq \nu(\mathcal{K})\frac{3^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma}(10n)^{n+p}\kappa^\gamma\lambda^\gamma \leq \frac{\nu(\mathcal{K})}{8} \, .$$

Using this estimate for $h \in \{1,2\}$ in (4.56) and then absorbing terms gives (4.62). $\qquad\square$

4.3.7. *Determining κ.* At this point the constant κ introduced in (4.7) can be completely determined. We choose

(4.64)
$$\kappa := \min\{\kappa_0,\kappa_1,\kappa_2\} \equiv \min\left\{\frac{\varepsilon^{1/p'}}{(2C_d)^{1/p}}, \frac{\varepsilon^{1/\gamma}}{2^{1/\gamma}3C_{nd}}, \frac{\varepsilon^{1/\gamma}}{8^{1/\gamma}3C_{nd}(10n)^{(n+p)/\gamma}}\right\},$$

so that conditions (4.47), (4.55), and (4.61) are satisfied. Since the constant C_d defined in Theorem 3.1 depends only on n and p, and since C_{nd} defined in Lemma 4.6 depends only on `data`, we conclude that there exists $C_\kappa = C_\kappa(\mathtt{data})$ such that

(4.65)
$$\kappa \geq \frac{\varepsilon^{1/\gamma}}{C_\kappa} \, .$$

4.3.8. *Summation in $\mathscr{B}_{\lambda,nd}$.* Dealing with the cubes from $\mathscr{B}_{\lambda,nd}$ requires delicate estimates and combinatorial arguments. We will first set up notation designed to describe the cubes more precisely. We will then demonstrate that the choice of cutoff $(10n)^{n+p}\kappa^\gamma\lambda^\gamma$ gives us an upper bound on the distance of a cube from the diagonal.

Define the "problematic" projections of the "bad" off-diagonal set

$$\pi_h\mathscr{B}_\lambda := \{\pi_h\mathcal{K} : \mathcal{K} \in \mathscr{B}_\lambda^h\}, \quad h \in \{1,2\} \, .$$

Since all cubes belonging to the collection $\pi_1\mathscr{B}_\lambda \cup \pi_2\mathscr{B}_\lambda$ are dyadic cubes, it follows that a disjoint subfamily of cubes always exists. We denote this disjoint subfamily of $\pi_1\mathscr{B}_\lambda \cup \pi_2\mathscr{B}_\lambda$ by $\pi\mathscr{B}_\lambda$. By definition, all cubes of $\pi\mathscr{B}_\lambda$ belong to $\pi_1\mathscr{B}_\lambda \cup \pi_2\mathscr{B}_\lambda$ and are therefore diagonal cubes.

Lemma 4.8. *Let $\mathcal{K} \in \mathscr{B}_{\lambda,nd}$ be a cube such that $\pi_h\mathcal{K} \subset \mathcal{M}$ for some $\mathcal{M} \in \pi\mathscr{B}_\lambda$ and some $h \in \{1,2\}$. Then* $\mathrm{dist}(K_1, K_2) \geq 2^{-k(\mathcal{M})}$.

Proof. First, consider $\mathcal{M} \in \pi\mathscr{B}_\lambda$. Take the diagonal ball $\mathcal{B}(\mathcal{M}) \equiv \mathcal{B}(x_\mathcal{M}, 2^{-(k(\mathcal{M})+1)})$ with $x_\mathcal{M}$ being the center of $\mathcal{M}$. Thus,

$$\mathcal{B}(\mathcal{M}) \subset \mathcal{M} \subset \sqrt{n}\mathcal{B}(\mathcal{M}).$$

Thus, by (3.2), Hölder's inequality, and by the definition of $\pi\mathscr{B}_\lambda$

$$(10n)^{(n+p)/\gamma}\kappa\lambda < \left(\fint_\mathcal{M} H^\gamma \, \mathrm{d}\nu \right)^{1/\gamma}$$

$$\leq \left(\frac{\nu(10n\mathcal{B}(\mathcal{M}))}{\nu(\mathcal{B}(\mathcal{M}))} \fint_\mathcal{M} H^\gamma \, \mathrm{d}\nu \right)^{1/\gamma}$$

$$\leq (10n)^{(n+p)\gamma} \left(\fint_\mathcal{M} H^{p'} \, \mathrm{d}\nu \right)^{1/p'}.$$

By the definition of $D_{\kappa\lambda}$ in (4.8) we have $(x_\mathcal{M}, x_\mathcal{M}) \in D_{\kappa\lambda}$, and then the exit-time condition (4.10) gives $\mathcal{B}(\mathcal{M}) \subset \mathcal{B}(x_\mathcal{M}, R(\mathcal{M}))$. We can use the exit time condition since the radius of the ball $10n\mathcal{B}(\mathcal{M})$ is smaller than $\frac{\alpha-\beta}{40^n}$, which is in turn a consequence of the fact that $k(\mathcal{M}) + 1 \geq k_0$ and k_0 has been chosen as in (4.34). Therefore, by (4.11)

$$(4.66) \qquad 10n\mathcal{B}(\mathcal{M}) \subset \bigcup_j 10\mathcal{B}_j.$$

Now, assume by contradiction that $\mathrm{dist}(K_1, K_2) < 2^{-k(\mathcal{M})}$. We will show that

$$(4.67) \qquad \mathcal{K} \subset 10n\mathcal{B}(\mathcal{M}),$$

which then contradicts the assumption $\mathcal{M} \in \mathscr{B}_{\lambda,nd}$ by (4.66). To show (4.67) we use Proposition 4.2 and that $\pi_h\mathcal{K} \subset \mathcal{M}$ to get

$$\mathrm{dist}(\mathcal{K}, \mathcal{M}) \leq \mathrm{dist}(\mathcal{K}, \pi_h\mathcal{K}) = \mathrm{dist}(K_1, K_2) \leq 2^{-k(\mathcal{M})}.$$

Again using Proposition 4.2 we have $k(\mathcal{K}) = k(\pi_h\mathcal{K})$ and $k(\mathcal{K}) \geq k(\mathcal{M})$. Therefore, since $\mathcal{M} \subset \sqrt{n}\mathcal{B}(\mathcal{M})$ and the radius of $\mathcal{B}(\mathcal{M})$ is $2^{-(k(\mathcal{M})+1)}$, then (4.67) must hold. $\qquad\square$

Lemma 4.9 (Second (harder) off-diagonal estimate). *There exists a constant $C = C(\mathtt{data})$ such that the estimate*

$$(4.68) \qquad \sum_{\mathcal{K} \in \mathscr{B}_{\lambda,nd}} \nu(\mathcal{K}) \leq \frac{C}{\lambda^\gamma} \int_{\mathcal{B}(x_0,\alpha) \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$

holds, where κ is as in (4.64).

Proof. <u>Step 1: Classification.</u> We classify cubes from $\mathscr{B}_{\lambda,nd}$ according to the location of their projections, their size, and their distance from the diagonal. Lemma 4.8, in summary, correlates the distance of a cube from the diagonal with the location of its projection. This allows us to consider only cubes in $\pi\mathscr{B}_\lambda$.

We will partition $\mathscr{B}_{\lambda,nd}$ into suitable disjoint subfamilies. Define the collections

$$\mathscr{B}_{\lambda,nd}^h := \mathscr{B}_{\lambda,nd} \cap \mathscr{B}_\lambda^h, \qquad h \in \{1,2\}$$

For every $\mathcal{M} \in \pi\mathscr{B}_\lambda$, set

$$\mathscr{B}^h_{\lambda,nd}(\mathcal{M}) := \{\mathcal{K} \in \mathscr{B}_{\lambda,nd} : \pi_h\mathcal{K} \subset \mathcal{M}\}, \quad h \in \{1,2\}.$$

Thus, we have the decomposition into disjoint subcollections

(4.69) $$\mathscr{B}^h_{\lambda,nd} = \bigcup_{\mathcal{M}\in\pi\mathscr{B}_\lambda} \mathscr{B}^h_{\lambda,nd}(\mathcal{M}).$$

These subcollections are disjoint in the sense that for $\mathcal{M}_1, \mathcal{M}_2 \in \pi\mathscr{B}_\lambda$, $\mathscr{B}^h_{\lambda,nd}(\mathcal{M}_1) \cap \mathscr{B}^h_{\lambda,nd}(\mathcal{M}_2) \neq \emptyset$ implies $\mathcal{M}_1 = \mathcal{M}_2$. This follows from the definition of $\mathscr{B}^h_{\lambda,nd}(\mathcal{M})$ and from the fact that all elements in $\pi\mathscr{B}_\lambda$ are pairwise disjoint.

Next we classify cubes according to their size. For each $\mathcal{K} \in \mathscr{B}^h_{\lambda,nd}(\mathcal{M})$ we have $k(\mathcal{K}) = k(\pi_h\mathcal{K}) \geq k(\mathcal{M})$, so we can define the collections

$$[\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_i := \{\mathcal{K} \in \mathscr{B}^h_{\lambda,nd}(\mathcal{M}) : k(\mathcal{K}) = i + k(\mathcal{M})\}, \qquad h \in \{1,2\}, \quad i \in \mathbb{Z}_+.$$

We again have a decomposition in mutually disjoint subcollections

$$\mathscr{B}^h_{\lambda,nd}(\mathcal{M}) = \bigcup_{i\in\mathbb{Z}_+} [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_i,$$

in the sense that $[\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i_1} \cap [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i_2} \neq \emptyset$ implies that $i_1 = i_2$.

We further classify the cubes according to their distance from the diagonal. We now use the combinatorial information obtained from Lemma 4.8. Take $\mathcal{M} \in \pi\mathscr{B}_\lambda$. If $\mathcal{K} \in \mathscr{B}^h_{\lambda,nd}(\mathcal{M})$; that is, if $\pi_h\mathcal{K} \subset \mathcal{M}$, then by Lemma 4.8 it follows that $\mathrm{dist}(K_1, K_2) \geq 2^{-k(\mathcal{M})}$. This leads us to the definition of the subcollections

$$[\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i,j} := \{\mathcal{K} \in [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_i : 2^{j-k(\mathcal{M})} \leq \mathrm{dist}(K_1,K_2) < 2^{j+1-k(\mathcal{M})}\}, \qquad h \in \{1,2\}, \quad i,j \in \mathbb{Z}_+,$$

We yet again have the decomposition in mutually disjoint subcollections

$$\mathscr{B}^h_{\lambda,nd}(\mathcal{M}) = \bigcup_{i,j\in\mathbb{Z}_+} [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i,j}$$

and these are mutually disjoint in the sense that $[\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i_1,j_1} \cap [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i_2,j_2} \neq \emptyset$ implies that $(i_1,j_1) = (i_2,j_2)$. In summary, we have the decomposition

$$\mathscr{B}^h_{\lambda,nd} = \bigcup_{\mathcal{M}\in\pi\mathscr{B}_\lambda} \bigcup_{i,j\in\mathbb{Z}_+} [\mathscr{B}^h_{\lambda,nd}(\mathcal{M})]_{i,j}.$$

Step 2: Summation and Further Partitioning. Fix $\mathcal{M} \in \pi\mathscr{B}_\lambda$. Our goal in this step is to show that for $h \in \{1,2\}$

(4.70)

$$\frac{1}{\varepsilon} \sum_{\mathcal{K}\in\mathscr{B}^h_{\lambda,nd}(\mathcal{M})} \frac{\nu(\mathcal{K})}{\nu(\pi_h\mathcal{K})} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \int_{\pi_h\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu \leq \frac{C(\mathtt{data})}{s^2} \int_{\mathcal{M}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu.$$

To begin, note that for $\mathcal{K} \in \mathscr{B}^h_{\lambda,nd}$ we have $\mathrm{dist}(K_1,K_2) \geq 2^{-k(\mathcal{K})}$, and so we can apply the inequality from Lemma 4.3 to get that

$$\frac{1}{\varepsilon}\frac{\nu(\mathcal{K})}{\nu(\pi_h\mathcal{K})} \leq C_{dd}\left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{n-\varepsilon p}$$

for $h \in \{1,2\}$, and further if $\mathcal{K} \in [\mathscr{B}^h_{\lambda,nd}]_{i,j}$ then

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

With these we can estimate

$$
\frac{1}{\varepsilon} \sum_{\mathcal{K} \in \mathscr{B}_{\lambda,nd}^h(\mathcal{M})} \frac{\nu(\mathcal{K})}{\nu(\pi_h \mathcal{K})} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{\eta(s+\varepsilon)} \int_{\pi_h \mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu
$$

$$
\leq C_{dd} \sum_{\mathcal{K} \in \mathscr{B}_{\lambda,nd}^h(\mathcal{M})} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{n + \eta(s+\varepsilon) - \varepsilon p} \int_{\pi_h \mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu
$$

(4.71)

$$
= C_{dd} \sum_{i,j=0}^\infty \sum_{\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{n + \eta(s+\varepsilon) - \varepsilon p} \int_{\pi_h \mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu
$$

$$
\leq C_{dd} \sum_{i,j=0}^\infty \left(\frac{1}{2^{i+j}} \right)^{n + \eta(s+\varepsilon) - \varepsilon p} \sum_{\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}} \int_{\pi_h \mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,.
$$

In order to evaluate the last sum we must further partition $[\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}$. In order to do so, note that for each $i \in \mathbb{Z}_+$ the cube $\mathcal{M}$ contains precisely 2^{2ni} disjoint dyadic cubes belonging to $\Delta_{i+k(\mathcal{M})}$ and precisely 2^{ni} disjoint dyadic diagonal cubes from $\Delta_{i+k(\mathcal{M})}^d$; see the definition (4.36) and the text immediately following. Thus, $\mathcal{M}$ contains at most 2^{ni} dyadic cubes from the class $\Delta_{i+k(\mathcal{M})}^d \cap (\pi_1 \mathscr{B}_\lambda \cap \pi_2 \mathscr{B}_\lambda)$. In any case we consider all the diagonal cubes from $\mathcal{M}$ in $\Delta_{i+k(\mathcal{M})}^d$ and label them as

$$
\{\widetilde{\mathcal{M}} \in \Delta_{i+k(\mathcal{M})}^d \;:\; \widetilde{\mathcal{M}} \subset \mathcal{M}\} \equiv \{\mathcal{M}_i^m \;:\; 1 \leq m \leq 2^{ni}\} \,,
$$

so that

(4.72)
$$
\sum_{m=1}^{2^{ni}} \int_{\mathcal{M}_i^m \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \leq \int_{\mathcal{M} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,.
$$

For any $\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}$ with $h \in \{1, 2\}$ there exists a unique diagonal cube from the collection Δ_k^d, which we denote by $\mathcal{M}_i^m(\mathcal{K})$, such that $\pi_h \mathcal{K} = \mathcal{M}_i^m(\mathcal{K})$. Then we can split $[\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}$ into subsets

$$
[\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j,m} := \{\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j} \;:\; \pi_h \mathcal{K} = \mathcal{M}_i^m\}, \qquad m \in \{1, \ldots, 2^{ni}\} \,.
$$

Since $\mathscr{B}_{\lambda,nd}^1$ is a family of dyadic cubes, if $\mathcal{K}_1, \mathcal{K}_2 \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j,m}$ and $\mathcal{K}_1 \neq \mathcal{K}_2$, then $\pi_2 \mathcal{K}_1 \cap \pi_2 \mathcal{K}_2 = \emptyset$, else the two cubes would coincide.

Thus for each i, $j \geq 0$ and for $m \in \{1, \ldots, 2^{ni}\}$ the number of cubes $\mathcal{K}$ of side length $2^{i+k(\mathcal{M})}$ with $\mathrm{dist}(K_1, K_2) \in [2^{j-k(\mathcal{M})}, 2^{j+1-k(\mathcal{M})})$ that project into $\mathcal{M}_i^m$ is bounded from above by $C(n) \left(\frac{\text{length of half-open interval defining the distance from the diagonal}}{\text{cube side length}} \right)^n$; that is,

(4.73)
$$
\#[\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j,m} \leq C(n) 2^{n(i+j)} \,, \qquad h \in \{1, 2\} \,.
$$

Thus, using (4.72) and (4.73) we estimate

$$
\sum_{\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}} \int_{\pi_h \mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu = \sum_{m=1}^{2^{ni}} \sum_{\mathcal{K} \in [\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j,m}} \int_{\mathcal{M}_i^m \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu
$$

$$
\leq C(n) 2^{n(i+j)} \sum_{m=1}^{2^{ni}} \int_{\mathcal{M}_i^m \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu
$$

$$
\leq C(n) 2^{n(i+j)} \int_{\mathcal{M} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,.
$$

Then using the series estimate (2.1) it follows that

$$\sum_{i,j=0}^{\infty} \left(\frac{1}{2^{i+j}}\right)^{n+\eta(s+\varepsilon)-\varepsilon p} \sum_{\mathcal{K}\in[\mathscr{B}_{\lambda,nd}^h(\mathcal{M})]_{i,j}} \int_{\pi_h\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$\leq C(n) \sum_{i,j=0}^{\infty} \left(\frac{1}{2^{i+j}}\right)^{\eta(s+\varepsilon)-\varepsilon p} \int_{\mathcal{M}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$\leq C(n) \left[\frac{2^{\eta(s+\varepsilon)-\varepsilon p}}{\ln(2)(\eta(s+\varepsilon)-\varepsilon p)}\right]^2 \int_{\mathcal{M}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$\leq \frac{C(n,p)}{[\eta(s+\varepsilon)-\varepsilon p]^2} \int_{\mathcal{M}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu \leq \frac{C(n,p)}{s^2} \int_{\mathcal{M}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu \,.$$

In the last line we used that $\eta > 1$ and that $\varepsilon < \frac{s}{p}$, which gives $\eta(s+\varepsilon) - \varepsilon p > \frac{s}{p}$. Combining this inequality with (4.71) results in (4.70).

Step 3: Further Summation. Let $\mathcal{K} \in \mathscr{B}_{\lambda,nd}^1$. Then, either $\mathcal{K} \in \mathscr{G}_\lambda^2$ or $\mathcal{K} \in \mathscr{B}_\lambda^2$ (see (4.58) and (4.59)). If $\mathcal{K} \in \mathscr{G}_\lambda^2$, then using the almost-reverse Hölder inequality (4.56) along with the estimate (4.63) for cubes in $\mathscr{G}_\lambda^2$ and reabsorbing terms,

$$\nu(\mathcal{K}) \leq \frac{6^\gamma C_{nd}^\gamma}{\lambda^\gamma} \int_{\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu + \frac{6^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \frac{\nu(\mathcal{K})}{\nu(\pi_1\mathcal{K})} \int_{\pi_1\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu \,.$$

On the other hand, if $\mathcal{K} \in \mathscr{B}_\lambda^2$ then we can use the almost-reverse Hölder inequality (4.56) to obtain

$$\nu(\mathcal{K}) \leq \frac{3^\gamma C_{nd}^\gamma}{\lambda^\gamma} \int_{\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$+ \frac{3^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \left[\frac{\nu(\mathcal{K})}{\nu(\pi_1\mathcal{K})} \int_{\pi_1\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu + \frac{\nu(\mathcal{K})}{\nu(\pi_2\mathcal{K})} \int_{\pi_2\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu\right] \,.$$

A similar reasoning holds for $\mathcal{K} \in \mathscr{B}_\lambda^2$. Summing over the cubes $\mathcal{K} \in \mathscr{B}_{\lambda,nd} = \mathscr{B}_{\lambda,nd}^1 \cup \mathscr{B}_{\lambda,nd}^2$ gives us

(4.74)
$$\sum_{\mathcal{K}\in\mathscr{B}_{\lambda,nd}} \nu(\mathcal{K}) \leq \frac{6^\gamma C_{nd}^\gamma}{\lambda^\gamma} \sum_{\mathcal{K}\in\mathscr{B}_{\lambda,nd}} \int_{\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$+ \frac{6^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \sum_{\mathcal{K}\in\mathscr{B}_{\lambda,nd}^1} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \frac{\nu(\mathcal{K})}{\nu(\pi_1\mathcal{K})} \int_{\pi_1\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu$$

$$+ \frac{6^\gamma C_{nd}^\gamma}{\varepsilon\lambda^\gamma} \sum_{\mathcal{K}\in\mathscr{B}_{\lambda,nd}^2} \left(\frac{2^{-k}}{\mathrm{dist}(K_1,K_2)}\right)^{\eta(s+\varepsilon)} \frac{\nu(\mathcal{K})}{\nu(\pi_2\mathcal{K})} \int_{\pi_2\mathcal{K}\cap\{H>\kappa\lambda\}} H^\gamma \, d\nu \,.$$

The point of this argument beginning in Step 3 is that terms involving the projections $\pi_h\mathcal{K}$ appear if and only if $\mathcal{K} \in \mathscr{B}_{\lambda,nd}^h$ for $h \in \{1,2\}$. We will now argue that the last two terms in the above inequality coincide.

For a cube $\mathcal{K} = K_1 \times K_2 \in \Delta$, define

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

Then by definition

(4.75)
$$\pi_1(\mathcal{K}) = \pi_1(K_1 \times K_2) = \pi_2(K_2 \times K_1) = \pi_2(\mathrm{Symm}(\mathcal{K})) \,.$$

Further, by symmetry of H we have

$$\int_{\mathcal{K}} H^r \, \mathrm{d}\nu = \int_{\mathrm{Symm}(\mathcal{K})} H^r \, \mathrm{d}\nu \,, \qquad r \in [1, \infty)$$

and

$$(4.76) \qquad \int_{\mathcal{K} \cap \{H > \kappa\lambda\}} H^r \, \mathrm{d}\nu = \int_{\mathrm{Symm}(\mathcal{K}) \cap \{H > \kappa\lambda\}} H^r \, \mathrm{d}\nu \,, \qquad r \in [1, \infty) \,.$$

Therefore, by definition of $\mathscr{B}_\lambda^h$

$$\mathcal{K} \in \mathscr{B}_\lambda^1 \quad \Leftrightarrow \quad \mathrm{Symm}(\mathcal{K}) \in \mathscr{B}_\lambda^2 \,,$$

and vice versa, and so

$$(4.77) \qquad \mathcal{K} \in \mathscr{B}_{\lambda, nd}^1 \quad \Leftrightarrow \quad \mathrm{Symm}(\mathcal{K}) \in \mathscr{B}_{\lambda, nd}^2 \,,$$

and vice versa. It follows from (4.75)-(4.77) that the last two terms in (4.74) coincide. Recalling the decomposition (4.69), (4.74) can be written as

$$\sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \nu(\mathcal{K}) \leq \frac{C}{\lambda^\gamma} \sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \int_{\mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$

$$+ \frac{C}{\varepsilon\lambda^\gamma} \sum_{\mathcal{M} \in \pi\mathscr{B}_\lambda} \sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}^1(\mathcal{M})} \left(\frac{2^{-k}}{\mathrm{dist}(K_1, K_2)} \right)^{\eta(s+\varepsilon)} \frac{\nu(\mathcal{K})}{\nu(\pi_1\mathcal{K})} \int_{\pi_1\mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,,$$

for a constant $C \equiv C(\mathtt{data})$. Using the estimate (4.70) proved in Step 2 yields

$$\sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \nu(\mathcal{K}) \leq \frac{C}{\lambda^\gamma} \sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \int_{\mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu + \frac{C}{\lambda^\gamma} \sum_{\mathcal{M} \in \pi\mathscr{B}_\lambda} \int_{\mathcal{M} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,.$$

Then since the collections $\mathscr{B}_{\lambda, nd}$ and $\pi\mathscr{B}_\lambda$ are comprised of mutually disjoint dyadic cubes all of which are contained in $\mathcal{B}(x_0, \alpha)$ (see (4.35)) we can estimate

$$\sum_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \int_{\mathcal{K} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu + \sum_{\mathcal{M} \in \pi\mathscr{B}_\lambda} \int_{\mathcal{M} \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \leq 2 \int_{\mathcal{B}(x_0, \alpha) \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu \,.$$

Thus (4.68) is proved, and the proof of Lemma 4.9 is complete. $\qquad\square$

4.3.9. *Off-Diagonal Conclusion.* The next lemma summarizes the off-diagonal estimate obtained for H.

Lemma 4.10. *The inequality*

$$(4.78) \qquad \int_{\mathcal{B}(x_0, \beta) \cap \{H > \lambda\}} H^{p'} \, \mathrm{d}\nu \leq 10^{n+p} \kappa^{p'} \lambda^{p'} \sum_j \nu(\mathcal{B}_j) + C\lambda^{p'-\gamma} \int_{\mathcal{B}(x_0, \alpha) \cap \{H > \kappa\lambda\}} H^\gamma \, \mathrm{d}\nu$$

holds for a constant $C \equiv C(\mathtt{data})$, while κ has been chosen in (4.64) and satisfies (4.65).

Proof. We have the decomposition $\mathcal{H}_\lambda = \mathcal{H}_\lambda^d \cup \mathcal{H}_\lambda^{nd}$ with $\mathcal{H}_\lambda^{nd} = \mathscr{G}_\lambda \cup \mathscr{B}_{\lambda, d} \cup \mathscr{B}_{\lambda, nd}$. Recall from the relation (4.48) and from the definition (4.60) that

$$\bigcup_{\mathcal{K} \in \mathcal{H}_\lambda^d} \mathcal{K} \cup \bigcup_{\mathcal{K} \in \mathscr{B}_{\lambda, d}} \mathcal{K} \subset \bigcup_j 10\mathcal{B}_j \,.$$

Thus

$$\bigcup_{\mathcal{K} \in \mathcal{H}_\lambda} \subset \bigcup_j 10\mathcal{B}_j \cup \bigcup_{\mathcal{K} \in \mathscr{G}_\lambda} \mathcal{K} \cup \bigcup_{\mathcal{K} \in \mathscr{B}_{\lambda, nd}} \mathcal{K} \,.$$

Now, with this and with (4.46), we estimate

$$\int_{\mathcal{B}(x_0,\beta)\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu \le \sum_j \int_{10\mathcal{B}_j\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu + \sum_{\mathcal{K}\in\mathscr{G}_\lambda\cup\mathscr{B}_{\lambda,nd}} \int_{\mathcal{K}\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu\,.$$

Now, by the choice of Calderón-Zygmund splitting in (4.44), for $\mathcal{K}\in\mathscr{G}_\lambda\cup\mathscr{B}_{\lambda,nd}\subset\mathcal{H}^{nd}_\lambda$

$$\fint_{\mathcal{K}} H^{p'}\,\mathrm{d}\nu \le \frac{\nu(\widetilde{\mathcal{K}})}{\nu(\mathcal{K})}\fint_{\widetilde{\mathcal{K}}} H^{p'}\,\mathrm{d}\nu \le \frac{\nu(\widetilde{\mathcal{K}})}{\nu(\mathcal{K})}\lambda^{p'}\,.$$

Since $\mathcal{K}\in\mathcal{H}^{nd}_\lambda$ it follows from the definition (4.49) that $\mathrm{dist}(\widetilde{K}_1,\widetilde{K}_2)\ge 2^{-k(\mathcal{K})}$, so we can use the geometric estimate (4.41) from Lemma 4.4 to obtain the bound $\frac{\nu(\widetilde{\mathcal{K}})}{\nu(\mathcal{K})}\le C_{ddd}$. In summary,

$$\mathcal{K}\in\mathscr{G}_\lambda\cup\mathscr{B}_{\lambda,nd} \quad\Rightarrow\quad \int_{\mathcal{K}\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu \le C_{ddd}\lambda^{p'}\nu(\mathcal{K})\,.$$

Using this inequality in conjunction with (4.13) results in

$$\int_{\mathcal{B}(x_0,\beta)\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu \le 10^{n+\varepsilon p}\kappa^{p'}\lambda^{p'}\sum_j \nu(\mathcal{B}_j) + C_{ddd}\lambda^{p'}\sum_{\mathcal{K}\in\mathscr{G}_\lambda\cup\mathscr{B}_{\lambda,nd}}\nu(\mathcal{K})\,.$$

Then (4.78) follows from the conclusions of Lemmas 4.7 and 4.9. $\qquad\square$

4.4. **Conclusion of the Proof.** We now come to the conclusion of the proof of Proposition 4.1. We start by combining the diagonal estimate (4.33) with the off-diagonal estimate (4.78). We use the elementary estimate

$$\int_{\mathcal{B}(x_0,\beta)\cap\{H>\widetilde{\kappa}\kappa\lambda\}} H^{p'}\,\mathrm{d}\nu \le \lambda^{p'-\gamma}\int_{\mathcal{B}(x_0,\beta)\cap\{H>\widetilde{\kappa}\kappa\lambda\}} H^{\gamma}\,\mathrm{d}\nu + \int_{\mathcal{B}(x_0,\beta)\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu\,;$$

recall that $\widetilde{\kappa},\kappa\in(0,1]$. Then (4.33) and (4.78) give, after some elementary algebraic manipulations, the estimate

$$
\begin{aligned}
(4.79)\quad \int_{\mathcal{B}(x_0,\beta)\cap\{H>\widetilde{\kappa}\kappa\lambda\}} H^{p'}\,\mathrm{d}\nu &\le \frac{C(\mathtt{data})}{\varepsilon^{2-2\gamma/p'}(\widetilde{\kappa}\kappa)^{p'-\gamma}}(\widetilde{\kappa}\kappa\lambda)^{p'-\gamma}\int_{\mathcal{B}(x_0,\alpha)\cap\{H>\widetilde{\kappa}\kappa\lambda\}} H^{\gamma}\,\mathrm{d}\nu \\
&\quad + \frac{C_5\lambda_1^{\vartheta_f}}{\widehat{\kappa}^{p'}(\widehat{\kappa}\kappa\lambda)^{\widetilde{\vartheta}_f-p'}}\int_{\mathcal{B}(x_0,\alpha)\cap\{F>\widehat{\kappa}\kappa\lambda\}} F^{p_*}\,\mathrm{d}\nu\,.
\end{aligned}
$$

Recall that C_5 is defined in (4.32). We can reformulate this estimate as

$$
\begin{aligned}
(4.80)\quad \int_{\mathcal{B}(x_0,\beta)\cap\{H>\lambda\}} H^{p'}\,\mathrm{d}\nu &\le \frac{C}{\varepsilon^{2-2\gamma/p'}(\widetilde{\kappa}\kappa)^{p'-\gamma}}\lambda^{p'-\gamma}\int_{\mathcal{B}(x_0,\alpha)\cap\{H>\lambda\}} H^{\gamma}\,\mathrm{d}\nu \\
&\quad + \frac{C_6\lambda_1^{\vartheta_f}}{\lambda^{\widetilde{\vartheta}_f-p'}}\int_{\mathcal{B}(x_0,\alpha)\cap\{F>\widehat{\kappa}\lambda/\widetilde{\kappa}\}} F^{p_*}\,\mathrm{d}\nu\,.
\end{aligned}
$$

The constants C and C_6 satisfy the following dependencies:

$$C\equiv C(\mathtt{data})\,,\qquad C_6\equiv C_6(\mathtt{data},\varepsilon)\,.$$

Since (4.79) holds for all $\lambda\ge\lambda_2$ where λ_2 has been defined in (4.43) we have that (4.80) holds for all $\lambda\ge\widetilde{\kappa}\kappa\lambda_2$. Recall again that $\widetilde{\kappa},\widehat{\kappa},\kappa\in(0,1]$ have been defined in (4.25), (4.30) and (4.64) respectively.

In order to conclude with the level set estimate (4.5) we need to estimate several constants. We need to obtain the specific dependence on ε of the constant appearing in front of the first integrand on the right-hand side. By using (4.25) and (4.64), we see that we can choose $\widetilde{\kappa}$, κ to satisfy

$$\widetilde{\kappa}\kappa = \frac{\varepsilon^{3/\gamma - 2/p'}}{C} \,,$$

where $C \equiv C(\texttt{data})$. With this choice, we can estimate the constant appearing in front of the second integral in (4.80) and therefore arrive at a choice of ϑ exactly as in (4.4). We next set

$$\kappa_f := \widehat{\kappa}/\widetilde{\kappa} \,,$$

and making note of (4.30) we can additionally choose $\widehat{\kappa}$ small so that $\kappa_f \in (0,1)$.

Last, we need to find an upper bound on the numbers λ_1 and λ_2 defined in (4.7) and (4.43) respectively, so that the level set estimate (4.5) can be verified for the range prescribed by (4.6). If $x \in B(x_0, \beta)$ and $\frac{\alpha-\beta}{40^n} \leq R \leq \frac{\varrho_0}{2}$, then $\mathcal{B}(x,R) \subset \mathcal{B}(x_0, 2\varrho_0)$. Recalling the doubling property (3.2), whenever $\widetilde{H}$ is a ν-integrable function we can estimate

$$\fint_{\mathcal{B}(x,R)} \widetilde{H} \, \mathrm{d}\nu \leq \frac{\nu(\mathcal{B}(x_0, 2\varrho_0))}{\nu(\mathcal{B}(x,R))} \fint_{\mathcal{B}(x_0, 2\varrho_0)} \widetilde{H} \, \mathrm{d}\nu \leq C \left(\frac{\varrho_0}{\alpha - \beta} \right)^{n+\varepsilon p} \fint_{\mathcal{B}(x_0, 2\varrho_0)} \widetilde{H} \, \mathrm{d}\nu$$

where $C = C(n)$. Applying this inequality to $H^{p'}$, F^{p_*}, H^{γ}, and $F^{p_* + \delta_f}$, as well as on different balls $2^k \mathcal{B}(x,R) \subset 2^k \mathcal{B}(x_0, 2\varrho_0)$, we get

(4.81)
$$\kappa^{-1} \big\{ \Psi_M(x,R) + \Upsilon_0(x,R) + Tail(x,R) \big\}$$
$$\leq \frac{C}{\varepsilon^{1/\gamma}} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{n+\varepsilon p} \big\{ \Psi_M(x_0, 2\varrho_0) + \Upsilon_0(x_0, 2\varrho_0) + Tail(x_0, 2\varrho_0) \big\}$$
$$\leq \frac{C}{\varepsilon^{1/\gamma}} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{n+\varepsilon p} \Theta(x_0, 2\varrho_0) \leq \frac{C}{\varepsilon} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{2n+p} \Theta(x_0, 2\varrho_0) \,,$$

where $C \equiv C(\texttt{data})$. We also used (4.64) to remove dependence on κ, (4.22) to remove dependence on M, that $\varepsilon < 1$, and that $\frac{\varrho_0}{\alpha - \beta} \geq 2$. Recall also that Θ has been defined in (4.1). Thus we have obtained the desired upper bound on λ_1. To estimate λ_2, note that for $\mathcal{K} = K_1 \times K_2 \in \Delta_{k_0}$ with k_0 as in (4.34) we have $\mathcal{K} \subset \mathcal{B}(x_0, \alpha) \subset \mathcal{B}(x_0, 2\varrho_0)$ and therefore

$$\nu(\mathcal{K}) \geq \frac{C}{\varrho_0^{n+\varepsilon p}} \int_{K_1} \int_{K_2} \mathrm{d}y \, \mathrm{d}x = \frac{C(\alpha - \beta)^{2n}}{\varrho_0^{n+\varepsilon p}} \,, \qquad C \equiv C(n,p) \,.$$

Thus, for any cube $\mathcal{K} \in \Delta_{k_0}$ we can estimate
(4.82)
$$\left(\fint_{\mathcal{K}} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \leq \left(\frac{\nu(\mathcal{B}(x_0, 2\varrho_0))}{\nu(\mathcal{K})} \fint_{\mathcal{B}(x_0, 2\varrho_0)} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} \leq \frac{C}{\varepsilon^{1/p'}} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{2n/p'} \left(\fint_{\mathcal{B}(x_0, 2\varrho_0)} H^{p'} \, \mathrm{d}\nu \right)^{1/p'} .$$

Then using (4.81) and (4.82) (also using that $\varepsilon < 1$ and $\frac{\varrho_0}{\alpha - \beta} \geq 2$) we get that

$$\lambda_2 \leq \frac{C}{\varepsilon} \left(\frac{\varrho_0}{\alpha - \beta} \right)^{2n+p} \Theta(x_0, 2\varrho_0) \,,$$

where $C \equiv C(\texttt{data})$. We therefore take λ_0 as in (4.6) so that $\lambda_0 \geq \max\{\lambda_1, \lambda_2\}$. We finally arrive at (4.5) with λ prescribed as in (4.6), and the proof of Proposition 4.1 is complete.

Remark 4.11. Note that in the proof of Proposition 4.1 the off-diagonal analysis in Section 4.3 does not make use of the assumption (3.9). Thus, the conclusions of Lemma 4.10 hold for general fractional Sobolev functions u, and not just solutions to (1.1).

References

[1] Richard F Bass and Moritz Kassmann. Hölder continuity of harmonic functions with respect to operators of variable order. *Communications in Partial Differential Equations*, 30(8):1249–1259, 2005.

[2] Cristiana De Filippis and Giampiero Palatucci. Hölder regularity for nonlocal double phase equations. *Journal of Differential Equations*, 267(1):547–586, 2019.

[3] Tuomo Kuusi, Giuseppe Mingione, and Yannick Sire. A fractional Gehring lemma, with applications to nonlocal equations. *Rendiconti Lincei-Matematica e Applicazioni*, 25(4):345–358, 2014.

[4] Tuomo Kuusi, Giuseppe Mingione, and Yannick Sire. Nonlocal self-improving properties. *Anal. PDE*, 8(1):57–114, 2015.

[5] Tadele Mengesha and James M. Scott. A fractional korn-type inequality for smooth domains and a regularity estimate for nonlinear nonlocal systems of equations, 2020.

[6] Tadele Mengesha and James M. Scott. Self-improving inequalities for bounded weak solutions to nonlocal double-phase equations. https://arxiv.org/abs/2011.11466, 2020.

[7] Giuseppe Mingione. The singular set of solutions to non-differentiable elliptic systems. *Archive for rational mechanics and analysis*, 166(4):287–301, 2003.

[8] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[9] Simon Nowak, Regularity theory for nonlocal equations with VMO coefficients. *To appear in Ann. Inst. H. Poincaré Anal. Non Linéaire, arXiv:2101.11690.*

[10] Simon Nowak, Improved S obolev regularity for linear nonlocal equations with VMO coefficients *To appear in Mathematische Annalen, 2022, DOI 10.1007/s00208-022-02369-w*

[11] Armin Schikorra. Nonlinear commutators for the fractional p-Laplacian and applications. *Mathematische Annalen*, 366(1-2):695–720, 2016.

[12] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

(Tadele Mengesha) Department of Mathematics, University of Tennessee Knoxville, mengesha@utk.edu.

(James M. Scott) Department of Applied Physics and Applied Mathematics, Columbia University.