DE GRUYTER Advances in Nonlinear Analysis 2022; 11: 1307–1331

Research Article

Jamil Chaker and Minhyun Kim*

Regularity estimates for fractional orthotropic p-Laplacians of mixed order

https://doi.org/10.1515/anona-2022-0243
received April 27, 2021; accepted February 4, 2022

Abstract: We study robust regularity estimates for a class of nonlinear integro-differential operators with anisotropic and singular kernels. In this paper, we prove a Sobolev-type inequality, a weak Harnack inequality, and a local Hölder estimate.

Keywords: nonlocal operators, divergence form, regularity theory, anisotropic measures, weak Harnack inequality

MSC 2020: 35B65, 47G20, 31B05, 42B25

1 Introduction

In this paper, we investigate regularity estimates for weak solutions to nonlocal equations

$$\mathrm{L}u = f \quad \text{in } Q = (-1, 1)^n, \tag{1.1}$$

where L is a nonlinear integro-differential operator of the form

$$\mathrm{L}u(x) = \mathrm{PV} \int_{\mathbb{R}^n} |u(y) - u(x)|^{p-2}(u(y) - u(x))\mu(x, \mathrm{d}y) \tag{1.2}$$

for $p > 1$ and $f \in L^q(Q)$ for some sufficiently large q. The operator L is clearly determined by the family of measures $(\mu(x, \mathrm{d}y))_{x\in\mathbb{R}^n}$. In the special case, when L is the generator of a Lévy process, $\mu(x, A)$ measures the number of expected jumps from x into the set A within the unit time interval. However, the class of operators that we consider in this paper is more involved, and for that reason, we first take a look at an important example. Let $n \in \mathbb{N}$. For $s_1, \dots, s_n \in (0, 1)$, we define

$$\mu_{\mathrm{axes}}(x, \mathrm{d}y) = \sum_{k=1}^{n} s_k(1 - s_k)|x_k - y_k|^{-1-s_k p}\mathrm{d}y_k \prod_{i\neq k}\delta_{x_i}(\mathrm{d}y_i).$$

This family plays a central role in our paper since admissible operators resp. families of measures will be defined on the basis of μ_{axes}. Given $x \in \mathbb{R}^n$, the measure $\mu_{\mathrm{axes}}(x, \cdot)$ only charges differences that occur along the axes

$$\{x + te_k \mid t \in \mathbb{R}\} \quad \text{for } k \in \{1, \dots, n\}.$$

Hence, we can think of the operator Lu for $\mu(x, \cdot) = \mu_{\mathrm{axes}}(x, \cdot)$ as the sum of one-dimensional fractional p-Laplacian in $\mathbb{R}^n$ with orders of differentiability $s_1, \dots, s_n \in (0, 1)$ depending on the respective direction. In particular, $\mu_{\mathrm{axes}}(x, \cdot)$ does not possess a density with respect to the Lebesgue measure. An interesting

*** Corresponding author: Minhyun Kim,** Fakultät für Mathematik, Universität Bielefeld, 33615 Bielefeld, Germany, e-mail: minhyun.kim@uni-bielefeld.de
Jamil Chaker: Fakultät für Mathematik, Universität Bielefeld, 33615 Bielefeld, Germany, e-mail: jchaker@math.uni-bielefeld.de

phenomenon for the case $p = 2$ and $s = s_1 = \dots = s_n$ is that on the one hand the corresponding energies for the fractional Laplacian and L are comparable. On the other hand (for sufficiently good functions), both operators converge to the Laplace operator as $s \nearrow 1$. It is known that the fractional p-Laplacian converges to the p-Laplacian (see [11, Theorem 2.8] or [20, Lemma 5.1] for details), which is defined by

$$\Delta_p u(x) = \operatorname{div}(|\nabla u(x)|^{p-2}\nabla u(x)).$$

However, the operator L for $\mu(x,\cdot) = \mu_{\text{axes}}(x,\cdot)$ converges for any $p > 1$ and $s = s_1 = \dots = s_n$ to the following local operator (up to a constant depending on p only)

$$A^p_{\text{loc}} u(x) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\left|\frac{\partial u(x)}{\partial x_i}\right|^{p-2} \frac{\partial u(x)}{\partial x_i}\right) = \operatorname{div}(a(\nabla u(x))) \tag{1.3}$$

as $s \nearrow 1$, where $a : \mathbb{R}^n \to \mathbb{R}^n$ with $a(z) = (|z_i|^{p-2} z_i)_{i\in\{1,\dots,n\}}$. This convergence is a direct consequence of the convergence for the one-dimensional fractional p-Laplacian and the summation structure of the operator for μ_{axes}. For details, we refer the reader to Proposition D.1. The operator A^p_{loc} is known as orthotropic p-Laplacian and is a well-known operator in the analysis (see, for instance, [41, Chapter 1, Section 8]). This operator is sometimes also called pseudo p-Laplacian. Minimizers for the corresponding energies have been studied in [6], where the authors prove Hölder continuity of minimizers. In [9], local Lipschitz regularity for weak solutions to orthotropic p-Laplace equations for $p \geq 2$ and every dimension is proved. The case, when p is allowed to be different in each direction, is also studied in several papers. For instance, in [3], the authors introduce anisotropic De Giorgi classes and study related problems. Another interesting paper studying such operators with nonstandard growth condition is [8], where the authors show that bounded local minimizers are locally Lipschitz continuous. For further results, we refer the reader to the references given in the previously mentioned papers.

The two local operators Δ_p and A^p_{loc} are substantially different, as Δ_p is invariant under orthogonal transformation, while A^p_{loc} is not. One strength of our results is that they are robust, and we can recover results for the orthotropic p-Laplacian by taking the limit.

One way to deal with the anisotropy of μ_{axes} is to consider for given $s_1, \dots, s_n \in (0,1)$ a class of suitable rectangles instead of cubes or balls. Hence, we define $s_{\max} = \max\{s_1, \dots, s_n\}$.

Definition 1.1. For $r > 0$ and $x \in \mathbb{R}^n$, we define

$$M_r(x) = \mathop{\times}_{k=1}^{n} \left(x_k - r^{\frac{s_{\max}}{s_k}}, x_k + r^{\frac{s_{\max}}{s_k}}\right) \quad \text{and} \quad M_r = M_r(0).$$

The advantage of taking these cubes is that they take the anisotropy of the measures resp. operators into account and the underlying metric measure space is a doubling space. The choice of $s_{\max}$ in the definition of $M_r(x)$ is not important. It can be replaced by any positive number $\varsigma \geq s_{\max}$. We only need to ensure that $M_r(x)$ are balls in a metric measure space with radius $r > 0$ and center $x \in \mathbb{R}^n$. This allows us to use known results on doubling spaces like the John-Nirenberg inequality or results on the Hardy-Littlewood maximal function.

In the spirit of [15], for each $k \in \{1, \dots, n\}$, we define $E^k_r(x) = \{y \in \mathbb{R}^n : |x_k - y_k| < r^{s_{\max}/s_k}\}$. Note, that

$$M_r(x) = \bigcap_{k=1}^{n} E^k_r(x). \tag{1.4}$$

We consider families of measures $\mu(x, \mathrm{d}y)$, which are given through certain properties regarding the reference family $\mu_{\text{axes}}(x, \mathrm{d}y)$. Let us introduce and briefly discuss our assumptions on the families $(\mu(x,\cdot))_{x\in\mathbb{R}^n}$.

Assumption 1. We assume

$$\sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} (|x-y|^p \wedge 1)\mu(x,\mathrm{d}y) < \infty,$$

and for all sets $A, B \in \mathcal{B}(\mathbb{R}^n)$:

$$\int_A \int_B \mu(x, \mathrm{d}y)\mathrm{d}x = \int_B \int_A \mu(x, \mathrm{d}y)\mathrm{d}x.$$

Assumption 1 provides integrability and symmetry of the family of measures. Furthermore, we assume the following tail behavior of $(\mu(x, \cdot))_{x\in\mathbb{R}^n}$.

Assumption 2. There is $\Lambda \geq 1$ such that for every $x_0 \in \mathbb{R}^n$, $k \in \{1, \dots, n\}$ and all $r > 0$

$$\mu(x_0, \mathbb{R}^n \setminus E_r^k(x_0)) \leq \Lambda(1 - s_k)r^{-ps_{\max}}.$$

Note that Assumption 2 is a stronger assumption than an assumption on the volume on the complement of every $M_r(x_0)$. It gives an appropriate tail behavior for the family of measures in each direction separately and allows us to control the appearing constants in our tail estimate in all directions. This is necessary to prove robust estimates for the corresponding operators.

Note that by Assumption 2 and (1.4), we have

$$\mu(x_0, \mathbb{R}^n \setminus M_\rho(x_0)) \leq \sum_{k=1}^n \mu(x_0, \mathbb{R}^n \setminus E_\rho^k(x_0)) \leq \Lambda \sum_{k=1}^n (1 - s_k)\rho^{-ps_{\max}} \leq \Lambda n \rho^{-ps_{\max}}. \tag{1.5}$$

Hence, (1.5) shows that Assumption 2 implies $\mu(x_0, \mathbb{R}^n \setminus M_\rho(x_0)) \leq c\mu_{\text{axes}}(x_0, \mathbb{R}^n \setminus M_\rho(x_0))$ for all $x_0 \in \mathbb{R}^n$.

Finally, we assume the local comparability of corresponding functionals. Hence, we define for any open and bounded $\Omega \subset \mathbb{R}^n$

$$\mathcal{E}_\Omega^\mu(u, v) = \int_\Omega \int_\Omega |u(y) - u(x)|^{p-2}(u(y) - u(x))(v(y) - v(x))\mu(x, \mathrm{d}y)\mathrm{d}x$$

and $\mathcal{E}^\mu(u, v) = \mathcal{E}_{\mathbb{R}^n}^\mu(u, v)$ whenever these quantities are finite.

Assumption 3. There is $\Lambda \geq 1$ such that for every $x_0 \in \mathbb{R}^n$, $\rho \in (0, 3)$ and every $u \in L^p(M_\rho(x_0))$:

$$\Lambda^{-1}\mathcal{E}_{M_\rho(x_0)}^\mu(u, u) \leq \mathcal{E}_{M_\rho(x_0)}^{\mu_{\text{axes}}}(u, u) \leq \Lambda \mathcal{E}_{M_\rho(x_0)}^\mu(u, u). \tag{1.6}$$

Local comparability of the functionals is an essential assumption on the family of measures. It tells us that our family of measures can vary from our reference family in the given sense of local functionals without losing crucial information on $(\mu(x, \cdot))_{x\in\mathbb{R}^n}$ like functional inequalities, which we deduce for the explicitly known family $(\mu_{\text{axes}}(x, \cdot))_{x\in\mathbb{R}^n}$. This assumption allows us, for instance, to study operators of the form (1.2) for μ_{axes} in the general framework of bounded and measurable coefficients. We emphasize that further examples of families of measures satisfying (1.6) can be constructed similarly to the case $p = 2$ (see [15, Section 9]). In this paper, we study nonlocal operators of the form (1.2) for families of measures that satisfy the previously given assumptions.

Definition 1.2. Let $p > 1$, $\Lambda \geq 1$, and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0, 1)$. We call a family of measures $(\mu(x, \cdot))_{x\in\mathbb{R}^n}$ admissible with regard to $(\mu_{\text{axes}}(x, \cdot))_{x\in\mathbb{R}^n}$, if it satisfies Assumptions 1–3. We denote the class of such measures by $\mathcal{K}(p, s_0, \Lambda)$.

It is not hard to see that the family $(\mu_{\text{axes}}(x, \cdot))_{x\in\mathbb{R}^n}$ is admissible in the aforementioned sense. Note that Assumptions 1 and 3 are clearly satisfied. Furthermore, for every $x_0 \in \mathbb{R}^n$, $k \in \{1, \dots, n\}$ and all $r > 0$,

$$\mu_{\text{axes}}(x_0, \mathbb{R}^n \setminus E_r^k(x_0)) = 2s_k(1 - s_k) \int_{r^{s_{\max}/s_k}}^\infty h^{-1-s_k p} = \frac{2(1 - s_k)}{p} r^{-s_{\max} p},$$

which shows Assumption 2 for $\Lambda = \frac{2}{p}$.

The purpose of this paper is to study weak solutions to nonlocal equations governed by the class of operators L as in (1.2). In order to study weak solutions, we need appropriate Sobolev-type function spaces that guarantee regularity and integrability with respect to μ.

Definition 1.3. Let $\Omega \subset \mathbb{R}^n$ open and $p > 1$. We define the function spaces

$$V^{p,\mu}(\Omega|\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} \text{ meas. } |u|_\Omega \in L^p(\Omega), (u,u)_{V^{p,\mu}(\Omega|\mathbb{R}^n)} < \infty\},$$
$$H^{p,\mu}_\Omega(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} \text{ meas. } |u \equiv 0 \text{ on } \mathbb{R}^n \setminus \Omega, \|u\|_{H^{p,\mu}_\Omega(\mathbb{R}^n)} < \infty\},$$

where

$$(u,v)_{V^{p,\mu}(\Omega|\mathbb{R}^n)} = \int_\Omega \int_{\mathbb{R}^n} |u(y)-u(x)|^{p-2}(u(y)-u(x))(v(y)-v(x))\mu(x,\mathrm{d}y)\mathrm{d}x,$$
$$\|u\|^p_{H^{p,\mu}_\Omega(\mathbb{R}^n)} = \|u\|^p_{L^p(\Omega)} + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |u(y)-u(x)|^p \mu(x,\mathrm{d}y)\mathrm{d}x.$$

The space $V^{p,\mu}(\Omega|\mathbb{R}^n)$ can be seen as a nonlocal analog of the space $H^{1,p}(\Omega)$. It provides fractional regularity (measured in terms of μ) inside of Ω and integrability on $\mathbb{R}^n \setminus \Omega$. The space $V^{p,\mu}(\Omega|\mathbb{R}^n)$ will serve as solution space. On the other hand, the space $H^{p,\mu}_\Omega(\mathbb{R}^n)$ can be seen as a nonlocal analog of $H^{1,p}_0(\Omega)$. See [28] and [24] for further studies of these spaces in the case $p = 2$.

We are interested in finding robust regularity estimates for weak solutions to a class of nonlocal equations. This means that the constants in the regularity estimates do not depend on the orders of differentiability of the integro-differential operator itself but only on a lower bound of the orders. Let us formulate the main results of this paper. For this purpose, we define $\bar{s}$ to be the harmonic mean of the orders $s_1, \dots, s_n$, that is,

$$\bar{s} = \left(\frac{1}{n}\sum_{k=1}^n \frac{1}{s_k}\right)^{-1}.$$

It is well known that the Harnack inequality fails for weak solutions to singular equations of the type (1.1). Even in the case $p = 2$ and $s_1 = \dots = s_n$, the Harnack inequality does not hold (see, for instance, [4,7]). Our first main result is a weak Harnack inequality for weak supersolutions to equations of the type (1.1). Throughout the paper, we denote by $p_\star = np/(n - p\bar{s})$ the Sobolev exponent, which will appear in Theorem 2.7.

Theorem 1.4. (Weak Harnack inequality) *Let $\Lambda \geq 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Let $1 < p < n/\bar{s}$ and $f \in L^{q/(p\bar{s})}(M_1)$ for some $q > n$. There are $p_0 = p_0(n, p, p_\star, s_0, q, \Lambda) \in (0,1)$ and $C = C(n, p, p_\star, s_0, q, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $u \in V^{p,\mu}(M_1|\mathbb{R}^n)$ satisfying $u \geq 0$ in M_1 and*

$$\mathcal{E}^\mu(u,\varphi) \geq (f,\varphi) \quad \text{for every non-negative } \varphi \in H^{p,\mu}_{M_1}(\mathbb{R}^n),$$

the following holds:

$$\inf_{M_{1/4}} u \geq C\left(\fint_{M_{1/2}} u^{p_0}(x)\mathrm{d}x\right)^{1/p_0} - \sup_{x\in M_{15/16}} 2\left(\int_{\mathbb{R}^n\setminus M_1} u^-(z)^{p-1}\mu(x,\mathrm{d}z)\right)^{1/(p-1)} - \|f\|_{L^{q/(p\bar{s})}(M_{15/16})}. \tag{1.7}$$

Although the weak Harnack inequality provides an estimate on the infimum only, it is sufficient to prove a decay of oscillation for bounded weak solutions and therefore a local Hölder estimate.

Theorem 1.5. (Local Hölder estimate) *Let $\Lambda \geq 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Let $1 < p < n/\bar{s}$ and $f \in L^{q/(p\bar{s})}(M_1)$ for some $q > n$. There are $\alpha = \alpha(n, p, p_\star, s_0, q, \Lambda) \in (0,1)$ and $C = C(n, p, p_\star, s_0, q, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $u \in V^{p,\mu}(M_1|\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ satisfying*

$$\mathcal{E}^\mu(u,\varphi) = (f,\varphi) \quad \text{for every } \varphi \in H^{p,\mu}_{M_1}(\mathbb{R}^n),$$

the following holds: $u \in C^{\alpha}(\overline{M}_{1/2})$ *and*

$$\|u\|_{C^{\alpha}(\overline{M}_{1/2})} \leq C(\|u\|_{L^{\infty}(\mathbb{R}^n)} + \|f\|_{L^{q/(p\bar{s})}(M_{15/16})}).$$

Note that the result needs global boundedness of weak solutions. The same assumption is also needed in the previous works [14,15,25]. Global and local boundedness of weak solutions to anisotropic nonlocal equations are nontrivial open questions.

Furthermore, note that the general case, replacing $M_{\frac{1}{4}}$, $M_{\frac{1}{2}}$, $M_{\frac{15}{16}}$, M_1 by $M_{\frac{r}{4}}(x_0)$, $M_{\frac{r}{2}}(x_0)$, $M_{\frac{15r}{16}}(x_0)$,$M_r(x_0)$, for $x_0 \in \mathbb{R}^n$ and $r \in (0,1]$ follows by a translation and anisotropic scaling argument introduced in Section 4. See also [14].

Let us comment on related results in the literature. The underlying ideas in developing regularity results for uniformly elliptic operators in divergence form with bounded and measurable coefficients go back to the influential contributions by De Giorgi, Nash, and Moser (see [19,45,46]). These works led to many further results in various directions. Similar results for nonlocal operators in divergence form have been obtained by several authors including these works: [1, 2,5,12,16–18,22,24,26,27, 33,34,38,42–44,54,55]. See also the references therein. For further regularity results concerning nonlocal equations governed by fractional p-Laplacians, we refer the reader to [10,38,39,48–51].

In [22], the authors extend the De Giorgi-Nash-Moser theory to a class of fractional p-Laplace equations. They provide the existence of a unique minimizer to homogenous equations and prove local regularity estimates for weak solutions. Moreover, in [21], the same authors prove a general Harnack inequality for weak solutions.

Nonlocal operators with anisotropic and singular kernels of the type μ_{axes} are studied in various mathematical areas such as stochastic differential equations and potential theory. In [4], the authors study regularity estimates for harmonic functions for systems of stochastic differential equations $\mathrm{d}X_t = A(X_{t-})\mathrm{d}Z_t$ driven by Lévy processes Z_t with Lévy measure $\mu_{\text{axes}}(0, \mathrm{d}y)$, where $2s_1 = \cdots = 2s_n = \alpha$ and $p = 2$. See also [13,56,52]. Sharp two-sided bounds for the heat kernels are established in [35,37]. In [36], the authors prove the existence of transition density of the process X_t and establish semigroup properties of solutions. The existence of densities for solutions to stochastic differential equations with Hölder continuous coefficients driven by Lévy processes with anisotropic jumps has been proved in [30]. Such types of anisotropies also appear in the study of the anisotropic stable JCIR process, see [29].

Our approach follows mainly the ideas of [14,25] and [15]. In [25], the authors develop a local regularity theory for a class of linear nonlocal operators, which covers the case $s = s_1 = \cdots = s_n \in (0,1)$ and $p = 2$. Based on the ideas of [25], the authors in [14] establish regularity estimates in the case $p = 2$ for weak solutions in a more general framework, which allows the orders of differentiability $s_1, \ldots, s_n$ to be different. In [15], parabolic equations in the case $p = 2$ and possible different orders of differentiability are studied. That paper provides robust regularity estimates, which means the constants in the weak Harnack inequality and Hölder regularity estimate do not depend on the orders of differentiability but on their lower one, only. This allows us to recover regularity results for local operators from the theory of nonlocal operators by considering the limit.

The purpose of this paper is to provide local regularity estimates as in [14] for operators which are allowed to be nonlinear. This nonlinearity leads to several difficulties like the need for a different proof for the discrete gradient estimate (see Lemma 3.4). Since we cannot use the helpful properties of Hilbert spaces (like Plancherel's theorem), we also need an approach different from the one in [14] to prove the Sobolev-type inequality. One strength of this paper is the robustness of all results. This allows us to recover regularity estimates for the limit operators such as for the orthotropic p-Laplacian.

Finally, we would like to point out that it is also interesting to study such operators in nondivergence form. We refer the reader to [53] for regularity results concerning the fractional Laplacian and to [40] for the fractional p-Laplacian. See also [23] for the anisotropic case.

Even in the most simple case, that is, $p = 2$ and $s = s_1 = s_2 = \cdots = s_n$, regularity estimates for operators in nondivergence form of the type (1.2) with $\mu = \mu_{\text{axes}}$ lead to various open problems such as an Alexandrov-Bakelmann-Pucci estimate.

The authors wish to express their thanks to Lorenzo Brasco for helpful comments.

Outline: This paper is organized as follows. In Section 2, we introduce appropriate cut-off functions and prove auxiliary results concerning functionals for admissible families of measures. One main result of that section is the Sobolev-type inequality (see Theorem 2.7). In Section 3, we prove the weak Harnack inequality, and Section 4 presents the proof of the local Hölder estimate. In Appendix A, we prove some auxiliary algebraic inequalities, and in Appendix B, we briefly sketch the construction of appropriate anisotropic "dyadic" rectangles. In Appendix C, we use the anisotropic dyadic rectangles to sketch the proof of a suitable sharp maximal function theorem.

2 Auxiliary results

This section is devoted to providing some general properties for the class of nonlocal operators that we study in the scope of this paper. The main auxiliary result is a robust Sobolev-type inequality.

Let us first introduce a class of suitable cut-off functions that will be useful for appropriate localization.

Definition 2.1. We say that $(\tau_{x_0,r,\lambda})_{x_0,r,\lambda} \subset C^{0,1}(\mathbb{R}^n)$ is an admissible family of cut-off functions if there is $c \geq 1$ such that for all $x_0 \in \mathbb{R}^n$, $r \in (0,1]$ and $\lambda \in (1,2]$, it holds that

$$\begin{cases} \operatorname{supp}(\tau) \subset M_{\lambda r}(x_0), \\ \|\tau\|_\infty \leq 1, \\ \tau \equiv 1 \text{ on } M_r(x_0), \\ \|\partial_k \tau\|_\infty \leq c(\lambda^{s_{\max}/s_k} - 1)^{-1} r^{-s_{\max}/s_k} \text{ for every } k \in \{1 \dots n\}. \end{cases}$$

For brevity, we simply write τ for any such function from $(\tau_{x_0,r,\lambda})_{x_0,r,\lambda}$, if the respective choice of x_0, r and λ is clear. The existence of such functions is standard.

Recall the definition of admissible families of measures $\mathcal{K}(p, s_0, \Lambda)$ from Definition 1.2.

Lemma 2.2. *Let $p > 1$, $\Lambda \geq 1$, and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. There is $C = C(n, p, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$, every $x_0 \in \mathbb{R}^n$, $r \in (0,1]$, $\lambda \in (1,2]$ and every admissible cut-off function τ, the following is true*:

$$\sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} |\tau(y) - \tau(x)|^p \mu(x, \mathrm{d}y) \leq C\left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-ps_k}\right) r^{-ps_{\max}}.$$

Proof. We skip the proof. One can follow the lines of the proof from [15, Lemma 3.1] and will get the same result with the factor n^{p-1} instead of n. □

For future purposes, we deduce the following observation. It is an immediate consequence of the foregoing lemma.

Corollary 2.3. *Let $p > 1$, $\Lambda \geq 1$, and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. There is a constant $C = C(n, p, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $x_0 \in \mathbb{R}^n$, $r \in (0,1]$, $\lambda \in (1,2]$ and every admissible cut-off function τ and every $u \in L^p(M_{\lambda r}(x_0))$, it holds true that*

$$\int_{M_{\lambda r}(x_0)} \int_{\mathbb{R}^n \setminus M_{\lambda r}(x_0)} |u(x)|^p |\tau(x)|^p \mu(x, \mathrm{d}y) \mathrm{d}x \leq C\left(\sum_{k=1}^n \left(\lambda^{s_{\max}/s_k} - 1\right)^{-ps_k}\right) r^{-ps_{\max}} \|u\|^p_{L^p(M_{\lambda r}(x_0))}.$$

Note that the constants in Lemma 2.2 and Corollary 2.3 do not depend on s_0. Therefore, the lower bound $s_0 \leq s_k$ for all $k \in \{1, \dots, n\}$ can be dropped here.

2.1 Functional inequalities

This section is devoted to the proofs of a Sobolev and a Poincaré-type inequality. We start our analysis by first proving a technical lemma, see also [15, Lemma 4.1].

Lemma 2.4. *Let $p > 1$, $a \in (0,1]$, $b \geq 1$, $N \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, and $s_k \in (0,1)$. For any $u \in L^p(\mathbb{R}^n)$*

$$\int_{\mathbb{R}^n} \sup_{\rho>0} \frac{1}{\rho^{(1+ps_k)b}} \int_{\mathbb{R}} |u(x) - u(x+he_k)|^p \mathbf{1}_{[a\rho^b, 2a\rho^b)}(|h|) \mathrm{d}h \mathrm{d}x$$
$$\leq (2a)^{1+ps_k} N^{p(1-s_k)} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \frac{|u(x) - u(x+he_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{\left[\frac{a}{N}\rho^b, \frac{2a}{N}\rho^b\right)}(|h|) \mathrm{d}h \mathrm{d}x.$$

Proof. Let $I_a = [a\rho^b, 2a\rho^b)$. By the triangle inequality and a simple change of variables, we have

$$\begin{aligned}
&\int_{\mathbb{R}} |u(x) - u(x+he_k)|^p \mathbf{1}_{I_a}(|h|) \mathrm{d}h \\
&\quad \leq N^{p-1} \sum_{j=1}^{N} \int_{\mathbb{R}} \left| u\left(x + \frac{j-1}{N} he_k\right) - u\left(x + \frac{j}{N} he_k\right)\right|^p \mathbf{1}_{I_a}(|h|) \mathrm{d}h \\
&\quad = N^p \sum_{j=1}^{N} \int_{\mathbb{R}} |u(x + (j-1)he_k) - u(x + jhe_k)|^p \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}h.
\end{aligned}$$

Since $|h| < \frac{2a}{N}\rho^b$, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n} \sup_{\rho>0} \frac{1}{\rho^{(1+ps_k)b}} \int_{\mathbb{R}} |u(x) - u(x+he_k)|^p \mathbf{1}_{I_a}(|h|) \mathrm{d}h \mathrm{d}x \\
&\quad \leq N^p \left(\frac{2a}{N}\right)^{1+ps_k} \sum_{j=1}^{N} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \frac{|u(x+(j-1)he_k) - u(x+jhe_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}h \mathrm{d}x.
\end{aligned}$$

We change the order of integration by Fubini's theorem and then use the change of variables $y = x + (j-1)he_k$ to conclude that

$$\begin{aligned}
&\sum_{j=1}^{N} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \frac{|u(x+(j-1)he_k) - u(x+jhe_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}h \mathrm{d}x \\
&\quad = \sum_{j=1}^{N} \int_{\mathbb{R}} \int_{\mathbb{R}^n} \frac{|u(x+(j-1)he_k) - u(x+jhe_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}x \mathrm{d}h \\
&\quad = \sum_{j=1}^{N} \int_{\mathbb{R}} \int_{\mathbb{R}^n} \frac{|u(y) - u(y+he_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}y \mathrm{d}h \\
&\quad = N \int_{\mathbb{R}^n} \int_{\mathbb{R}} \frac{|u(y) - u(y+he_k)|^p}{|h|^{1+ps_k}} \mathbf{1}_{I_{a/N}}(|h|) \mathrm{d}h \mathrm{d}y.
\end{aligned}$$ □

Using the foregoing result for $b = s_{\max}/s_k$ allows us to prove a robust Sobolev-type inequality. Robust in this context means that the appearing constant in the Sobolev-type inequality is independent of $s_1, \ldots, s_n$ and depends on the lower bound s_0 only.

Before we prove a robust Sobolev-type inequality, we recall the definition of the Hardy–Littlewood maximal function and sharp maximal function. For $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$,

$$\mathbf{M}u(x) = \sup_{\rho>0} \fint_{M_\rho(x)} u(y)\mathrm{d}y \quad \text{and} \quad \mathbf{M}^\sharp u(x) = \sup_{\rho>0} \fint_{M_\rho(x)} |u(y) - (u)_{M_\rho(x)}|\mathrm{d}y,$$

where $(u)_\Omega = \fint_\Omega u(z)\mathrm{d}z$. We will use the maximal function theorem and the sharp maximal function theorem. Note that $\mathbb{R}^n$ is equipped with the metric induced by rectangles of the form $M_r(x)$ and the standard Lebesgue measure. Since $|M_{2r}| = 2^n(2r)^{ns_{\max}/\bar{s}} \le 2^{n/s_0}|M_r|$, this space is a doubling space with the doubling constant $2^{n/s_0}$.

Theorem 2.5. [32, Theorem 2.2] *Let $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Then, there is a constant $C_1 = C_1(n, s_0) > 0$ such that*

$$|\{x \in \mathbb{R}^n : \mathbf{M}u(x) > t\}| \le \frac{C_1}{t}\|u\|_{L^1(\mathbb{R}^n)}$$

for all $t > 0$ and $u \in L^1(\mathbb{R}^n)$. For $p > 0$, there is a constant $C_p = C_p(n, p, s_0) > 0$ such that

$$\|\mathbf{M}u(x)\|_{L^p(\mathbb{R}^n)} \le C_p\|u\|_{L^p(\mathbb{R}^n)}$$

for all $u \in L^p(\mathbb{R}^n)$.

We were not able to find a reference for the sharp maximal function theorem for sets of the type M_ρ. Actually, we are not sure whether such a result is available in the literature. However, one can follow the ideas of [31, Section 3.4], where the L^p bound is established for the sharp maximal function with cubes (instead of anisotropic rectangles). In order to prove the same result for the sharp maximal function with anisotropic rectangles, dyadic cubes have to be replaced by appropriate anisotropic “dyadic” rectangles. We construct the anisotropic dyadic rectangles in Appendix B and prove the following theorem in Appendix C. See also Appendix C for the definition of the dyadic maximal function $\mathbf{M}_d u$.

Theorem 2.6. *Let $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$ and let $0 < p_0 \le p < \infty$. Then, there is a constant $C = C(n, p, s_0) > 0$ such that for all $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ with $\mathbf{M}_d u \in L^{p_0}(\mathbb{R}^n)$,*

$$\|u\|_{L^p(\mathbb{R}^n)} \le C\|\mathbf{M}^\sharp u\|_{L^p(\mathbb{R}^n)}.$$

We are now in a position to prove a robust Sobolev-type inequality by using Lemma 2.4 and Theorems 2.5 and 2.6.

Theorem 2.7. *Let $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Suppose that $1 < p < n/\bar{s}$ and let $p_\star = np/(n - p\bar{s})$. Then, there is a constant $C = C(n, p, p_\star, s_0) > 0$ such that for every $u \in V^{p,\mu_{\mathrm{axes}}}(\mathbb{R}^n|\mathbb{R}^n)$*

$$\|u\|^p_{L^{p_\star}(\mathbb{R}^n)} \le C \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |u(x) - u(y)|^p \mu_{\mathrm{axes}}(x, \mathrm{d}y)\mathrm{d}x. \tag{2.1}$$

Proof. This proof is based on the technique of [47], which uses the maximal and sharp maximal inequalities. Note that by definition of $V^{p,\mu_{\mathrm{axes}}}(\mathbb{R}^n|\mathbb{R}^n)$ and Hölder's inequality, $V^{p,\mu_{\mathrm{axes}}}(\mathbb{R}^n|\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \subset L^p_{\mathrm{loc}}(\mathbb{R}^n) \subset L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Hence, the maximal and sharp maximal functions are well defined for every function $u \in V^{p,\mu_{\mathrm{axes}}}(\mathbb{R}^n|\mathbb{R}^n)$. For $x \in \mathbb{R}^n$ and $\rho > 0$, we have

$$\fint_{M_\rho(x)} |u(y) - (u)_{M_\rho(x)}|\mathrm{d}y \le \fint_{M_\rho(x)}\fint_{M_\rho(x)} |u(y) - u(z)|\mathrm{d}z\mathrm{d}y. \tag{2.2}$$

Let us consider as in [14, Lemma 2.1] a polygonal chain $\ell = (\ell_0(y,z), \dots, \ell_n(y,z)) \in \mathbb{R}^{n(n+1)}$ connecting y and z with

$$\ell_k(y,z) = (l_1^k, \dots, l_n^k), \quad \text{where } l_j^k = \begin{cases} z_j, & \text{if } j \le k, \\ y_j, & \text{if } j > k, \end{cases}$$

then $y = \ell_0(y,z)$, $z = \ell_n(y,z)$, and $|\ell_{k-1}(y,z) - \ell_k(y,z)| = |y_k - z_k|$ for all $k = 1, \dots, n$. By the triangle inequality, we have

$$\fint_{M_\rho(x)}\fint_{M_\rho(x)} |u(y)-u(z)|\,\mathrm{d}z\mathrm{d}y \le \sum_{k=1}^{n} \fint_{M_\rho(x)}\fint_{M_\rho(x)} |u(\ell_{k-1}(y,z)) - u(\ell_k(y,z))|\,\mathrm{d}z\mathrm{d}y. \tag{2.3}$$

For a fixed k, we set $w = \ell_{k-1}(y,z) = (z_1, \ldots, z_{k-1}, y_k, \ldots, y_n)$ and $v = y + z - w = (y_1, \ldots, y_{k-1}, z_k, \ldots, z_n)$, then $\ell_k(y,z) = w + e_k(v_k - w_k)$. By Fubini's theorem, we obtain

$$\fint_{M_\rho(x)}\fint_{M_\rho(x)} |u(\ell_{k-1}(y,z)) - u(\ell_k(y,z))|\,\mathrm{d}z\mathrm{d}y \le \fint_{M_\rho(x)} \fint_{x_k-\rho^{s_{\max}/s_k}}^{x_k+\rho^{s_{\max}/s_k}} |u(w) - u(w + e_k(v_k - w_k))|\,\mathrm{d}v_k\mathrm{d}w. \tag{2.4}$$

Moreover, using the inequality $|v_k - w_k| \le |v_k - x_k| + |w_k - x_k| < 2\rho^{s_{\max}/s_k}$, we make the inner integral on the right-hand side of (2.4) independent of x. Namely, we have

$$\begin{aligned}\fint_{x_k-\rho^{s_{\max}/s_k}}^{x_k+\rho^{s_{\max}/s_k}} |u(w) - u(w + e_k(v_k - w_k))|\,\mathrm{d}v_k &\le 2 \fint_{w_k-2\rho^{s_{\max}/s_k}}^{w_k+2\rho^{s_{\max}/s_k}} |u(w) - u(w + e_k(v_k - w_k))|\,\mathrm{d}v_k \\ &= 2 \fint_{-2\rho^{s_{\max}/s_k}}^{2\rho^{s_{\max}/s_k}} |u(w) - u(w + he_k)|\,\mathrm{d}h.\end{aligned} \tag{2.5}$$

Combining (2.2)–(2.5), we arrive at

$$\fint_{M_\rho(x)} |u(y) - (u)_{M_\rho(x)}|\,\mathrm{d}y \le \sum_{k=1}^{n} \rho^{s_{\max}} \fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w, \tag{2.6}$$

where the function F_k is defined by

$$F_k(w) := \sup_{\rho>0}\left(2\rho^{-s_{\max}} \fint_{-2\rho^{s_{\max}/s_k}}^{2\rho^{s_{\max}/s_k}} |u(w) - u(w + he_k)|\,\mathrm{d}h\right).$$

By Hölder's inequality,

$$\left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p_*} \le \left(\fint_{M_\rho(x)} F_k^p(w)\,\mathrm{d}w\right)^{\frac{p_*-p}{p}} \left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p} \le |M_\rho|^{-\frac{p_*-p}{p}} \|F_k\|_{L^p(\mathbb{R}^n)}^{p_*-p} \left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p}. \tag{2.7}$$

Thus, it follows from (2.6) and (2.7) that

$$\begin{aligned}\left(\fint_{M_\rho(x)} |u(y) - (u)_{M_\rho(x)}|\,\mathrm{d}y\right)^{p_*} &\le n^{p_*-1} \sum_{k=1}^{n} \rho^{p_* s_{\max}} \left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p_*} \\ &\le n^{p_*-1} \sum_{k=1}^{n} \rho^{p_* s_{\max}} |M_\rho|^{-\frac{p_*-p}{p}} \|F_k\|_{L^p(\mathbb{R}^n)}^{p_*-p} \left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p} \\ &\le \frac{n^{p_*-1}}{2^{n\frac{p_*-p}{p}}} \sum_{k=1}^{n} \|F_k\|_{L^p(\mathbb{R}^n)}^{p_*-p} \left(\fint_{M_\rho(x)} F_k(w)\,\mathrm{d}w\right)^{p}.\end{aligned}$$

Taking the supremum over $\rho > 0$, we obtain

$$(\mathbf{M}^\sharp u(x))^{p_*} \le \frac{n^{p_*-1}}{2^{n\frac{p_*-p}{p}}} \sum_{k=1}^{n} \|F_k\|_{L^p(\mathbb{R}^n)}^{p_*-p} (\mathbf{M}F_k(x))^p. \tag{2.8}$$

We now use Theorems 2.5 and 2.6. By (C.1) and Theorem 2.5, we know that $\mathbf{M}_d u \in L^p(\mathbb{R}^n)$. Thus, Theorem 2.6 yields that

$$\|u\|_{L^{p_*}(\mathbb{R}^n)}^{p_*} \le C\|\mathbf{M}^\sharp u\|_{L^{p_*}(\mathbb{R}^n)}^{p_*}$$

for some $C = C(n, p_*, s_0) > 0$. Moreover, assuming $F_k \in L^p(\mathbb{R}^n)$, we have by Theorem 2.5 and (2.8)

$$\|\mathbf{M}^\sharp u\|^{p_*}_{L^{p_*}(\mathbb{R}^n)} \le C\sum_{k=1}^n \|F_k\|^{p_*-p}_{L^p(\mathbb{R}^n)}\|\mathbf{M}F_k\|^p_p \le C\sum_{k=1}^n \|F_k\|^{p_*}_{L^p(\mathbb{R}^n)}$$

for some $C = C(n, p, p_*, s_0) > 0$. Therefore, it only remains to show that

$$\|F_k\|^p_{L^p(\mathbb{R}^n)} \le Cs_k(1-s_k)\int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{|u(x)-u(x+he_k)|^p}{|h|^{1+ps_k}}\,\mathrm{d}h\mathrm{d}x \tag{2.9}$$

for each $k = 1,\dots,n$.

Let us fix k. By using Hölder's inequality, we have

$$\begin{aligned}\|F_k\|^p_{L^p(\mathbb{R}^n)} &\le \int_{\mathbb{R}^n} \sup_{\rho>0} \frac{2^p}{\rho^{ps_{\max}}} \fint_{-2\rho^{s_{\max}/s_k}}^{2\rho^{s_{\max}/s_k}} |u(x)-u(x+he_k)|^p\,\mathrm{d}h\mathrm{d}x \\ &= \sum_{i=0}^\infty \int_{\mathbb{R}^n} \sup_{\rho>0} \frac{2^{p-2}}{\rho^{(1+ps_k)s_{\max}/s_k}} \int_{\mathbb{R}} |u(x)-u(x+he_k)|^p \mathbf{1}_{I_i}(|h|)\,\mathrm{d}h\mathrm{d}x,\end{aligned}$$

where $I_i = [2^{-i}\rho^{s_{\max}/s_k}, 2^{-i+1}\rho^{s_{\max}/s_k})$. For each i, let $\{\beta_{j,i}\}_{j=0}^\infty$ be a sequence such that $\sum_j \beta_{j,i} \ge 1$, which will be chosen later. Then,

$$\|F_k\|^p_{L^p(\mathbb{R}^n)} \le \sum_{i,j=0}^\infty \beta_{j,i}\int_{\mathbb{R}^n} \sup_{\rho>0} \frac{2^{p-2}}{\rho^{(1+ps_k)s_{\max}/s_k}} \int_{\mathbb{R}} |u(x)-u(x+he_k)|^p \mathbf{1}_{I_i}(|h|)\,\mathrm{d}h\mathrm{d}x.$$

By Lemma 2.4 for $N = 2^j$, $a = 2^{-i} \in (0,1]$ and $b = s_{\max}/s_k$, we obtain

$$\|F_k\|^p_{L^p(\mathbb{R}^n)} \le \sum_{i,j=0}^\infty 2^{p-2+(1+ps_k)(1-i)+p(1-s_k)j}\beta_{j,i}\int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{|u(x)-u(x+he_k)|^p}{|h|^{1+ps_k}}\mathbf{1}_{I_{i+j}}(|h|)\,\mathrm{d}h\mathrm{d}x.$$

We rearrange the double sums to have

$$\|F_k\|^p_{L^p(\mathbb{R}^n)} \le \sum_{i=0}^\infty\sum_{j=0}^i 2^{p-2+(1+ps_k)(1-i+j)+p(1-s_k)j}\beta_{j,i-j}\int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{|u(x)-u(x+he_k)|^p}{|h|^{1+ps_k}}\mathbf{1}_{I_i}(|h|)\,\mathrm{d}h\mathrm{d}x.$$

Let $\beta_{j,i} = p(\log 2)(1-s_k)2^{-p(1-s_k)j}$, then

$$1 \le \sum_{j=0}^\infty \beta_{j,i} = \frac{p(\log 2)(1-s_k)}{1-2^{-p(1-s_k)}} \le 2p < +\infty.$$

Since

$$\begin{aligned}\sum_{j=0}^i 2^{p-2+(1+ps_k)(1-i+j)+p(1-s_k)j}\beta_{j,i-j} &= p(\log 2)(1-s_k)2^{p-2+(1+ps_k)(1-i)}\sum_{j=0}^i 2^{(1+ps_k)j} \\ &= p(\log 2)(1-s_k)2^{p-2+(1+ps_k)(1-i)}\frac{2^{(1+ps_k)(i+1)}}{2^{1+ps_k}-1} \\ &\le p(1-s_k)2^{p(1+s_k)},\end{aligned}$$

we arrive at

$$\begin{aligned}\|F_k\|^p_{L^p(\mathbb{R}^n)} &\le \sum_{i=0}^\infty p(1-s_k)2^{p(1+s_k)}\int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{|u(x)-u(x+he_k)|^p}{|h|^{1+ps_k}}\mathbf{1}_{I_i}(|h|)\,\mathrm{d}h\mathrm{d}x \\ &\le p2^{2p}\frac{s_k}{s_0}(1-s_k)\int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{|u(x)-u(x+he_k)|^p}{|h|^{1+ps_k}}\,\mathrm{d}h\mathrm{d}x,\end{aligned}$$

which proves (2.9). □

Next, we can make use of appropriate cut-off functions to prove a localized version of the foregoing Sobolev-type inequality.

Corollary 2.8. *Let $\Lambda \geq 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Suppose that $1 < p < n/\bar{s}$. There is $C = C(n, p, p_\star, s_0, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $x_0 \in \mathbb{R}^n$, $r \in (0,1]$, $\lambda \in (1,2]$ and $u \in H^{p,\mu}_{M_{\lambda r}(x_0)}(\mathbb{R}^n)$ it holds*

$$\|u\|^p_{L^{p_\star}(M_r(x_0))} \leq C \int\limits_{M_{\lambda r}(x_0)} \int\limits_{M_{\lambda r}(x_0)} |u(x) - u(y)|^p \mu(x, \mathrm{d}y)\mathrm{d}x + C\left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-s_k p}\right) r^{-ps_{\max}} \|u\|^p_{L^p(M_{\lambda r}(x_0))},$$

where $p_\star$ is defined as in Theorem 2.7.

Proof. Let $\tau : \mathbb{R}^n \to \mathbb{R}$ be an admissible cut-off function in the sense of Definition 2.1. For simplicity of notation, we write $M_r = M_r(x_0)$.

By Theorem 2.7, there is a constant $c_1 = c_1(n, p, p_\star, s_0) > 0$ such that

$$\begin{aligned}\|u\tau\|^p_{L^{p_\star}(\mathbb{R}^n)} &\leq c_1\left(\int\limits_{M_{\lambda r}} \int\limits_{M_{\lambda r}} |u(x)\tau(x) - u(x)\tau(y)|^p \mu_{\text{axes}}(x, \mathrm{d}y)\mathrm{d}x + 2\int\limits_{M_{\lambda r}} \int\limits_{(M_{\lambda r})^c} |u(x)\tau(x) - u(x)\tau(y)|^p \mu_{\text{axes}}(x, \mathrm{d}y)\mathrm{d}x\right)\\ &=: c_1(I_1 + 2I_2).\end{aligned}$$

We have

$$\begin{aligned} I_1 &\leq \frac{1}{2^p}\left(\int\limits_{M_{\lambda r}} \int\limits_{M_{\lambda r}} 2^{p-1} \,|(u(y) - u(x))(\tau(x) + \tau(y))|^p \mu_{\text{axes}}(x, \mathrm{d}y)\mathrm{d}x\right.\\ &\qquad \left. + \int\limits_{M_{\lambda r}} \int\limits_{M_{\lambda r}} 2^{p-1} \,|(u(x) + u(y))(\tau(x) - \tau(y))|^p \mu_{\text{axes}}(x, \mathrm{d}y)\mathrm{d}x\right)\\ &= \frac{1}{2}(J_1 + J_2).\end{aligned}$$

Since $(\tau(x) + \tau(y)) \leq 2$ for all $x, y \in M_{\lambda r}$, we get

$$J_1 \leq 2^p \int\limits_{M_{\lambda r}} \int\limits_{M_{\lambda r}} |u(y) - u(x)|^p \mu_{\text{axes}}(x, \mathrm{d}y)\mathrm{d}x \leq \Lambda 2^p \int\limits_{M_{\lambda r}} \int\limits_{M_{\lambda r}} |u(y) - u(x)|^p \mu(x, \mathrm{d}y)\mathrm{d}x,$$

where we used Assumption 3 in the second inequality.

Moreover, since $|u(x) + u(y)|^p |\tau(x) - \tau(x)|^p \leq 2^{p-1}|u(x)|^p |\tau(x) - \tau(x)|^p + 2^{p-1}|u(y)|^p |\tau(y) - \tau(x)|^p$, we can again apply Assumption 3, and by Lemma 2.2, we get

$$J_2 \leq 2^p\left(\sup_{x \in \mathbb{R}^n} \int\limits_{\mathbb{R}^n} |\tau(y) - \tau(x)|^p \mu(x, \mathrm{d}y)\right)\|u\|^p_{L^p(M_{\lambda r})} \leq c_2\left(\sum_{k=1}^n \left(\lambda^{\frac{s_{\max}}{s_k}} - 1\right)^{-ps_k}\right) r^{-ps_{\max}} \|u\|^p_{L^p(M_{\lambda r})}$$

for some $c_2 > 0$, depending on n, p, s_0, and Λ. Moreover, by Corollary 2.3, there is $c_3 = c_3(n, p, \Lambda) > 0$ such that

$$I_2 \leq c_3\left(\sum_{k=1}^n \left(\lambda^{\frac{s_{\max}}{s_k}} - 1\right)^{-ps_k}\right) r^{-ps_{\max}} \|u\|^p_{L^p(M_{\lambda r})}.$$

Combining these estimates, we find a constant $C = C(n, p, p_\star, s_0, \Lambda) > 0$ such that

$$\|u\|^p_{L^{p_\star}(M_r)} \le \|u\tau\|^p_{L^{p_\star}(\mathbb{R}^n)}$$
$$\le C\left(\int_{M_{\lambda r}}\int_{M_{\lambda r}} |u(x)-u(y)|^p \mu_{\mathrm{axes}}(x,\mathrm{d}y)\mathrm{d}x + \left(\sum_{k=1}^n \left(\lambda^{\frac{s_{\max}}{s_k}}-1\right)^{-ps_k}\right) r^{-ps_{\max}}\|u\|^p_{L^p(M_{\lambda r})}\right). \qquad \square$$

Applying the same method as in the proof of the Sobolev-type inequality Theorem 2.7, we can deduce a Poincaré inequality.

Theorem 2.9. *Let $p > 1$, $\Lambda \ge 1$, and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. There is $C = C(n, p, s_0, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $x_0 \in \mathbb{R}^n$, $r \in (0,1]$ and $u \in L^p(M_r(x_0))$,*

$$\|u-(u)_{M_r(x_0)}\|^p_{L^p(M_r(x_0))} \le Cr^{ps_{\max}}\mathcal{E}^{\mu}_{M_r(x_0)}(u,u).$$

The proof is analog to the proof of the Poincaré inequality for the case $p = 2$, see [15, Theorem 4.2].

3 Weak Harnack inequality

In this section, we prove Theorem 1.4. The proof is based on Moser's iteration technique. We first need to verify a few properties for weak supersolutions to (1.1).

Lemma 3.1. *Let $\Lambda \ge 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. Let $1 < p \le n/\bar{s}$, $x_0 \in \mathbb{R}^n$, $r \in (0,1]$, and $\lambda \in (1,2]$. Set $M_r = M_r(x_0)$ and assume $f \in L^{q/(p\bar{s})}(M_{\lambda r})$ for some $q > n$. There is $C = C(n, p, s_0, \Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $u \in V^{p,\mu}(M_{\lambda r}|\mathbb{R}^n)$ that satisfies*

$$\mathcal{E}^{\mu}(u,\varphi) \ge (f,\varphi) \quad \text{for any nonnegative } \varphi \in H^{p,\mu}_{M_{\lambda r}}(\mathbb{R}^n), \quad u(x) \ge \varepsilon \quad \text{a.e. in } M_{\lambda r} \text{ for some } \varepsilon > 0,$$

the following holds:

$$\begin{aligned}\int_{M_r}\int_{M_r} &|\log u(y) - \log u(x)|^p \mu(x,\mathrm{d}y)\mathrm{d}x \\ \le\ & C\left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k}-1)^{-ps_k}\right) r^{-ps_{\max}}|M_{\lambda r}| \\ &+ \varepsilon^{1-p}\|f\|_{L^{q/(p\bar{s})}(M_{\lambda r})}|M_{\lambda r}|^{\frac{q-p\bar{s}}{q}} + 2\varepsilon^{1-p}|M_{\lambda r}| \sup_{x\in M_{(\lambda+1)r/2}} \int_{\mathbb{R}^n\setminus M_{\lambda r}} u_-^{p-1}(y)\mu(x,\mathrm{d}y).\end{aligned}$$

Proof. Let τ be an admissible cut-off function in the sense of Definition 2.1, and let $\varphi(x) = \tau^p(x)u^{1-p}(x)$, which is well defined since $\mathrm{supp}(\tau) \subset M_{\lambda r}$. Then, we have

$$\begin{aligned}(f,\varphi) \le & \int_{M_{\lambda r}}\int_{M_{\lambda r}} |u(x)-u(y)|^{p-2}(u(x)-u(y))\left(\frac{\tau^p(x)}{u^{p-1}(x)} - \frac{\tau^p(y)}{u^{p-1}(y)}\right)\mu(x,\mathrm{d}y)\mathrm{d}x \\ &+ 2\int_{M_{\lambda r}}\int_{\mathbb{R}^n\setminus M_{\lambda r}} |u(x)-u(y)|^{p-2}(u(x)-u(y))\frac{\tau^p(x)}{u^{p-1}(x)}\mu(x,\mathrm{d}y)\mathrm{d}x \\ =:& I_1 + I_2.\end{aligned} \tag{3.1}$$

Similar to the proof of [22, Lemma 1.3], we get the inequality

$$|u(x)-u(y)|^{p-2}(u(x)-u(y))\left(\frac{\tau^p(x)}{u^{p-1}(x)} - \frac{\tau^p(y)}{u^{p-1}(y)}\right) \le -c_1|\log u(x) - \log u(y)|^p\tau^p(y) + c_2|\tau(x)-\tau(y)|^p, \tag{3.2}$$

where $c_1, c_2 > 0$ are constants depending only on p. Hence, by (3.2) and Lemma 2.2,

$$I_1 \le -c_1 \int_{M_r}\int_{M_r} |\log u(x) - \log u(y)|^p \mu(x, dy)dx + C\left(\sum_{k=1}^{n} (\lambda^{s_{\max}/s_k} - 1)^{-ps_k}\right) r^{-ps_{\max}} |M_{\lambda r}|. \tag{3.3}$$

For I_2, again by Lemma 2.2,

$$\begin{aligned} I_2 &\le 2 \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} (u(x) - u(y))^{p-1} \frac{\tau^p(x)}{u^{p-1}(x)} \mathbf{1}_{\{u(x) \ge u(y)\}} \mu(x, dy)dx \\ &\le 2 \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} |\tau(x) - \tau(y)|^p \mu(x, dy)dx + 2 \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} u_-^{p-1}(y) \frac{\tau^p(x)}{\varepsilon^{p-1}} \mu(x, dy)dx \\ &\le C\left(\sum_{k=1}^{n} (\lambda^{s_{\max}/s_k} - 1)^{-ps_k}\right) r^{-ps_{\max}} |M_{\lambda r}| + \frac{2|M_{\lambda r}|}{\varepsilon^{p-1}} \sup_{x \in M_{(\lambda+1)r/2}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} u_-^{p-1}(y) \mu(x, dy), \end{aligned} \tag{3.4}$$

where we assumed that $\operatorname{supp}(\tau) \subset M_{(\lambda+1)r/2}$. Combining (3.1), (3.3), and (3.4), and using Hölder's inequality, we conclude that

$$\begin{aligned} &\int_{M_r}\int_{M_r} |\log u(y) - \log u(x)|^p \mu(x, dy)dx \\ &\le C\left(\sum_{k=1}^{n} (\lambda^{s_{\max}/s_k} - 1)^{-ps_k}\right) r^{-ps_{\max}} |M_{\lambda r}| + \varepsilon^{1-p} \|f\|_{L^{q/(p\bar{s})}(M_{\lambda r})} |M_{\lambda r}|^{\frac{q-p\bar{s}}{q}} \\ &\quad + 2\varepsilon^{1-p} |M_{\lambda r}| \sup_{x \in M_{(\lambda+1)/2}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} u_-^{p-1}(y) \mu(x, dy). \end{aligned}$$

□

The next theorem is an essential result to prove the weak Harnack inequality.

Theorem 3.2. *Let $\Lambda \ge 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0, 1)$. Let $1 < p < n/\bar{s}$, $x_0 \in \mathbb{R}^n$, and $r \in (0, 1]$. Set $M_r = M_r(x_0)$ and assume $f \in L^{q/(p\bar{s})}(M_{5r/4})$ for some $q > n$. There are $C = C(n, p, s_0, q, \Lambda) > 0$ and $\bar{p} = \bar{p}(n, p, s_0, q, \Lambda) \in (0, 1)$ such that for each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $u \in V^{p,\mu}(M_{5r/4}|\mathbb{R}^n)$ that satisfies*

$$\mathcal{E}^\mu(u, \varphi) \ge (f, \varphi) \quad \text{for any nonnegative } \varphi \in H^{p,\mu}_{M_{5r/4}}(\mathbb{R}^n), \quad u(x) \ge \varepsilon \quad \text{a.e. in } M_{5r/4},$$

for

$$\varepsilon > r^\delta \|f\|^{\frac{1}{p-1}}_{L^{q/(p\bar{s})}(M_{5r/4})} + \left(r^{ps_{\max}} \sup_{x \in M_{9r/8}} \int_{\mathbb{R}^n \setminus M_{5r/4}} u_-^{p-1}(y) \mu(x, dy)\right)^{\frac{1}{p-1}},$$

where $\delta = \frac{ps_{\max}}{p-1} \frac{q-n}{q}$, the following holds:

$$\left(\fint_{M_r} u^{\bar{p}}(x)dx\right)^{1/\bar{p}} \le C \left(\fint_{M_r} u^{-\bar{p}}(x)dx\right)^{-1/\bar{p}}.$$

Proof. We only need to prove that $\log u \in \mathrm{BMO}(M_r)$. The rest of the proof is standard. The Poincaré inequality (see Theorem 2.9) and Lemma 3.1 imply

$$\begin{aligned} \|\log u - (\log u)_{M_r}\|^p_{L^p(M_r)} &\le C r^{ps_{\max}} \mathcal{E}^\mu_{M_r}(\log u, \log u) \\ &\le C\left(\sum_{k=1}^{n} \left(\left(\frac{5}{4}\right)^{\frac{s_{\max}}{s_k}} - 1\right)^{-s_k p}\right) |M_{5r/4}| + C\varepsilon^{1-p} \|f\|_{L^{q/(p\bar{s})}(M_{5r/4})} r^{ps_{\max}} |M_{5r/4}|^{\frac{q-p\bar{s}}{q}} \\ &\quad + C\varepsilon^{1-p} r^{ps_{\max}} |M_{5r/4}| \sup_{x \in M_{9r/8}} \int_{\mathbb{R}^n \setminus M_{5r/4}} u_-^{p-1}(y) \mu(x, dy) \le C|M_r|, \end{aligned}$$

where we used the bound on ε in the last inequality. Finally, by Hölder's inequality, we obtain

$$\|\log u\|_{\mathrm{BMO}(M_r)} \leq \left(\fint_{M_r} |\log u - (\log u)_{M_r}|^p \mathrm{d}x \right)^{1/p} \leq C,$$

which shows that $\log u \in \mathrm{BMO}(B_r)$. □

In order to apply Moser's iteration for negative exponents, we prove the following lemma.

Lemma 3.3. *Let $\Lambda \geq 1$ and $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0, 1)$. Let $1 < p < n/\bar{s}, x_0 \in \mathbb{R}^n$, $r \in (0, 1]$, and $\lambda \in (1, 2]$. Set $M_r = M_r(x_0)$ and assume $f \in L^{q/(p\bar{s})}(M_{\lambda r})$ for some $q > n$. For each $\mu \in \mathcal{K}(p, s_0, \Lambda)$ and every $u \in V^{p,\mu}(M_{\lambda r}|\mathbb{R}^n)$ that satisfies*

$$\mathcal{E}^\mu(u, \varphi) \geq (f, \varphi) \quad \text{for any nonnegative } \varphi \in H^{p,\mu}_{M_{\lambda r}}(\mathbb{R}^n), \quad u(x) \geq \varepsilon \quad \text{a.e. in } M_{\lambda r},$$

for

$$\varepsilon > r^\delta \|f\|^{\frac{1}{p-1}}_{L^{q/(p\bar{s})}(M_{\lambda r/4})} + \left(r^{ps_{\max}} \sup_{x \in M_{(\lambda+1)r/2}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} u_-^{p-1}(y) \mu(x, \mathrm{d}y) \right)^{\frac{1}{p-1}},$$

the following is true for any $t > p - 1$,

$$\|u^{-1}\|^{t-p+1}_{L^{(t-p+1)\gamma}(M_r)} \leq C \left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-s_k p} \right) r^{-s_{\max} p} \|u^{-1}\|^{t-p+1}_{L^{t-p+1}(M_{\lambda r})},$$

where $\delta = \frac{ps_{\max}}{p-1} \frac{q-n}{q}$, $\gamma = n/(n - p\bar{s})$, and $C = C(n, p, p_\star, q, t, s_0, \Lambda) > 0$ is a constant that is bounded when t is bounded away from $p - 1$.

To prove Lemma 3.3, we need the following algebraic inequality.

Lemma 3.4. *Let $a, b > 0$, $\tau_1, \tau_2 \in [0, 1]$, and $t > p - 1 > 0$. Then,*

$$|b - a|^{p-2}(b - a)(\tau_1^p a^{-t} - \tau_2^p b^{-t}) \geq c_1 \left| \tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}} \right|^p - c_2 |\tau_1 - \tau_2|^p (a^{-t+p-1} + b^{-t+p-1}),$$

where $c_i = c_i(p, t) > 0$, $i = 1, 2$, is bounded when t is bounded away from $p - 1$.

Note that Lemma 3.4 is a discrete version of

$$|\nabla v|^{p-2} \nabla v \cdot \nabla(-v^{-t}\tau^p) \geq c_1 \left| \nabla \left(v^{\frac{-t+p-1}{p}} \tau \right) \right|^p - c_2 |\nabla \tau|^p v^{-t+p-1}.$$

The proof of Lemma 3.4 is provided in Appendix A.

Proof of Lemma 3.3. Let τ be an admissible cut-off function in the sense of Definition 2.1. Since $\tau = 0$ outside $M_{\lambda r}$, the function $\varphi = -\tau^p u^{-t}$ is well defined. By using Lemma 3.4, we have

$$\begin{aligned}
(f, -\tau^p u^{-t}) &\geq \mathcal{E}(u, -\tau^p u^{-t}) \\
&= \int_{M_{\lambda r}} \int_{M_{\lambda r}} |u(y) - u(x)|^{p-2}(u(y) - u(x))(\tau^p(x) u^{-t}(x) - \tau^p(y) u^{-t}(y)) \mu(x, \mathrm{d}y) \mathrm{d}x \\
&\quad + 2 \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} |u(y) - u(x)|^{p-2}(u(y) - u(x)) \tau^p(x) u^{-t}(x) \mu(x, \mathrm{d}y) \mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&\geq c_1 \int_{M_{\lambda r}} \int_{M_{\lambda r}} \left| \tau(x) u^{\frac{-t+p-1}{p}}(x) - \tau(y) u^{\frac{-t+p-1}{p}}(y) \right|^p \mu(x, \mathrm{d}y)\mathrm{d}x \\
&\quad - c_2 \int_{M_{\lambda r}} \int_{M_{\lambda r}} |\tau(x) - \tau(y)|^p (u^{-t+p-1}(x) + u^{-t+p-1}(y)) \mu(x, \mathrm{d}y)\mathrm{d}x \\
&\quad - 2 \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} (u(x) - u(y))^{p-1} \tau^p(x) u^{-t}(x) \mathbf{1}_{\{u(y) \leq u(x)\}} \mu(x, \mathrm{d}y)\mathrm{d}x \\
&=: c_1 I_1 - c_2 I_2 - I_3,
\end{aligned}
$$

where c_1 and c_2 are constants given in Lemma 3.4. By Theorem 2.8, we obtain

$$
\begin{aligned}
I_1 &= \int_{M_{\lambda r}} \int_{M_{\lambda r}} \left| \tau(x) u^{\frac{-t+p-1}{p}}(x) - \tau(y) u^{\frac{-t+p-1}{p}}(y) \right|^p \mu(x, \mathrm{d}y)\mathrm{d}x \\
&\geq C \left\| \tau u^{\frac{-t+p-1}{p}} \right\|^p_{L^{p_\star}(M_r)} - C r^{-p s_{\max}} \left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-s_k p} \right) \left\| \tau u^{\frac{-t+p-1}{p}} \right\|^p_{L^p(M_{\lambda r})},
\end{aligned}
$$

where $p_\star = \frac{np}{n - p\bar{s}}$.

For I_2, we use Lemma 2.2 again to have

$$
I_2 = 2 \int_{M_{\lambda r}} \int_{M_{\lambda r}} |\tau(x) - \tau(y)|^p u^{-t+p-1}(x) \mu(x, \mathrm{d}y)\mathrm{d}x \leq C r^{-p s_{\max}} \left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-s_k p} \right) \|u^{-t+p-1}\|_{L^1(M_{\lambda r})}.
$$

For I_3, assuming that $\operatorname{supp}(\tau) \subset M_{(\lambda+1)r/2}$ and using Lemma 2.2, we deduce

$$
\begin{aligned}
I_3 &\leq C \int_{M_{\lambda r}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} (u^{p-1}(x) + u_-^{p-1}(y)) \tau^p(x) u^{-t}(x) \mu(x, \mathrm{d}y)\mathrm{d}x \\
&\leq C r^{-p s_{\max}} \left(\sum_{k=1}^n (\lambda^{s_{\max}/s_k} - 1)^{-s_k p} \right) \|u^{-t+p-1}\|_{L^1(M_{\lambda r})} + C \varepsilon^{1-p} \left(\sup_{x \in M_{(\lambda+1)r/2}} \int_{\mathbb{R}^n \setminus M_{\lambda r}} u_-^{p-1}(y) \mu(x, \mathrm{d}y) \right) \|u^{-t+p-1}\|_{L^1(M_{\lambda r})}.
\end{aligned}
$$

Moreover, we estimate

$$
\begin{aligned}
|(f, -\tau^p u^{-t})| &\leq \varepsilon^{1-p} \int_{\mathbb{R}^n} |f| \tau^p u^{-t+p-1} \mathrm{d}x \\
&\leq \varepsilon^{1-p} \|f\|_{L^{q/(p\bar{s})}(M_{\lambda r})} \|\tau^p u^{-t+p-1}\|_{L^{q/(q-p\bar{s})}(M_{\lambda r})} \\
&= \varepsilon^{1-p} \|f\|_{L^{q/(p\bar{s})}(M_{\lambda r})} \left\| \tau u^{\frac{-t+p-1}{p}} \right\|^p_{L^{pq/(q-p\bar{s})}(M_{\lambda r})}.
\end{aligned}
$$

By using Lyapunov's inequality and Young's inequality, we have

$$
\|v\|^p_{pq/(q-p\bar{s})} \leq \|v\|^{np/q}_{p_\star} \|v\|^{(qp-np)/q}_p \leq \frac{n}{q} \omega \|v\|^p_{p_\star} + \frac{q-n}{q} \omega^{-n/(q-n)} \|v\|^p_p
$$

for any $v \in L^{p_\star} \cap L^p$ and any $\omega > 0$. This yields that

$$
\begin{aligned}
|(f, -\tau^p u^{-t})| &\leq \varepsilon^{1-p} \|f\|_{L^{q/(p\bar{s})}(M_{\lambda r})} \left(\frac{n}{q} \omega \left\| \tau u^{\frac{-t+p-1}{p}} \right\|^p_{L^{p_\star}} + \frac{q-n}{q} \omega^{-n/(q-n)} \left\| \tau u^{\frac{-t+p-1}{p}} \right\|^p_{L^p} \right) \\
&\leq r^{-p s_{\max} \frac{q-n}{q}} \left(\frac{n}{q} \omega \|\tau^p u^{-t+p-1}\|_{L^\gamma} + \frac{q-n}{q} \omega^{-n/(q-n)} \|\tau^p u^{-t+p-1}\|_{L^1} \right).
\end{aligned}
$$

Combining all the estimates, we have

$$\|\tau^p u^{-t+p-1}\|_{L^\gamma(M_{\lambda r})} \leq Cr^{-ps_{\max}}\left(1 + \sum_{k=1}^{n}(\lambda^{s_{\max}/s_k} - 1)^{-s_k p}\right)\|u^{-t+p-1}\|_{L^1(M_{\lambda r})}$$
$$+ Cr^{-ps_{\max}\frac{q-n}{q}}\left(\frac{n}{q}\omega\,\|\tau^p u^{-t+p-1}\|_{L^\gamma(M_{\lambda r})} + \frac{q-n}{q}\omega^{-n/(q-n)}\,\|\tau^p u^{-t+p-1}\|_{L^1(M_{\lambda r})}\right).$$

Taking $\omega = \varepsilon_0 r^{-ps_{\max}\frac{q-n}{q}}$ with $\varepsilon_0 > 0$ small enough, we arrive at

$$\|u^{-1}\|^{t-p+1}_{L^{(t-p+1)\gamma}(M_r)} \leq \|\tau^p u^{-t+p-1}\|_{L^\gamma(M_{\lambda r})} \leq C\left(1 + \sum_{k=1}^{n}(\lambda^{s_{\max}/s_k} - 1)^{-s_k p}\right)r^{-ps_{\max}}\,\|u^{-1}\|^{t-p+1}_{L^{t-p+1}(M_{\lambda r})},$$

where C depends on n, p, $p_\star$, t, s_0, q, and Λ and is bounded when t is bounded away from $p-1$. Since $\lambda \leq 2$, we obtain

$$\sum_{k=1}^{n}(\lambda^{s_{\max}/s_k} - 1)^{-s_k p} \geq \sum_{k=1}^{n}(\lambda^{1/s_k} - 1)^{-s_k p} \geq \sum_{k=1}^{n}2^{-p} = n2^{-p},$$

from which the desired result follows. □

The standard iteration technique proves the following lemma, see [25,14].

Lemma 3.5. *Under the same assumptions as in* Lemma 3.3, *for any* $p_0 > 0$, *there is a constant* $C = C(n, p, p_\star, q, p_0, s_0, \Lambda) > 0$ *such that*

$$\inf_{M_r} u \geq C\left(\fint_{M_{2r}} u(x)^{-p_0}\mathrm{d}x\right)^{-1/p_0}. \tag{3.5}$$

The proof of Theorem 1.4 follows from Theorem 3.2, Lemma 3.5, and the triangle inequality.

4 Hölder estimates

This section is devoted to the proof of Theorem 1.5. The general scheme for the derivation of a priori interior Hölder estimates from the weak Harnack inequality in the nonlocal setting has been developed in [25] and applied successfully to the anisotropic setting [14] when $p = 2$. We extend the result presented in [14] to the general case $p > 1$.

Recall that the rectangles in Definition 1.1 satisfy the following property. For $\lambda > 0$ and $\Omega \subset \mathbb{R}^n$ open, we have

$$\mathcal{E}^{\mu_{\text{axes}}}_{\Omega}(u \circ \Psi, v \circ \Psi) = \lambda^{-(n-\bar{s}p)s_{\max}/\bar{s}}\mathcal{E}^{\mu_{\text{axes}}}_{\Psi(\Omega)}(u, v) \quad \text{for every } u, v \in V^{p,\mu_{\text{axes}}}(\Omega|\mathbb{R}^n)$$

and

$$(f \circ \Psi, \varphi \circ \Psi) = \lambda^{-ns_{\max}/\bar{s}}(f, \varphi) \quad \text{for every } f \in L^{q/(p\bar{s})}(\Omega),\ \varphi \in H^{p,\mu_{\text{axes}}}_{\Omega}(\mathbb{R}^n),$$

where $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism given by

$$\Psi(x) = \begin{pmatrix} \lambda^{s_{\max}/s_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda^{s_{\max}/s_n} \end{pmatrix} x. \tag{4.1}$$

The rectangles from Definition 1.1 are balls in a metric space $(\mathbb{R}^n, d)$, where the metric $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ is defined as follows:

$$d(x,y) = \sup_{k\in\{1,\dots,n\}} |x_k - y_k|^{s_k/s_{\max}}.$$

By the scaling property and covering arguments provided in [14], it is enough to show the following theorem.

Theorem 4.1. *Let $\Lambda \ge 1$ and $s_1,\dots,s_n \in [s_0,1)$ be given for some $s_0 \in (0,1)$. Let $1 < p < n/\bar{s}$. Assume $f \in L^{q/(p\bar{s})}(M_1)$ for some $q > n$. There are $\alpha = \alpha(n,p,p_\star,q,s_0,\Lambda) \in (0,1)$ and $C = C(n,p,p_\star,q,s_0,\Lambda) > 0$ such that for each $\mu \in \mathcal{K}(p,s_0,\Lambda)$ and every $u \in V^{p,\mu}(M_1|\mathbb{R}^n)$ satisfying*

$$\mathcal{E}^\mu(u,\varphi) = (f,\varphi) \quad \text{for any } \varphi \in H^{p,\mu}_{M_1}(\mathbb{R}^n),$$

we have $u \in C^\alpha$ at 0 and

$$|u(x) - u(0)| \le C(\|u\|_{L^\infty(\mathbb{R}^n)} + \|f\|_{L^{q/(p\bar{s})}(M_{15/16})}) d(x,0)^\alpha$$

for all $x \in M_1$.

Proof. We assume that $2\|u\|_{L^\infty(\mathbb{R}^n)} + \kappa^{-1}\|f\|_{L^{q/(p\bar{s})}(M_{15/16})} \le 1$ for some $\kappa > 0$, which will be chosen later. It is enough to construct sequences $\{a_k\}_{k=0}^\infty$ and $\{b_k\}_{k=0}^\infty$ such that $a_k \le u \le b_k$ in $M_{1/4^k}$ and $b_k - a_k = 4^{-\alpha k}$ for some $\alpha > 0$. For $k = 0$, we set $a_0 = -1/2$ and $b_0 = 1/2$. Assume that we have constructed such sequences up to k and let us choose a_{k+1} and b_{k+1}.

We assume

$$|\{u \ge (b_k + a_k)/2\} \cap M_{\frac{1}{2}4^{-k}}| \ge |M_{\frac{1}{2}4^{-k}}|/2, \tag{4.2}$$

and then prove that we can choose a_{k+1} and b_{k+1}. If (4.2) does not hold, then we can consider $-u$ instead of u. Let Ψ be the diffeomorphism given by (4.1) with $\lambda = 4^{-k}$ and define

$$v(x) = \frac{u(\Psi(x)) - a_k}{(b_k - a_k)/2} \quad \text{and} \quad g(x) = \frac{\lambda^{ps_{\max}} f(\Psi(x))}{(b_k - a_k)/2}.$$

Then, $v \ge 0$ in M_1 and $\mathcal{E}^\mu_{M_1}(v,\varphi) = (g,\varphi)$ for every $\varphi \in H^\mu_{M_1}(\mathbb{R}^n)$. Moreover, it is easy to see that $v \ge 2(1 - 4^{\alpha j})$ in M_{4^j} for every $j \ge 0$ by induction hypothesis. By applying Theorem 1.4, we obtain

$$\left(\fint_{M_{1/2}} v^{p_0}(x)\,\mathrm{d}x \right)^{1/p_0} \le C \inf_{M_{1/4}} v + C \sup_{x\in M_{15/16}} \left(\int_{\mathbb{R}^n\setminus M_1} v^-(y)^{p-1}\mu(x,\mathrm{d}y) \right)^{1/(p-1)} + \|g\|_{L^{q/(p\bar{s})}(M_{15/16})}. \tag{4.3}$$

By taking $\alpha < ps_0\frac{q-n}{q}$, we have

$$\|g\|_{L^{q/(p\bar{s})}(M_{15/16})} = 2 \cdot 4^{(\alpha - ps_{\max}\frac{q-n}{q})k} \|f\|_{L^{q/(p\bar{s})}\left(M_{4^{-k}\cdot 15/16}\right)} \le 2\kappa. \tag{4.4}$$

For $x \in M_{15/16}$ and for each $j \ge 1$, we have $M_{4^j}\setminus M_{4^{j+1}} \subset \mathbb{R}^n\setminus M_{4^j}(x)$. Hence, by (1.5),

$$\begin{aligned} \int_{\mathbb{R}^n\setminus M_1} v^-(y)^{p-1}\mu(x,\mathrm{d}y) &\le \sum_{j=1}^\infty \int_{M_{4^j}\setminus M_{4^{j+1}}} (2(4^{\alpha j} - 1))^{p-1}\mu(x,\mathrm{d}y) \\ &\le \sum_{j=1}^\infty (2(4^{\alpha j} - 1))^{p-1}\mu\big(x, \mathbb{R}^n\setminus M_{4^j}(x)\big) \\ &\le \sum_{j=1}^l \Lambda(2(4^{\alpha j} - 1))^{p-1} 4^{-ps_0 j} + 2^{p-1}\Lambda \sum_{j=l+1}^\infty 4^{(\alpha(p-1) - ps_0)j}. \end{aligned} \tag{4.5}$$

If we assume that $\alpha < \frac{ps_0}{2(p-1)}$, then we can make the last term in (4.5) as small as we want by taking $l = l(p,s_0)$ sufficiently large. Since the first term in (4.5) converges to 0 as $\alpha \to 0$, we have

$$C \sup_{x \in M_{15/16}} \left(\int_{\mathbb{R}^n \setminus M_1} v^-(y)^{p-1} \mu(x, \mathrm{d}y) \right)^{1/(p-1)} \leq \kappa \tag{4.6}$$

by assuming further that $\alpha = \alpha(n, p, p_\star, q, s_0, \Lambda)$ is sufficiently small.

On the other hand, it follows from (4.2) that

$$\left(\fint_{M_{1/2}} v^{p_0}(x) \mathrm{d}x \right)^{1/p_0} \geq \left(\frac{1}{|M_{1/2}|} \int_{M_{1/2} \cap \{v \geq 1\}} v^{p_0}(x) \mathrm{d}x \right)^{1/p_0} \geq 2^{-1/p_0}. \tag{4.7}$$

Combining (4.3), (4.4), (4.6), and (4.7), and choosing $\kappa > 0$ sufficiently small, we arrive at $\inf_{M_{1/4}} v \geq \kappa_0$ for some κ_0. We take $a_{k+1} = a_k + \kappa_0(b_k - a_k)/2$ and $b_{k+1} = b_k$, and make α and κ_0 small so that $1 - \kappa_0/2 = 4^{-\alpha}$. Then, $a_{k+1} \leq u \leq b_{k+1}$ in $M_{4^{-(k+1)}}$ and $b_{k+1} - a_{k-1} = 4^{-\alpha(k+1)}$, which finishes the proof. □

Funding information: Jamil Chaker gratefully acknowledges support by the Deutsche Forschungsgemeinschaft (SFB 1283/2 2021 – 317210226). Minhyun Kim gratefully acknowledges funding by the Deutsche Forschungsgemeinschaft (GRK 2235/2 2021 - 282638148).

Conflict of interest: The authors state no conflict of interest.

References

[1] A. Banerjee, P. Garain, and J. Kinnunen, *Some local properties of subsolution and supersolutions for a doubly nonlinear nonlocal p-Laplace equation*, Annali di Matematica Pura ed Applicata (2021), 1–35.

[2] M. T. Barlow, R. F. Bass, Z.-Q. Chen, and M. Kassmann, *Non-local Dirichlet forms and symmetric jump processes*, Trans. Amer. Math. Soc. **361** (2009), no. 4, 1963–1999.

[3] P. Baroni, A. DiCastro, and G. Palatucci, *Intrinsic geometry and De Giorgi classes for certain anisotropic problems*, Discrete Contin. Dyn. Syst. Ser. S **10** (2017), no. 4, 647–659.

[4] R. F. Bass and Z.-Q. Chen, *Regularity of harmonic functions for a class of singular stable-like processes*, Math. Z. **266** (2010), no. 3, 489–503.

[5] R. F. Bass and D. A. Levin, *Harnack inequalities for jump processes*, Potential Anal. **17** (2002), no. 4, 375–388.

[6] M. Belloni and B. Kawohl, *The pseudo-p-Laplace eigenvalue problem and viscosity solutions as p → ∞*, ESAIM Control Optim. Calc. Var. **10** (2004), no. 1, 28–52.

[7] K. Bogdan and P. Sztonyk, *Estimates of the potential kernel and Harnackas inequality for the anisotropic fractional Laplacian*, Studia Math. **181** (2007), no. 2, 101–123.

[8] P. Bousquet and L. Brasco, *Lipschitz regularity for orthotropic functionals with nonstandard growth conditions*, Rev. Mat. Iberoam. **36** (2020), no. 7, 1989–2032.

[9] P. Bousquet, L. Brasco, C. Leone, and A. Verde, *On the Lipschitz character of orthotropic p-harmonic functions*, Calc. Var. Partial Differ. Equ. **57** (2018), no. 3, Paper No. 88, 33.

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

[11] C. Bucur and M. Squassina, *An asymptotic expansion for the fractional p-Laplacian and for gradient-dependent nonlocal operators*, Commun. Contemp. Math. (2021), 2150021.

[12] L. Caffarelli, C. H. Chan, and A. Vasseur, *Regularity theory for parabolic nonlinear integral operators*, J. Amer. Math. Soc. **24** (2011), no. 3, 849–869.

[13] J. Chaker, *Regularity of solutions to anisotropic nonlocal equations*, Math. Z. **296** (2020), no. 3–4, 1135–1155.

[14] J. Chaker and M. Kassmann, *Nonlocal operators with singular anisotropic kernels*, Comm. Partial Differ. Equ. **45** (2020), no. 1, 1–31.

[15] J. Chaker, M. Kassmann, and M. Weidner, *Robust Hölder estimates for parabolic nonlocal operators*, Preprint (2019), https://arxiv.org/abs/1912.09919.

[16] Z.-Q. Chen and T. Kumagai, *Heat kernel estimates for stable-like processes on d-sets*, Stochastic Process. Appl. **108** (2003), no. 1, 27–62.

[17] Z.-Q. Chen, T. Kumagai, and J. Wang, *Elliptic Harnack inequalities for symmetric non-local Dirichlet forms*, J. Math. Pures Appl. **125** (2019), no. 9, 1–42.
[18] M. Cozzi, *Regularity results and Harnack inequalities for minimizers and solutions of nonlocal problems: a unified approach via fractional De Giorgi classes*, J. Funct. Anal. **272** (2017), no. 11, 4762–4837.
[19] E. De Giorgi, *Sulla differenziabilità e laanaliticità delle estremali degli integrali multipli regolari*, Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. **3** (1957), no. 3, 25–43.
[20] F. del Teso, D. Gómez-Castro, and J. L. Vázquez, *Three representations of the fractional p-Laplacian: semigroup, extension and Balakrishnan formulas*, Fract. Calc. Appl. Anal. **24** (2021), no. 4, 966–1002.
[21] A. Di Castro, T. Kuusi, and G. Palatucci, *Nonlocal Harnack inequalities*, J. Funct. Anal. **267** (2014), no. 6, 1807–1836.
[22] A. Di Castro, T. Kuusi, and G. Palatucci, *Local behavior of fractional p-minimizers*, Ann. Inst. H. Poincaré Anal. Non Linéaire. **33** (2016), no. 5, 1279–1299.
[23] E. dos Santos and R. Leitaao, *On the Hölder regularity for solutions of integro-differential equations like the anisotropic fractional Laplacian*, SN Partial Differ. Eq. Appl. **2** (2021), no. 2, 1–34.
[24] B. Dyda and M. Kassmann, *Function spaces and extension results for nonlocal Dirichlet problems*, J. Funct. Anal. **277** (2019), no. 11, 108134, 22.
[25] B. Dyda and M. Kassmann, *Regularity estimates for elliptic nonlocal operators*, Anal. PDE **13** (2020), no. 2, 317–370.
[26] M. M. Fall, *Regularity results for nonlocal equations and applications*, Calc. Var. Partial Differ. Equ. **59** (2020), no. 5, Paper No. 181, 53.
[27] M. Felsinger and M. Kassmann, *Local regularity for parabolic nonlocal operators*, Comm. Partial Differ. Equ. **38** (2013), no. 9, 1539–1573.
[28] M. Felsinger, M. Kassmann, and P. Voigt, *The Dirichlet problem for nonlocal operators*, Math. Z. **279** (2015), no. 3–4, 779–809.
[29] M. Friesen and P. Jin, *On the anisotropic stable JCIR process*, ALEA Lat. Am. J. Probab. Math. Stat. **17** (2020), no. 2, 643–674.
[30] M. Friesen, P. Jin, and B. Ruuudiger, *Existence of densities for stochastic differential equations driven by Lévy processes with anisotropic jumps*, Ann. Inst. Henri Poincaré Probab. Stat. **57** (2021), no. 1, 250–271.
[31] L. Grafakos, *Modern Fourier analysis*, Graduate Texts in Mathematics, Third edition, vol. 250, Springer, New York, 2014.
[32] J. Heinonen, *Lectures on Analysis on Metric Spaces*, Universitext, Springer-Verlag, New York, 2001.
[33] M. Kassmann, *A priori estimates for integro-differential operators with measurable kernels*, Calc. Var. Partial Differ. Equ. **34** (2009), no. 1, 1–21.
[34] M. Kassmann and R. W. Schwab, *Regularity results for nonlocal parabolic equations*, Riv. Math. Univ. Parma (N.S.) **5** (2014), no. 1, 183–212.
[35] M. Kassmann, K.-Y. Kim, and T. Kumagai, *Heat kernel bounds for nonlocal operators with singular kernels*, (2019), Preprint, https://arxiv.org/abs/1910.04242.
[36] T. Kulczycki and M. Ryznar, *Semigroup properties of solutions of SDEs driven by Lévy processes with independent coordinates*, Stochastic Process. Appl. **130** (2020), no. 12, 7185–7217.
[37] T. Kulczycki and M. Ryznar, *Transition density estimates for diagonal systems of SDEs driven by cylindrical α-stable processes*, ALEA Lat. Am. J. Probab. Math. Stat. **15** (2018), no. 2, 1335–1375.
[38] T. Kuusi, G. Mingione, and Y. Sire, *Nonlocal equations with measure data*, Comm. Math. Phys. **337** (2015), no. 3, 1317–1368.
[39] T. Leonori, I. Peral, A. Primo, and F. Soria, *Basic estimates for solutions of a class of nonlocal elliptic and parabolic equations*, Discrete Contin. Dyn. Syst. **35** (2015), no. 12, 6031–6068.
[40] E. Lindgren, *Hölder estimates for viscosity solutions of equations of fractional p-Laplace type*, NoDEA Nonlinear Differ. Equ. Appl. **23** (2016), no. 1, 55, 18.
[41] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969.
[42] C. W. Lo and J. F. Rodrigues, *On a class of fractional obstacle type problems related to the distributional Riesz derivative*, Preprint (2021), https://arxiv.org/abs/2101.06863.
[43] G. Mingione, *Gradient potential estimates*, J. Eur. Math. Soc. (JEMS), **13** (2011), no. 2, 459–486.
[44] S. J. N. Mosconi, *Optimal elliptic regularity: a comparison between local and nonlocal equations*, Discrete Contin. Dyn. Syst. Ser. S **11** (2018), no. 3, 547–559.
[45] J. Moser, *On Harnackas theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961), 577–591.
[46] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math. **80** (1958), 931–954.
[47] A. D. Nguyen, J. I. Díaz, and Q.-H. Nguyen, *Fractional Sobolev inequalities revisited: the maximal function approach*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. **31** (2020), no. 1, 225–236.
[48] S. Nowak, *Improved sobolev regularity for linear nonlocal equations with VMO coefficients*, Math. Ann. (2022), 1–56, https://arxiv.org/abs/2108.02856.
[49] S. Nowak, *$H^{s,p}$ regularity theory for a class of nonlocal elliptic equations*, Nonlinear Anal. **195** (2020), 111730, 28.
[50] S. Nowak, *Higher Hölder regularity for nonlocal equations with irregular kernel*, Calc. Var. Partial Differ. Equ. **60** (2021), no. 1, Paper No. 24, 37.

[51] S. Nowak, *Regularity theory for nonlocal equations with VMO coefficients*, Ann. Inst. H. Poincaré Anal. Non Linéaire. (2021), https://arxiv.org/abs/2101.11690.

[52] X. Ros-Oton and E. Valdinoci, *The Dirichlet problem for nonlocal operators with singular kernels: convex and nonconvex domains*, Adv. Math. **288** (2016), 732–790.

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

[54] M. Strömqvist, *Harnackas inequality for parabolic nonlocal equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **36** (2019), no. 6, 1709–1745.

[55] M. Strömqvist, *Local boundedness of solutions to non-local parabolic equations modeled on the fractional p-Laplacian*, J. Differ. Equ. **266** (2019), no. 12, 7948–7979.

[56] L. Wang and X. Zhang, *Harnack inequalities for SDEs driven by cylindrical α-stable processes*, Potential Anal. **42** (2015), no. 3, 657–669.

Appendix

A Algebraic inequalities

In this section, we prove Lemma 3.4 using the series of following lemmas.

Lemma A.1. *Let $a, b > 0$ and $t > p-1 > 0$. Then,*

$$|b-a|^{p-2}(b-a)(a^{-t}-b^{-t}) \geq t\left(\frac{p}{t-p+1}\right)^p \left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p.$$

Proof. We may assume that $b > a$. Let $f(x) = -x^{\frac{-t+p-1}{p}}$ and $g(x) = -x^{-t}$, and then by using Jensen's inequality, we have

$$\begin{aligned}\left|\frac{f(b)-f(a)}{b-a}\right|^p &= \left|\fint_a^b f'(x)\mathrm{d}x\right|^p \leq \fint_a^b (f'(x))^p \mathrm{d}x \\ &= \frac{1}{t}\left(\frac{t-p+1}{p}\right)^p \fint_a^b g'(x)\mathrm{d}x = \frac{1}{t}\left(\frac{t-p+1}{p}\right)^p \frac{g(b)-g(a)}{b-a},\end{aligned}$$

which proves the lemma. □

Lemma A.2. *Let $a, b > 0$ and $t > p-1 > 0$. Then,*

$$|b-a|^{p-1}\min\{a^{-t}, b^{-t}\} \leq \left(\frac{p}{t-p+1}\right)^{p-1} \left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^{p-1} \min\left\{a^{\frac{-t+p-1}{p}}, b^{\frac{-t+p-1}{p}}\right\}.$$

Proof. We may assume that $b > a$. Let $f(x) = -x^{\frac{-t+p-1}{p}}$, then

$$\begin{aligned}\left|\frac{f(b)-f(a)}{b-a}\right|^{p-1} &= \left|\fint_a^b f'(x)\mathrm{d}x\right|^{p-1} = \left(\frac{t-p+1}{p}\right)^{p-1}\left(\fint_a^b x^{\frac{-t-1}{p}}\mathrm{d}x\right)^{p-1} \\ &\geq \left(\frac{t-p+1}{p}\right)^{p-1}\left(\fint_a^b b^{\frac{-t-1}{p}}\mathrm{d}x\right)^{p-1} = \left(\frac{t-p+1}{p}\right)^{p-1}\frac{b^{-t}}{b^{\frac{-t+p-1}{p}}},\end{aligned}$$

which proves the lemma. □

Lemma A.3. *Let $\tau_1, \tau_2 \in [0,1]$ and $p > 1$. Then,*

$$|\tau_1^p - \tau_2^p| \leq p|\tau_1 - \tau_2|\max\{\tau_1^{p-1}, \tau_2^{p-1}\}.$$

Proof. The desired inequality follows from the convexity of the function $f(\tau) = \tau^p$. □

Lemma A.4. *Let $a, b > 0$, $\tau_1, \tau_2 \in [0,1]$, and $t > p-1 > 0$. Then,*

$$\min\{\tau_1^p, \tau_2^p\}\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p \geq 2^{1-p}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p - |\tau_1 - \tau_2|^p \max\{a^{-t+p-1}, b^{-t+p-1}\} \tag{A.1}$$

and

$$\max\{\tau_1^p, \tau_2^p\}\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p \leq 2^{p-1}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p + 2^{p-1}|\tau_1 - \tau_2|^p \max\{a^{-t+p-1}, b^{-t+p-1}\}. \tag{A.2}$$

Proof. For (A.1), we assume that $\tau_1 \geq \tau_2$. Then, we obtain from

$$\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}} = \tau_2\left(a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right) + (\tau_1 - \tau_2)a^{\frac{-t+p-1}{p}}$$

that

$$\begin{aligned} 2^{1-p}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p &\leq \tau_2^p\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p + |\tau_1 - \tau_2|^p a^{-t+p-1} \\ &\leq \tau_2^p\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p + |\tau_1 - \tau_2|^p \max\{a^{-t+p-1}, b^{-t+p-1}\}, \end{aligned}$$

from which (A.1) follows. The other case $\tau_1 < \tau_2$ can be proved in the same way.

For (A.2), we assume that $\tau_1 \geq \tau_2$. Then,

$$\begin{aligned} \max\{\tau_1^p, \tau_2^p\}\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p &= \left|\left(\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right) + (\tau_2 - \tau_1)b^{\frac{-t+p-1}{p}}\right|^p \\ &\leq 2^{p-1}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p + 2^{p-1}|\tau_2 - \tau_1|^p b^{-t+p-1}. \end{aligned}$$

The proof for the case $\tau_1 < \tau_2$ is the same. □

Proof of Lemma 3.4. We may assume that $b > a$. We begin with the equality

$$|b-a|^{p-2}(b-a)(\tau_1^p a^{-t} - \tau_2^p b^{-t}) = (b-a)^{p-1}(a^{-t} - b^{-t})\tau_1^p + (b-a)^{p-1}b^{-t}(\tau_1^p - \tau_2^p) =: A + B. \tag{A.3}$$

By Lemma A.1 and (A.1), we have

$$\begin{aligned} A &\geq t\left(\frac{p}{t-p+1}\right)^p \left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p \min\{\tau_1^p, \tau_2^p\} \\ &\geq t\left(\frac{p}{t-p+1}\right)^p\left(2^{1-p}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p - |\tau_1 - \tau_2|^p \max\{a^{-t+p-1}, b^{-t+p-1}\}\right). \end{aligned} \tag{A.4}$$

For B, we use Lemmas A.2, A.3, and Young's inequality to obtain

$$\begin{aligned} B &\geq -p\left(\frac{p}{t-p+1}\right)^{p-1}\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^{p-1} b^{\frac{-t+p-1}{p}}|\tau_1 - \tau_2|\max\{\tau_1^{p-1}, \tau_2^{p-1}\} \\ &\geq -(p-1)\left(\frac{p}{t-p+1}\right)^p \varepsilon^{p/(p-1)}\left|a^{\frac{-t+p-1}{p}} - b^{\frac{-t+p-1}{p}}\right|^p \max\{\tau_1^p, \tau_2^p\} - \frac{1}{\varepsilon^p} b^{-t+p-1}|\tau_1 - \tau_2|^p \end{aligned}$$

for any $\varepsilon > 0$. By using (A.2), we have

$$\begin{aligned} B &\geq -2^{p-1}(p-1)\left(\frac{p}{t-p+1}\right)^p \varepsilon^{p/(p-1)}\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p \\ &\quad - \left(2^{p-1}(p-1)\left(\frac{p}{t-p+1}\right)^p \varepsilon^{p/(p-1)} + \frac{1}{\varepsilon^p}\right)|\tau_1 - \tau_2|^p \max\{a^{-t+p-1}, b^{-t+p-1}\}. \end{aligned} \tag{A.5}$$

Combining (A.3)–(A.5), and then taking ε so that $2^{p-1}\varepsilon^{p/(p-1)} = 2^{1-p}$, we arrive at

$$|b-a|^{p-2}(b-a)(\tau_1^p a^{-t} - \tau_2^p b^{-t}) \geq c_1\left|\tau_1 a^{\frac{-t+p-1}{p}} - \tau_2 b^{\frac{-t+p-1}{p}}\right|^p - c_2|\tau_1 - \tau_2|^p(a^{-t+p-1} + b^{-t+p-1}),$$

where

$$c_1 = \frac{2^{1-p}p^p}{(t-p+1)^{p-1}} \quad \text{and} \quad c_2 = (t + 2^{1-p}(p-1))\left(\frac{p}{t-p+1}\right)^p + 2^{2(p-1)^2}.$$

Note that c_1 and c_2 are bounded when t is bounded away from $p-1$. □

B Anisotropic dyadic rectangles

Let us briefly sketch the construction of anisotropic "dyadic" rectangles. These objects can be used to prove the lower bound in L^p for the sharp maximal function $\mathbf{M}^\sharp u$.

We construct anisotropic dyadic rectangles having the following properties:

(i) For each integer $k \in \mathbb{Z}$, a countable collection $\{Q_{k,\alpha}\}_\alpha$ covers the whole space $\mathbb{R}^n$.

(ii) Each Q_k ($=Q_{k,\alpha}$ for some α) has an interior of the form $\operatorname{Int}(Q_k) = M_{2^{-k}}(x)$. We call Q_k an *anisotropic dyadic rectangle of generation k*.

(iii) Every $Q_{k,\alpha}$ is contained in $Q_{k-1,\beta}$ for some β. We call $Q_{k-1,\beta}$ a *predecessor of $Q_{k,\alpha}$*.

(iv) If $Q_{k,0}, \dots, Q_{k,2^n}$ are $2^n + 1$ different anisotropic dyadic rectangles of generation k, then $\cap_{i=0}^{2^n} Q_{k,i} = \varnothing$.

(v) If $2^n + 1$ different anisotropic dyadic rectangles $Q_{k_0}, \dots, Q_{k_{2^n}}$, $k_0 \leq \dots \leq k_{2^n}$, have a nonempty intersection, then $Q_j \subset Q_i$ for some $0 \leq i < j \leq 2^n$.

Remark B.1.

(1) A predecessor may not be unique.

(2) 2^n different anisotropic dyadic rectangles from the same generation may have a nonempty intersection.

Such a family of anisotropic dyadic rectangles can be easily constructed. Since the sets are rectangles, it is sufficient to exemplify the construction in one dimension. Let $Q_0 = [0, 1)$. Then, a countable collection $\{Q_0 + z\}_{z\in\mathbb{Z}}$ constitutes the zeroth generation. Let $N = \lfloor 2^{s_{\max}/s_1} \rfloor$. In order to construct the first generation, we take a disjoint family of (left-closed and right-opened) N intervals in Q_0 starting from 0 with length $2^{-s_{\max}/s_1}$ such that the following interval starts at the endpoint of the previous interval. If the right-endpoint of the last interval is 1, then these intervals constitute the first generation, and there is nothing to do. Thus, we assume from now on that $2^{s_{\max}/s_1} \notin \mathbb{Z}$. In this case, we add an interval $[1 - 2^{-s_{\max}/s_i}, 1)$ so that

$$Q_0 = \left(\bigcup_{i=0}^{N-1} Q_{1,i} \right) \cup Q_{1,N},$$

where

$$Q_{1,i} = [i2^{-s_{\max}/s_1}, (i+1)2^{-s_{\max}/s_1}) \quad \text{for } i = 0, \dots, N-1, \qquad Q_{1,N} = [1 - 2^{-s_{\max}/s_i}, 1),$$

and $N = \lfloor 2^{s_{\max}/s_1} \rfloor$. Then, the collection $\{Q_{1,i} + z\}_{0\leq i\leq N, z\in\mathbb{Z}}$ forms the first generation of intervals satisfying (i)–(iv) (Figure A1).

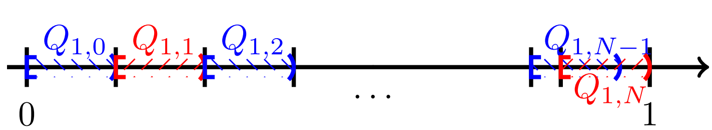

Figure A1: This figure shows the construction of the family $Q_{1,i}$.

We continue to construct the intervals of generation 2 that fill in $Q_{1,i}$ for each $0 \leq i \leq N-2$. However, we have to be careful in filling in $Q_{1,N-1}$ and $Q_{1,N}$ since $Q_{1,N-1} \cap Q_{1,N} \neq \varnothing$. Suppose that we filled in $Q_{1,N-1}$ and $Q_{1,N}$ as above, i.e.,

$$Q_{1,N-1} = \left(\bigcup_{i=0}^{N-1} Q_{2,i} \right) \cup Q_{2,N} \quad \text{and} \quad Q_{1,N} = \left(\bigcup_{i=0}^{N-1} \tilde{Q}_{2,i} \right) \cup \tilde{Q}_{2,N}$$

for some intervals $Q_{2,i}$ and $\tilde{Q}_{2,i}$, $0 \leq i \leq N$, of length $4^{-s_{\max}/s_1}$. Let K be the smallest integer such that $\overline{Q_{2,K}} \cap Q_{1,N} \neq \varnothing$. Then, we have

$$Q_{1,N-1} \cup Q_{1,N} = \left(\bigcup_{i=0}^{K} Q_{2,i} \right) \cup \left(\bigcup_{i=0}^{N-1} \tilde{Q}_{2,i} \right) \cup \tilde{Q}_{2,N}$$

and at most two different intervals among $\{Q_{2,0}, \dots, Q_{2,K}, \tilde{Q}_{2,0}, \dots, \tilde{Q}_{2,N}\}$ can intersect. Therefore, these intervals constitute the second generation satisfying (i)–(iv) (Figure A2).

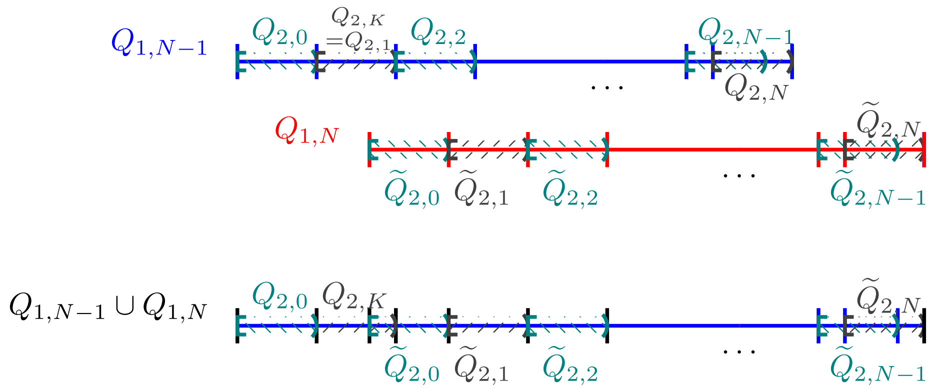

Figure A2: This figure shows the construction of generation 2.

In this way, we construct intervals of generation k for all $k \geq 0$.

Let us now construct intervals of generation $k < 0$. It is easy to observe that a collection $\{Q_{-1} + Nz\}_{z \in \mathbb{Z}}$ of intervals of generation -1 satisfies (i)–(iv), where

$$Q_{-1} = [0, 2^{s_{\max}/s_1}).$$

For the generation of -2, let K be the largest integer such that $Q_{-1} + NK \subset Q_{-2}$, where $Q_{-2} = [0, 4^{s_{\max}/s_1})$. Then, the intervals $Q_{-2} + N(K+1)z$, $z \in \mathbb{Z}$, form the generation of -2, which satisfying (i)–(iv). We continue this process to construct intervals of all generations $k < 0$.

We show that the intervals constructed in this way satisfy the property (v) as well. Suppose that three different intervals Q_{k_0}, Q_{k_1}, and Q_{k_2}, $k_0 \leq k_1 \leq k_2$, have a nonempty intersection. If $k_1 = k_2$, then $Q_{k_1} \subset Q_{k_0}$ or $Q_{k_2} \subset Q_{k_0}$. If $k_1 < k_2$, then either $Q_{k_2} \subset Q_{k_1}$ or not. In the former case, we are done. In the latter case, $Q_{k_2} \subset \tilde{Q}_{k_1}$ for some $\tilde{Q}_{k_1} \neq Q_{k_1}$, which reduces to the case $k_1 = k_2$.

C Sharp maximal function theorem

In this section, we prove Theorem 2.6 by using the anisotropic dyadic rectangles. For $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we define a dyadic maximal function $\mathbf{M}_d u$ by

$$\mathbf{M}_d u(x) = \sup_{x \in Q} \fint_Q |u(y)| \, dy,$$

where the supremum is taken over all anisotropic dyadic rectangles Q. Since

$$\mathbf{M}_d u \leq \mathbf{M} u, \tag{C.1}$$

Theorem 2.5 also holds for the dyadic maximal function $\mathbf{M}_d u$. We first prove a good-lambda estimate using the dyadic maximal function. See [31, Theorem 3.4.4].

Theorem C.1. *Let $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0,1)$. There exists a constant $C = C(n, s_0) > 0$ such that*

$$|\{x \in \mathbb{R}^n : \mathbf{M}_d u(x) > 2\lambda, \mathbf{M}^\sharp u(x) \leq \gamma\lambda\}| \leq C\gamma |\{x \in \mathbb{R}^n : \mathbf{M}_d u(x) > \lambda\}|$$

for all $\gamma > 0$, $\lambda > 0$, and $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$.

Proof. Let $\Omega_\lambda = \{x \in \mathbb{R}^n : M_d(f)(x) > \lambda\}$. We may assume that $|\Omega_\lambda| < +\infty$ since otherwise there is nothing to prove. For each $x \in \Omega_\lambda$, we find a maximal anisotropic dyadic rectangle Q^x such that

$$x \in Q^x \subset \Omega_\lambda \quad \text{and} \quad \fint_{Q^x} |f| > \lambda. \tag{C.2}$$

There are at most 2^n different maximal anisotropic dyadic rectangles of the same generation satisfying (C.2), but we can still choose anyone of them. Let Q_j be the collection of all such rectangles Q^x for all $x \in \Omega_\lambda$. Then, we have $\Omega_\lambda = \cup_j Q_j$. Note that different rectangles Q_j may have an intersection, but the intersection is contained in at most 2^n different maximal rectangles of the same generation. This is a consequence of the properties (iv) and (v) of anisotropic dyadic rectangles. Hence,

$$\sum_j |Q_j| \le 2^n |\Omega_\lambda|.$$

Therefore, the desired result follows once we have

$$|\{x \in Q_j : \mathbf{M}_d u(x) > 2\lambda, \mathbf{M}^\sharp u(x) \le \gamma\lambda\}| \le C\gamma |Q_j| \tag{C.3}$$

for some $C = C(n, s_0)$. Indeed, one can prove (C.3) by following the second paragraph of the proof of [31, Theorem 3.4.4], using Theorem 2.5 for $\mathbf{M}_d$, and replacing [31, equation (3.4.8)] by

$$\frac{1}{\lambda} \int_{Q_j} |u(y) - (u)_{Q_j'}| \mathrm{d}y \le \frac{2^{ns_{\max}/\bar{s}}}{\lambda} \frac{|Q_j|}{|Q_j'|} \int_{Q_j'} |u(y) - (u)_{Q_j'}| \mathrm{d}y \le \frac{2^{n/s_0}}{\lambda} |Q_j| \mathbf{M}^\sharp u(\xi_j)$$

for all $\xi_j \in Q_j$, where Q_j' is anyone of predecessors of Q_j. □

Theorem C.2. *Let $s_1, \dots, s_n \in [s_0, 1)$ be given for some $s_0 \in (0, 1)$, and let $0 < p_0 \le p < \infty$. Then, there is a constant $C = C(n, p, s_0) > 0$ such that for all functions $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ with $\mathbf{M}_d u \in L^{p_0}(\mathbb{R}^n)$, we have*

$$\|\mathbf{M}_d u\|_{L^p(\mathbb{R}^n)} \le C \|\mathbf{M}^\sharp u\|_{L^p(\mathbb{R}^n)}.$$

Theorem C.2 can be proved in the same way as in the proof of [31, Theorem 3.4.5] except that we use Theorem C.1 instead of [31, Theorem 3.4.4]. Finally, we combine the inequality

$$\|u\|_{L^p(\mathbb{R}^n)} \le \|\mathbf{M}_d u\|_{L^p(\mathbb{R}^n)},$$

which comes from the Lebesgue differentiation theorem and Theorem C.2 to conclude Theorem 2.6. See [31, Corollary 3.4.6].

D Pointwise convergence of the fractional orthotropic p-Laplacian

This section provides the proof of pointwise convergence of the fractional orthotropic p-Laplacian as $s \nearrow 1$.

Proposition D.1. *Let $u \in C^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ be such that $\partial_i u(x) \ne 0$ for all $i = 1, \dots, n$. Let $s_i = s$ for all $i = 1, \dots, n$. Let L be the operator in (1.2) with $\mu = \mu_{\mathrm{axes}}$ and A^p_{loc} be as in (1.3). Then, $\mathrm{L}u(x) \to A^p_{\mathrm{loc}} u(x)$ as $s \nearrow 1$ up to a constant.*

Proof. Let us fix a point $x \in \mathbb{R}^n$ with $\partial_i u(x) \ne 0$. For each $i = 1, \dots, n$, let us define $u_i : \mathbb{R} \to \mathbb{R}$ by $u_i(x_i) = u(x_1, \dots, x_i, \dots, x_n)$ as a function of one variable. Then, $u_i \in C^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $u'(x_i) \ne 0$. We write

$$\mathrm{L}u(x) = \sum_{i=1}^n s(1-s) \int_{\mathbb{R}} \frac{|u_i(y_i) - u(x_i)|^{p-2}(u_i(y_i) - u_i(x_i))}{|x_i - y_i|^{1+sp}} \mathrm{d}y_i = -\sum_{i=1}^n (-\partial^2)^s_p u_i(x_i),$$

which is the sum of one-dimensional fractional p-Laplacians. By [11, Theorem 2.8], we have

$$-(-\partial^2)^s_p u_i(x_i) \to \frac{\mathrm{d}}{\mathrm{d}x_i}\left(\left| \frac{\mathrm{d}u_i}{\mathrm{d}x_i}(x_i) \right|^{p-2} \frac{\mathrm{d}u_i}{\mathrm{d}x_i}(x_i) \right)$$

as $s \nearrow 1$, for each $i = 1, \dots, n$, up to a constant depending on p only. Consequently, by summing up, $\mathrm{L}u(x) \to A^p_{\mathrm{loc}} u(x)$ as $s \nearrow 1$. □