DE GRUYTER Forum Math. 2016; 28 (5):823–856

Research Article

Jun Cao, Svitlana Mayboroda and Dachun Yang*

Maximal function characterizations of Hardy spaces associated to homogeneous higher order elliptic operators

Abstract: Let L be a homogeneous divergence form higher order elliptic operator with complex bounded measurable coefficients and $(p_-(L), p_+(L))$ be the maximal interval of exponents $q \in [1, \infty]$ such that the semigroup $\{e^{-tL}\}_{t>0}$ is bounded on $L^q(\mathbb{R}^n)$. In this article, the authors establish the non-tangential maximal function characterizations of the associated Hardy spaces $H^p_L(\mathbb{R}^n)$ for all $p \in (0, p_+(L))$, which when $p = 1$, answers a question asked by Deng, Ding and Yao in [21]. Moreover, the authors characterize $H^p_L(\mathbb{R}^n)$ via various versions of square functions and Lusin-area functions associated to the operator L.

Keywords: Higher order elliptic operator, off-diagonal estimate, Hardy space, maximal function, square function, molecule, Riesz transform

MSC 2010: Primary 42B30; secondary 42B20, 42B35, 42B37, 46E30, 35J48, 47B06, 47B38

DOI: 10.1515/forum-2014-0127
Received July 15, 2014; revised April 21, 2015

Communicated by: Christopher D. Sogge

1 Introduction

Let $m \in \mathbb{N} := \{1, 2, \dots\}$ and L be a homogeneous higher order elliptic operator of the form

$$L := \sum_{|\alpha|=m=|\beta|} (-1)^m \partial^\alpha (a_{\alpha,\beta} \partial^\beta), \tag{1.1}$$

where $\alpha := (\alpha_1, \dots, \alpha_n)$ and $\beta := (\beta_1, \dots, \beta_n)$ belong to $(\mathbb{Z}_+)^n := (\mathbb{N} \cup \{0\})^n$ with $|\alpha| := \alpha_1 + \cdots + \alpha_n$ and $|\beta| := \beta_1 + \cdots + \beta_n$, and $\{a_{\alpha,\beta}\}_{|\alpha|=m=|\beta|}$ are bounded measurable functions mapping $\mathbb{R}^n$ into $\mathbb{C}$ (see Section 2.1 below for the exact definition of L in (1.1)). The aim of this article is to establish the maximal function characterizations of the associated Hardy space $H^p_L(\mathbb{R}^n)$ adapted to L, which when $p = 1$, answers a question asked by Deng, Ding and Yao in [21]. It is now well known that such a Hardy space adapted to L, which has appeared in [15, 21], is a good substitute of the Lebesgue space $L^p(\mathbb{R}^n)$, for smaller p, when studying the regularity of the solution to the corresponding elliptic equation (see, for example, [13, 14, 16–19, 22–25, 32, 34]).

Notice that, if

$$L \equiv -\Delta := -\sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$$

is the Laplace operator, then the Hardy space $H^p_{-\Delta}(\mathbb{R}^n)$ is just the classical Hardy space $H^p(\mathbb{R}^n)$ which has been

Jun Cao: School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, P. R. China. Current address: Department of Applied Mathematics, Zhejiang University of Technology, Hangzhou 310023, e-mail: caojun1860@zjut.edu.cn
Svitlana Mayboroda: School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA, e-mail: svitlana@math.umn.edu
***Corresponding author: Dachun Yang:** School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, P. R. China, e-mail: dcyang@bnu.edu.cn

systematically studied by Fefferman and Stein in their seminal paper [26]. In the same paper, Fefferman and Stein also established various real-variable characterizations of $H^p(\mathbb{R}^n)$, including their non-tangential maximal function characterization and Littlewood–Paley function characterizations. Recall that, for all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, the *non-tangential maximal function* $\mathcal{N}_\Delta(f)(x)$ is defined by

$$\mathcal{N}_\Delta(f)(x) := \sup_{(y,t)\in\Gamma(x)} |e^{-t\sqrt{\Delta}}(f)(y)|. \tag{1.2}$$

Here and hereafter, for all $x \in \mathbb{R}^n$,

$$\Gamma(x) := \{(y,t) \in \mathbb{R}^{n+1}_+ : |y-x| < t\} \tag{1.3}$$

denotes the *cone with vertex* x and $\mathbb{R}^{n+1}_+ := \mathbb{R}^n \times (0,\infty)$. Recall also that, for $n = 1$, the non-tangential maximal function characterization of $H^p(\mathbb{R}^n)$ was proved earlier by Burkholder, Gundy and Silverstein [11], which constitutes one of the motivations for Fefferman and Stein to study the real-variable theory of $H^p(\mathbb{R}^n)$.

Let $L \equiv -\operatorname{div}(A\nabla)$ be the second order elliptic operator, where $\nabla := (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ and $A := A(x)$ is an $n \times n$ matrix of complex bounded measurable coefficients defined on $\mathbb{R}^n$ and satisfies the *ellipticity condition*

$$\lambda|\xi|^2 \le \operatorname{Re}(A\xi\cdot\overline{\xi}) \quad \text{and} \quad |A\xi\cdot\overline{\zeta}| \le \Lambda|\xi||\zeta|$$

for all $\xi, \zeta \in \mathbb{C}^n$ and for some positive constants $0 < \lambda \le \Lambda < \infty$ independent of ξ and ζ. Hofmann and Mayboroda [28] (for $p = 1$), and Jiang and Yang [31] (for $p \in (0,1]$) established the non-tangential maximal function characterization of the associated Hardy space $H^p_L(\mathbb{R}^n)$. Recall that, for all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, the *non-tangential maximal function* $\mathcal{N}_L(f)$ is defined by

$$\mathcal{N}_L(f)(x) := \sup_{(y,t)\in\Gamma(x)} \left\{\frac{1}{t^n}\int_{B(y,t)} |e^{-t^2L}(f)(z)|^2\,dz\right\}^{\frac{1}{2}}. \tag{1.4}$$

Here and hereafter, for all $(y,t) \in \mathbb{R}^{n+1}_+$, $B(y,t) := \{z \in \mathbb{R}^n : |z-y| < t\}$. Observe that the non-tangential maximal function (1.4) is a little bit different from (1.2). The main reason for adding an extra averaging in the spatial variable in (1.4) is that we need to compensate for the lack of pointwise estimates of the heat semigroup (see [28] for more details).

Now, let L be a homogenous $2m$-th order elliptic operator as in (1.1), where $\{a_{\alpha,\beta}\}_{|\alpha|=m=|\beta|}$ are bounded measurable functions mapping $\mathbb{R}^n$ into $\mathbb{C}$ satisfying the ellipticity condition $(\mathcal{E}_0)$ or the strong ellipticity condition $(\mathcal{E}_1)$ (see Section 2.1 for their definitions). Some properties of the Hardy space $H^p_L(\mathbb{R}^n)$ associated with a homogeneous higher order elliptic operator L as in (1.1), for $p \in (0,1]$, have already been established in [15, 21]. To be precise, let L be the homogeneous higher order operator defined as in (1.1) that satisfies the ellipticity condition $(\mathcal{E}_0)$. For all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, the *L-adapted square function* $S_L(f)$ is defined by

$$S_L(f)(x) := \left\{\iint_{\Gamma(x)} |t^{2m}Le^{-t^{2m}L}(f)(y)|^2\,\frac{dy\,dt}{t^{n+1}}\right\}^{\frac{1}{2}}. \tag{1.5}$$

The following definition of Hardy spaces is motivated by [28, 29, 31]; see also [15, 21] for the case when $p \in (0,1]$.

Definition 1.1. Let $p \in (0,2]$ and L be as in (1.1) satisfying the ellipticity condition $(\mathcal{E}_0)$. A function $f \in L^2(\mathbb{R}^n)$ is said to be in the space $\mathbb{H}^p_L(\mathbb{R}^n)$ if $S_L(f) \in L^p(\mathbb{R}^n)$. Moreover, let $\|f\|_{H^p_L(\mathbb{R}^n)} := \|S_L(f)\|_{L^p(\mathbb{R}^n)}$. The Hardy space $H^p_L(\mathbb{R}^n)$ is then defined as the completion of $\mathbb{H}^p_L(\mathbb{R}^n)$ with respect to the *quasi-norm* $\|\cdot\|_{H^p_L(\mathbb{R}^n)}$.

For $p \in (2,\infty)$, the Hardy space $H^p_L(\mathbb{R}^n)$ is then defined as the dual space of the Hardy space $H^{p'}_{L^*}(\mathbb{R}^n)$, where L^* denotes the *adjoint operator* of L in $L^2(\mathbb{R}^n)$ and $p' := \frac{p}{p-1} \in (1,2)$ denotes the *conjugate exponent* of p.

For the Hardy space $H^p_L(\mathbb{R}^n)$ with $p \in (0,1]$, the authors in [15] established various characterizations of $H^p_L(\mathbb{R}^n)$ in terms of molecules, the generalized square function or the Riesz transform. Moreover, Deng, Ding and Yao in [21] also established some other interesting characterizations of these Hardy spaces for $p = 1$.

However, neither of the above articles gives the maximal function characterizations of $H_L^p(\mathbb{R}^n)$ for $p = 1$ and it has been raised by Deng, Ding and Yao [21] as an open question whether $H_L^1(\mathbb{R}^n)$ has the maximal function characterizations or not.

Motivated by the above articles, the main purpose of this article is to establish the maximal function characterizations of the Hardy space $H_L^p(\mathbb{R}^n)$ associated with L as in (1.1). Based on [28], we first introduce the following versions of maximal functions associated with L. For $\lambda \in (0,\infty)$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, the *radial maximal function* $\mathcal{R}_{h,L}^\lambda(f)$, associated with the heat semigroup generated by L, is defined by

$$\mathcal{R}_{h,L}^\lambda(f)(x) := \sup_{t\in(0,\infty)} \left\{ \frac{1}{(\lambda t)^n} \int_{B(x,\lambda t)} |e^{-t^{2m}L}(f)(y)|^2\, dy \right\}^{\frac{1}{2}}. \tag{1.6}$$

Similarly, the *non-tangential maximal function* $\mathcal{N}_{h,L}^\lambda(f)$, associated with the heat semigroup generated by L, is defined by

$$\mathcal{N}_{h,L}^\lambda(f)(x) := \sup_{(y,t)\in\Gamma^\lambda(x)} \left\{ \frac{1}{(\lambda t)^n} \int_{B(y,\lambda t)} |e^{-t^{2m}L}(f)(z)|^2\, dz \right\}^{\frac{1}{2}}, \tag{1.7}$$

where, for all $x \in \mathbb{R}^n$,

$$\Gamma^\lambda(x) := \{(y,t) \in \mathbb{R}_+^{n+1} : |y - x| < \lambda t\}. \tag{1.8}$$

In what follows, when $\lambda = 1$, we remove the superscript λ from $\mathcal{R}_{h,L}^\lambda(f)$ and $\mathcal{N}_{h,L}^\lambda(f)$ for simplicity. Observe also that, if $m = 1$, then the maximal functions defined in (1.6) and (1.7) coincide with those in [28, 31].

Definition 1.2. Let L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$. For all $p \in (0,\infty)$, the Hardy space $H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n)$ is defined as the completion of $\{f \in L^2(\mathbb{R}^n) : \mathcal{N}_{h,L}(f) \in L^p(\mathbb{R}^n)\}$ with respect to the *quasi-norm*

$$\|f\|_{H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n)} := \|\mathcal{N}_{h,L}(f)\|_{L^p(\mathbb{R}^n)}.$$

The Hardy space $H_{\mathcal{R}_{h,L}}^p(\mathbb{R}^n)$ is defined in the way same as $H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n)$ with $\mathcal{N}_{h,L}(f)$ replaced by $\mathcal{R}_{h,L}(f)$.

Remark 1.3. By the argument that has been used in the proof of [28, equation (6.50)] with a small modification, we know that, for all $p \in (0,\infty)$ and $f \in L^2(\mathbb{R}^n)$,

$$\|\mathcal{N}_{h,L}(f)\|_{L^p(\mathbb{R}^n)} \sim \|\mathcal{R}_{h,L}(f)\|_{L^p(\mathbb{R}^n)},$$

which implies that, for all $p \in (0,\infty)$, $H_{\mathcal{R}_{h,L}}^p(\mathbb{R}^n) = H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n)$ with equivalent quasi-norms.

Now, let $(p_-(L), p_+(L))$ be the *maximal interval* of exponents $q \in [1,\infty]$ such that the family $\{e^{-tL}\}_{t>0}$ of operators is bounded on $L^q(\mathbb{R}^n)$. The following theorem gives the maximal function characterizations of $H_L^p(\mathbb{R}^n)$ for all $p \in (0, p_+(L))$.

Theorem 1.4. *Let L be as in* (1.1) *and satisfying the strong ellipticity condition* $(\mathcal{E}_1)$ *(see Section 2.1 below for its definition). Then, for all $p \in (0, p_+(L))$, $H_L^p(\mathbb{R}^n) = H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n) = H_{\mathcal{R}_{h,L}}^p(\mathbb{R}^n)$ with equivalent quasi-norms, where $H_{\mathcal{N}_{h,L}}^p(\mathbb{R}^n)$ and $H_{\mathcal{R}_{h,L}}^p(\mathbb{R}^n)$ are defined as in Definition 1.2.*

The proof of Theorem 1.4 will be given in Section 3.

Before describing our method to prove Theorem 1.4, let us first recall some key points of the methods used to establish the maximal function characterizations in [26, 28].

For all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, let

$$S_\Delta(f)(x) := \left\{ \iint_{\Gamma(x)} |t\nabla e^{-t\sqrt{\Delta}}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}} \tag{1.9}$$

be the *Lusin-area function* of f associated to Δ, where for all $x \in \mathbb{R}^n$, $\Gamma(x)$ is as in (1.3). For convenience, throughout the article, we distinguish in terminology the *square function with gradient* from the one without gradient via calling the former the *Lusin-area function*.

Recall that Fefferman and Stein [26] established the maximal function characterizations of $H^p(\mathbb{R}^n)$ by developing the equivalence of the $L^p(\mathbb{R}^n)$ quasi-norms between $\mathcal{N}_\Delta(f)$ in (1.2) and $S_\Delta(f)$ in (1.9). The heart of their proof is to control the integral $\int_E [S_\Delta(f)(x)]^2\,dx$ for some set E. By Fubini's theorem, this is reduced to the corresponding estimates on a saw-tooth region $\mathcal{R} := \bigcup_{x\in E}\Gamma(x)$ based on E, namely, we need to control

$$\iint\limits_{\mathcal{R}} t|\nabla e^{-t\sqrt{\Delta}}(f)(y)|^2\,dy\,dt. \tag{1.10}$$

The main tool that they used to estimate (1.10) is Green's theorem. To this end, they first replaced the region $\mathcal{R}$ by an approximating family $\{\mathcal{R}_\epsilon\}_{\epsilon>0}$ of regions whose boundaries have certain uniform smoothness, and then applied Green's theorem to reduce the estimates on $\mathcal{R}_\epsilon$ to its boundary. Finally, they used some properties of harmonic functions to estimate the corresponding integral on the boundary.

Hofmann and Mayboroda [28] used the strategy similar to that of Fefferman and Stein [26]. However, there are some differences between these two methods. For all $f\in L^2(\mathbb{R}^n)$ and $x\in\mathbb{R}^n$, let

$$S_L(f)(x) := \left\{\iint\limits_{\Gamma(x)} |t^2Le^{-t^2L}(f)(y)|^2\,\frac{dy\,dt}{t^{n+1}}\right\}^{\frac{1}{2}} \tag{1.11}$$

be the square function of f associated to L, which was used in [28] to introduce the Hardy space $H^p_L(\mathbb{R}^n)$ associated with $L \equiv -\operatorname{div}(A\nabla)$. Notice that this square function, which is more convenient when introducing $H^p_L(\mathbb{R}^n)$, is different from the Lusin-area function (1.9). Then, to obtain the non-tangential maximal characterization of $H^p_L(\mathbb{R}^n)$, Hofmann and Mayboroda [28] used Caccioppoli's inequality to control $S_L(f)$ by another Lusin-type area function defined in a way similar to (1.9) with $e^{-t\sqrt{\Delta}}$ replaced by $e^{-t\sqrt{L}}$. Furthermore, they used the truncated cone to approximate the cone in (1.11) before applying Fubini's theorem. This reduces to estimating the following integral:

$$\iint\limits_{\mathcal{R}^{\alpha\epsilon,\alpha\epsilon,1/\alpha}(E^*)} t|\nabla e^{-t^2L}(f)(y)|^2\,dy\,dt, \tag{1.12}$$

where $\mathcal{R}^{\alpha\epsilon,\alpha\epsilon,1/\alpha}(E^*)$ denotes a truncated saw-tooth region. Finally, in the estimate of (1.12), since $e^{-t^2L}(f)$ is no longer a harmonic function and hence Green's theorem cannot be used directly, Hofmann and Mayboroda [28] made full use of the ellipticity condition of the operator $-\operatorname{div}(A\nabla)$ and the divergence theorem to reduce the corresponding estimates to the boundary of $\mathcal{R}^{\alpha\epsilon,\alpha\epsilon,1/\alpha}(E^*)$.

In the present article, to prove Theorem 1.4, we first point out that the proof of the inclusion

$$H^p_L(\mathbb{R}^n) \subset H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)$$

is relatively easy. Indeed, for $p\in(0,1]$, by the molecular characterization of $H^p_L(\mathbb{R}^n)$ (see Theorem 2.11 below), we only need to consider the action of the non-tangential maximal function $\mathcal{N}_{h,L}$ on each molecule of $H^p_L(\mathbb{R}^n)$. For $p\in[2,p_+(L))$, using the $L^2(\mathbb{R}^n)$ off-diagonal estimates, we show that the radial maximal function $\mathcal{R}_{h,L}$ is bounded on $L^p(\mathbb{R}^n)$, which together with relations between $H^p_L(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ (see Lemma 2.12 below), the complex interpolation of $H^p_L(\mathbb{R}^n)$ (see Proposition 2.8 below) and Remark 1.3, implies, for all $p\in(1,p_+(L))$, that $\mathcal{N}_{h,L}$ is bounded from $H^p_L(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. This furnishes the proof of the inclusion $H^p_L(\mathbb{R}^n)\subset H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)$.

For the proof of the converse inclusion of Theorem 1.4 when $p\in(0,2]$, we adapt the strategy of [26, 28]. The higher order setting produces new problems and requires new tools. To be precise, let S_L and $S_{h,L}$ be, respectively, the square function and the Lusin-area function as in (1.5) and (1.18). We obtain the converse inclusion by showing that, for all $p\in(0,2]$ and $f\in L^2(\mathbb{R}^n)$,

$$\|S_L(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|S_{h,L}(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|\mathcal{N}^\gamma_{h,L}(f)\|_{L^p(\mathbb{R}^n)}, \tag{1.13}$$

where $\gamma\in(0,\infty)$ and the implicit positive constants are independent of f. More precisely, in the proof of the first inequality of (1.13), we need a new higher order parabolic Caccioppoli's inequality (see (3.12) below).

To obtain (3.12), we first establish a parabolic Caccioppoli's inequality with gradient terms on the right-hand side of the inequality (see (3.2) below). Then, by an induction argument from Barton [7], we remove the gradient terms and obtain an improved Caccioppoli's inequality in (3.12). Also, in the proof of the second inequality of (1.13), in order to avoid the estimates on the boundary when applying the divergence theorem, we use some special cut-off functions. In this argument, the parabolic Caccioppoli's inequality (3.12) is also needed. The case $p \in (2, p_+(L))$ of the first inequality in (1.13) is obtained via duality, see Proposition 3.6 and Corollary 3.8 below.

With the help of Theorem 1.4, we point out that $H_L^p(\mathbb{R}^n)$ can also be characterized by another kind of maximal functions. To be precise, for $\lambda \in (0,\infty)$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, define the *radial maximal function*, $\tilde{\mathcal{R}}_{h,L}^\lambda(f)$, associated with the heat semigroup generated by L, by setting

$$\tilde{\mathcal{R}}_{h,L}^\lambda(f)(x) := \sup_{t\in(0,\infty)} \left\{ \frac{1}{(\lambda t)^n} \int_{B(x,\lambda t)} \sum_{k=0}^{m-1} |(t\nabla)^k e^{-t^{2m}L}(f)(y)|^2 \, dy \right\}^{\frac{1}{2}}. \tag{1.14}$$

Similarly, define the *non-tangential maximal function*, $\tilde{\mathcal{N}}_{h,L}^\lambda(f)$, associated with the heat semigroup generated by L, by setting

$$\tilde{\mathcal{N}}_{h,L}^\lambda(f)(x) := \sup_{(y,t)\in\Gamma^\lambda(x)} \left\{ \frac{1}{(\lambda t)^n} \int_{B(y,\lambda t)} \sum_{k=0}^{m-1} |(t\nabla)^k e^{-t^{2m}L}(f)(z)|^2 \, dz \right\}^{\frac{1}{2}}, \tag{1.15}$$

where for all $x \in \mathbb{R}^n$, $\Gamma^\lambda(x)$ is defined as in (1.8). In what follows, when $\lambda = 1$, we remove the superscript λ from $\tilde{\mathcal{R}}_{h,L}^\lambda(f)$ and $\tilde{\mathcal{N}}_{h,L}^\lambda(f)$ for simplicity. Note also that, if $m = 1$, (1.14) and (1.15) coincide with (1.6) and (1.7), respectively.

The following theorem characterizes $H_L^p(\mathbb{R}^n)$ via maximal functions defined as in (1.14) and (1.15).

Theorem 1.5. *Let L be as in* (1.1) *and satisfying the strong ellipticity condition* ($\mathcal{E}_1$). *Then, for all $p \in (0, p_+(L))$,*

$$H_L^p(\mathbb{R}^n) = H_{\tilde{\mathcal{N}}_{h,L}}^p(\mathbb{R}^n) = H_{\tilde{\mathcal{R}}_{h,L}}^p(\mathbb{R}^n)$$

with equivalent quasi-norms, where $H_{\tilde{\mathcal{N}}_{h,L}}^p(\mathbb{R}^n)$ and $H_{\tilde{\mathcal{R}}_{h,L}}^p(\mathbb{R}^n)$ are defined similarly as in Definition 1.2 with $\mathcal{N}_{h,L}$ and $\mathcal{R}_{h,L}$ therein replaced, respectively, by $\tilde{\mathcal{N}}_{h,L}$ and $\tilde{\mathcal{R}}_{h,L}$.

The proof of Theorem 1.5 will be given in Section 3.

Now we characterize $H_L^p(\mathbb{R}^n)$ by using the non-tangential maximal function with only the $(m-1)$-order gradients of the heat semigroup generated by L.

Let $\psi \in C_c^\infty(B(0,2))$ satisfy $0 \le \psi \le 1$, $\psi \equiv 1$ on $B(0,1)$ and, for all $k \in \{0,\dots,m\}$,

$$\|\nabla^k \psi\|_{L^\infty(\mathbb{R}^n)} \lesssim 1,$$

where the implicit positive constant may depend on ψ, m and n. For all $(x,t) \in \mathbb{R}_+^{n+1}$ and $y \in \mathbb{R}^n$, let

$$\psi_{x,t}(y) := \frac{1}{t^n}\psi\left(\frac{y-x}{t}\right). \tag{1.16}$$

Then, $\psi_{x,t} \in C_c^\infty(B(x,2t))$ with $0 \le \psi_{x,t} \le 1$, $\psi_{x,t} \equiv 1$ on $B(x,t)$ and, for all $k \in \{0,\dots,m\}$,

$$\|\nabla^k \psi_{x,t}\|_{L^\infty(\mathbb{R}^n)} \lesssim t^{-k}$$

with the implicit positive constant independent of x and t.

Having fixed any ψ as above, for any $f \in L^2(\mathbb{R}^n)$, $x \in \mathbb{R}^n$ and $\lambda \in (0,\infty)$, we introduce the following version of the *non-tangential maximal function*, $\mathcal{N}_{h,\psi,L}^\lambda(f)$, associated with the heat semigroup generated by L:

$$\mathcal{N}_{h,\psi,L}^\lambda(f)(x) := \sup_{(y,t)\in\Gamma^\lambda(x)} \left\{ \frac{1}{(\lambda t)^n} \int_{B(y,\lambda t)} |(t\nabla)^{m-1}(\psi_{x,t} e^{-t^{2m}L}(f))(z)|^2 \, dz \right\}^{\frac{1}{2}}.$$

When $\lambda = 1$, we remove the superscript λ from $\mathcal{N}_{h,\psi,L}^\lambda(f)$ for simplicity.

Proposition 1.6. *Let L be as in (1.1) and satisfying the strong ellipticity condition $(\mathcal{E}_1)$, and let ψ be a cut-off function defined as in (1.16). For any $p \in (0, p_+(L))$, denote by $H^p_{\mathcal{N}_{h,\psi,L}}(\mathbb{R}^n)$ the Hardy space defined as $H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)$ with $\mathcal{N}_{h,L}$ replaced by $\mathcal{N}_{h,\psi,L}$. Then,*

$$H^p_L(\mathbb{R}^n) = H^p_{\mathcal{N}_{h,\psi,L}}(\mathbb{R}^n)$$

with equivalent quasi-norms. In particular, different choices of ψ in the definition of $H^p_{\mathcal{N}_{h,\psi,L}}(\mathbb{R}^n)$ above yield equivalent quasi-norms.

The proof of Proposition 1.6 will be given in Section 3, where a higher order Poincaré's inequality from [35] is used.

By the method used in the proof of Theorem 1.4, we are able to characterize $H^p_L(\mathbb{R}^n)$ via some more general square functions and Lusin-area functions.

To be precise, for all $\lambda \in (0, \infty)$, $k \in \mathbb{Z}_+$ and $f \in L^2(\mathbb{R}^n)$, the *L-adapted square function* $S^\lambda_{L,k}(f)$ is defined by setting, for all $x \in \mathbb{R}^n$,

$$S^\lambda_{L,k}(f)(x) := \left\{ \iint_{\Gamma^\lambda(x)} |(t^{2m}L)^k e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}} \tag{1.17}$$

and the *Lusin-area function* $S^\lambda_{h,L,k}(f)$ by setting, for all $x \in \mathbb{R}^n$,

$$S^\lambda_{h,L,k}(f)(x) := \left\{ \iint_{\Gamma^\lambda(x)} |(t\nabla)^m (t^{2m}L)^k e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}}, \tag{1.18}$$

where Γ^λ is as in (1.8). For simplicity, if $k = 1$, we remove the subscript k from $S^\lambda_{L,k}(f)$ and, if $k = 0$, we remove the subscript k from $S^\lambda_{h,L,k}(f)$. Also, if $\lambda = 1$, we remove the superscript λ from both $S^\lambda_{L,k}(f)$ and $S^\lambda_{h,L,k}(f)$.

Definition 1.7. Let L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$. For all $k \in \mathbb{N}$ and $p \in (0, \infty)$, the Hardy space $H^p_{S_{L,k}}(\mathbb{R}^n)$ is defined as the completion of

$$\{f \in L^2(\mathbb{R}^n) : S_{L,k}(f) \in L^p(\mathbb{R}^n)\}$$

with respect to the quasi-norm

$$\|f\|_{H^p_{S_{L,k}}(\mathbb{R}^n)} := \|S_{L,k}(f)\|_{L^p(\mathbb{R}^n)}.$$

Moreover, for all $k \in \mathbb{Z}_+$ and $p \in (0, \infty)$, the Hardy space $H^p_{S_{h,L,k}}(\mathbb{R}^n)$ is defined in the way same as $H^p_{S_{L,k}}(\mathbb{R}^n)$ with $S_{L,k}(f)$ given in (1.17) replaced by $S_{h,L,k}(f)$ given in (1.18).

The following theorem establishes the characterization of $H^p_L(\mathbb{R}^n)$ via, respectively, some square functions and some Lusin-area functions.

Theorem 1.8. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Then, the following hold:*
(i) *For all $k \in \mathbb{N}$ and $p \in (0, p_+(L))$, $H^p_L(\mathbb{R}^n) = H^p_{S_{L,k}}(\mathbb{R}^n)$ with equivalent quasi-norms.*
(ii) *For all $k \in \mathbb{N}$ and $p \in (0, p_+(L))$, $H^p_L(\mathbb{R}^n) = H^p_{S_{h,L,k}}(\mathbb{R}^n)$ with equivalent quasi-norms.*

The proof of Theorem 1.8 will be given in Section 3.

Let us end this section by making some conventions on the notation. Throughout the paper, we always let $\mathbb{N} := \{1, 2, \dots\}$ and $\mathbb{Z}_+ := \mathbb{N} \cup \{0\}$. We use the shorthand notation

$$\partial^\alpha := \frac{\partial^\alpha}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}},$$

where $\alpha := (\alpha_1, \dots, \alpha_n)$ and $|\alpha| := \alpha_1 + \dots + \alpha_n$. We use C to denote a *positive constant* that is independent of the main parameters involved but whose value may differ from line to line, and $C_{(\alpha,\dots)}$ to denote a *positive constant* depending on the parameters $\alpha, \dots$. *Constants with subscripts*, such as C_1, do not change in different occurrences. If $f \le Cg$, we then write $f \lesssim g$, and if $f \lesssim g \lesssim f$, we then write $f \sim g$. For any $x \in \mathbb{R}^n$, $r \in (0, \infty)$ and $\lambda \in (0, \infty)$, we let $B(x, r) := \{y \in \mathbb{R}^n : |x - y| < r\}$ and $\lambda B := B(x, \lambda r)$. Also, for any set $E \subset \mathbb{R}^n$, χ_E denotes its *characteristic function* and, for all $z \in \mathbb{C}$, $\operatorname{Re} z$ denotes its *real part*.

2 The Hardy space $H_L^p(\mathbb{R}^n)$

In this section, we study the Hardy space $H_L^p(\mathbb{R}^n)$ associated with the homogeneous higher order elliptic operator L in (1.1). To this end, we first collect some known basic facts on L in Section 2.1. Then, in Section 2.2, we present some real-variable properties of the Hardy space $H_L^p(\mathbb{R}^n)$ associated with L for $p \in (0,\infty)$. Recall that, for $p \in (0,1]$, $H_L^p(\mathbb{R}^n)$ has been studied in [15, 21]. Our results here also include the case $p \in (1,\infty)$.

2.1 Homogeneous higher order elliptic operators

Let $m \in \mathbb{N}$ and $\dot{W}^{m,2}(\mathbb{R}^n)$ be the *m-order homogeneous Sobolev space* equipped with the usual norm

$$\|f\|_{\dot{W}^{m,2}(\mathbb{R}^n)} := \left[\sum_{|\alpha|=m} \|\partial^\alpha f\|^2_{L^2(\mathbb{R}^n)}\right]^{\frac{1}{2}}.$$

For all multi-indices α and β in $(\mathbb{Z}_+)^n$ satisfying $|\alpha| = m = |\beta|$, let $a_{\alpha,\beta}$ be a complex valued L^∞ function on $\mathbb{R}^n$. For all f and $g \in \dot{W}^{m,2}(\mathbb{R}^n)$, define the *sesquilinear form* $\mathfrak{a}_0$, mapping $\dot{W}^{m,2}(\mathbb{R}^n) \times \dot{W}^{m,2}(\mathbb{R}^n)$ into $\mathbb{C}$, by

$$\mathfrak{a}_0(f,g) := \sum_{|\alpha|=m=|\beta|} \int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta f(x)\overline{\partial^\alpha g(x)}\,dx. \tag{2.1}$$

The following ellipticity condition on $\{a_{\alpha,\beta}\}_{|\alpha|=m=|\beta|}$ is necessary.

Ellipticity condition $(\mathcal{E}_0)$: There exist constants $0 < \lambda_0 \le \Lambda_0 < \infty$ such that, for all f and $g \in \dot{W}^{m,2}(\mathbb{R}^n)$,

$$\left|\sum_{|\alpha|=m=|\beta|} \int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta f(x)\overline{\partial^\alpha g(x)}\,dx\right| \le \Lambda_0\|\nabla^m f\|_{L^2(\mathbb{R}^n)}\|\nabla^m g\|_{L^2(\mathbb{R}^n)}$$

and

$$\mathrm{Re}\left\{\sum_{|\alpha|=m=|\beta|} \int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta f(x)\overline{\partial^\alpha f(x)}\,dx\right\} \ge \lambda_0\,\|\nabla^m f\|^2_{L^2(\mathbb{R}^n)},$$

where

$$\|\nabla^m f\|_{L^2(\mathbb{R}^n)} := \left[\sum_{|\alpha|=m} \int_{\mathbb{R}^n} |\partial^\alpha f(x)|^2\,dx\right]^{\frac{1}{2}}.$$

We also need the following strong ellipticity condition on $\{a_{\alpha,\beta}\}_{|\alpha|=m=|\beta|}$.

Strong ellipticity condition $(\mathcal{E}_1)$: There exists a positive constant λ_1 such that, for all $\xi := \{\xi_\alpha\}_{|\alpha|=m}$ with $\xi_\alpha \in \mathbb{C}$ and almost every $x \in \mathbb{R}^n$,

$$\mathrm{Re}\left\{\sum_{|\alpha|=m=|\beta|} a_{\alpha,\beta}(x)\xi_\beta\overline{\xi_\alpha}\right\} \ge \lambda_1|\xi|^2 = \lambda_1\left\{\sum_{|\alpha|=m}|\xi_\alpha|^2\right\}.$$

Moreover, for all multi-indices α and β with $|\alpha| = m = |\beta|$, $a_{\alpha,\beta} \in L^\infty(\mathbb{R}^n)$.

Remark 2.1. It is easy to see that the strong ellipticity condition $(\mathcal{E}_1)$ implies the ellipticity condition $(\mathcal{E}_0)$. However, the equivalence between $(\mathcal{E}_1)$ and $(\mathcal{E}_0)$ is only a specific feature of second order operators (see, for example, [6, p. 15]). For more relationships on these two kinds of ellipticity conditions, we refer the reader to [4, p. 365].

Let us recall some basic facts on sesquilinear forms from [36, Section 1.2.1].

Definition 2.2 ([36]). Assume that $\mathfrak{a} : D(\mathfrak{a}) \times D(\mathfrak{a}) \to \mathbb{C}$ is a sesquilinear form in the Hilbert space $\mathcal{H}$.
(i) $\mathfrak{a}$ is said to be *densely defined* if the domain of $\mathfrak{a}$, $D(\mathfrak{a})$, is dense in $\mathcal{H}$.
(ii) $\mathfrak{a}$ is said to be *accretive* if, for all $u \in D(\mathfrak{a})$,

$$\mathrm{Re}(\mathfrak{a}(u,u)) \ge 0.$$

(iii) $\mathfrak{a}$ is said to be *continuous* if there exists a non-negative constant M such that, for all $u, v \in D(\mathfrak{a})$,

$$|\mathfrak{a}(u, v)| \le M\|u\|_{\mathfrak{a}}\|v\|_{\mathfrak{a}},$$

where $\|u\|_{\mathfrak{a}} := \sqrt{\mathrm{Re}(\mathfrak{a}(u, u)) + \|u\|_{\mathcal{H}}^2}$.

(iv) $\mathfrak{a}$ is said to be *closed* if $(D(\mathfrak{a}), \|\cdot\|_{\mathfrak{a}})$ is a complete space.

For a densely defined, accretive, continuous and closed sesquilinear form in the Hilbert space $\mathcal{H}$, we have the following conclusion from [36, Proposition 1.22]. Recall that $\|\cdot\|_{\mathcal{H}}$ and $(\cdot\,,\cdot)_{\mathcal{H}}$ denote, respectively, the inner product and the norm of $\mathcal{H}$.

Proposition 2.3 ([36]). *Assume that $\mathfrak{a}$ is a densely defined, accretive, continuous and closed sesquilinear form in the Hilbert space $\mathcal{H}$. Then, there exists a densely defined operator T, defined by setting*

$$D(T) := \{u \in \mathcal{H} : \text{there exists } v \in \mathcal{H} \text{ such that for all } \phi \in D(\mathfrak{a}),\ \mathfrak{a}(u, \phi) = (v, \phi)_{\mathcal{H}}\}$$

and $Tu := v$ for all $u \in D(T)$, such that for all $\lambda \in (0, \infty)$, $\lambda I + T$ is invertible (from $D(T)$ into $\mathcal{H}$) and $(\lambda I + T)^{-1}$ is bounded on $\mathcal{H}$. Moreover, for all $\lambda \in (0, \infty)$ and $f \in \mathcal{H}$,

$$\|\lambda(\lambda I + T)^{-1}(f)\|_{\mathcal{H}} \le \|f\|_{\mathcal{H}}.$$

For $\mathfrak{a}_0$ defined as in (2.1), from the fact that $\dot{W}^{m,2}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ and the ellipticity condition $(\mathcal{E}_0)$, we deduce that $\mathfrak{a}_0$ is a densely defined, accretive and continuous sesquilinear form.

Moreover, let $W^{m,2}(\mathbb{R}^n)$ be the m-order inhomogeneous Sobolev space equipped with the usual norm

$$\|f\|_{W^{m,2}(\mathbb{R}^n)} := \left[\sum_{0\le|\alpha|\le m} \|\partial^\alpha f\|_{L^2(\mathbb{R}^n)}^2\right]^{\frac{1}{2}}. \tag{2.2}$$

For all $f \in D(\mathfrak{a}_0)$, by the ellipticity condition $(\mathcal{E}_0)$ and Plancherel's theorem, it is easy to see that

$$\|f\|_{\mathfrak{a}_0} := \sqrt{\mathrm{Re}(\mathfrak{a}_0(f, f)) + \|f\|_{L^2(\mathbb{R}^n)}^2} \sim \|f\|_{W^{m,2}(\mathbb{R}^n)}.$$

This, combined with the fact that $W^{m,2}(\mathbb{R}^n)$ is a Banach space, further implies that

$$(\dot{W}^{m,2}(\mathbb{R}^n), \|\cdot\|_{\mathfrak{a}_0})$$

is complete. Thus, $\mathfrak{a}_0$ is closed. Using Proposition 2.3, we know that there exists a densely defined operator L in $L^2(\mathbb{R}^n)$ associated with $\mathfrak{a}_0$, which is formally written as in (1.1).

Let $\omega \in [0, \pi/2)$. Recall that an operator T in the Hilbert space $\mathcal{H}$ is said to be *m-ω-accretive* (or *maximal ω-accretive*) if the following hold:

(i) The range of the operator $T + I$, $R(T + I)$, is dense in $\mathcal{H}$.

(ii) For all $u \in D(T)$, $|\arg(T(u), u)_{\mathcal{H}}| \le \omega$, where $\arg(T(u), u)_{\mathcal{H}}$ denotes the argument of $(T(u), u)_{\mathcal{H}}$, see [27, p. 173].

It is known that, by [27, Proposition 7.1.1], every closed m-ω-accretive operator is of type ω in $L^2(\mathbb{R}^n)$, namely, the spectrum of T, $\sigma(T)$, is contained in the sector

$$S_\omega := \{z \in \mathbb{C} : |\arg z| \le \omega\},$$

and for each $\theta \in (\omega, \pi)$, there exists a non-negative constant C such that, for all $z \in \mathbb{C} \setminus S_\theta$,

$$\|(T - zI)^{-1}\|_{\mathcal{L}(L^2(\mathbb{R}^n))} \le C|z|^{-1},$$

where $\|S\|_{\mathcal{L}(\mathcal{H})}$ denotes the operator norm of the linear operator S on the normed linear space $\mathcal{H}$.

Moreover, by [27] we know that, if T is of type ω, then $-T$ generates a semigroup $\{e^{-tT}\}_{t>0}$, which can be extended to a bounded holomorphic semigroup $\{e^{-zT}\}_{z\in S^0_{\pi/2-\omega}}$ in the open sector

$$S^0_{\pi/2-\omega} := \{z \in \mathbb{C} \setminus \{0\} : |\arg z| < \pi/2 - \omega\}.$$

Recall that, by the ellipticity condition $(\mathcal{E}_0)$, we know that L is an m-arctan(Λ/λ)-accretive operator in $L^2(\mathbb{R}^n)$. Thus, $-L$ generates a bounded holomorphic semigroup in the open sector $S^0_{\pi/2-\arctan(\Lambda/\lambda)}$.

The following $L^2(\mathbb{R}^n)$ off-diagonal estimates of $\{e^{-zL}\}_{z\in S^0_{\pi/2-\arctan(\Lambda/\lambda)}}$ are well known (see, for example, [2, p. 66], [21, Theorem 3.2] or [15, Lemma 3.1]).

Proposition 2.4. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, *and let* $\omega := \arctan(\Lambda_0/\lambda_0)$, *where Λ_0 and λ_0 are as in the ellipticity condition* $(\mathcal{E}_0)$. *Then, for all* $\ell \in (0,1)$, $k \in \mathbb{Z}_+$, *the family of operators*, $\{(zL)^k e^{-zL}\}_{z \in S^0_{\ell(\pi/2-\omega)}}$, *satisfies the m-Davies–Gaffney estimates in z. That is, there exist positive constants C and* $\tilde{C}$ *such that, for all* $f \in L^2(\mathbb{R}^n)$ *supported in E and* $z \in S^0_{\ell(\pi/2-\omega)}$,

$$\|(zL)^k e^{-zL}(f)\|_{L^2(F)} \le C \exp\left\{-\tilde{C}\frac{[\operatorname{dist}(E,F)]^{2m/(2m-1)}}{|z|^{1/(2m-1)}}\right\}\|f\|_{L^2(E)}.$$

We now consider the $L^p(\mathbb{R}^n)$ theory of $\{e^{-tL}\}_{t>0}$. Let $(p_-(L), p_+(L))$ be the maximal interval of exponents $p \in [1,\infty]$ such that $\{e^{-tL}\}_{t>0}$ is bounded on $L^p(\mathbb{R}^n)$. Let $(q_-(L), q_+(L))$ be the maximal interval of exponents $q \in [1,\infty]$ such that $\{\sqrt{t}\nabla^m e^{-tL}\}_{t>0}$ is bounded on $L^q(\mathbb{R}^n)$. By [2, pp. 66–67] and [21, Theorem 3.2], we have the following proposition.

Proposition 2.5 ([2, 21]). *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Then, the following hold:*

(i) *We have that*

$$\begin{cases} (p_-(L), p_+(L)) = (1,\infty) & \text{if } n \le 2m,\\ \left[\dfrac{n}{n+2m}, \dfrac{n}{n-2m}\right] \subset (p_-(L), p_+(L)) & \text{if } n > 2m.\end{cases}$$

(ii) *We have that* $q_-(L) = p_-(L)$, $q_+(L) > 2$ *and* $p_+(L) \ge (q_+(L))^{*m}$, *where, for any* $q \in (1,\infty)$,

$$q^* := \begin{cases} \dfrac{np}{n-p} & \text{if } p < n,\\ \infty & \text{if } p \ge n\end{cases}$$

denotes the Sobolev exponent of q and q^{*m} means the m-th iteration of the operation $q \mapsto q^*$.

(iii) *For all* $k \in \mathbb{Z}_+$ *and* $p_-(L) < p \le q < p_+(L)$, *the family* $\{(tL)^k e^{-tL}\}_{t>0}$ *of operators satisfies the following m-L^p-L^q off-diagonal estimates: There exist positive constants C and $\tilde{C}$ such that, for any closed sets E, F in* $\mathbb{R}^n$, $t \in (0,\infty)$ *and* $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ *supported in E*,

$$\|(tL)^k e^{-tL}(f)\|_{L^q(F)} \le C t^{\frac{n}{2m}(\frac{1}{q}-\frac{1}{p})} \exp\left\{-\tilde{C}\frac{[d(E,F)]^{\frac{2m}{2m-1}}}{t^{\frac{1}{2m-1}}}\right\}\|f\|_{L^p(\mathbb{R}^n)}.$$

(iv) *For all* $p_-(L) < p \le q < q_+(L)$, *the family* $\{(t^{1/(2m)}\nabla)^k e^{-tL}\}_{t>0}$ *of operators satisfies the following m-L^p-L^q off-diagonal estimates: There exist positive constants C and $\tilde{C}$ such that, for any closed sets E, F in* $\mathbb{R}^n$, $t \in (0,\infty)$ *and* $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ *supported in E*,

$$\|(t^{1/(2m)}\nabla)^k e^{-tL}(f)\|_{L^q(F)} \le C t^{\frac{n}{2m}(\frac{1}{q}-\frac{1}{p})} \exp\left\{-\tilde{C}\frac{[d(E,F)]^{\frac{2m}{2m-1}}}{t^{\frac{1}{2m-1}}}\right\}\|f\|_{L^p(\mathbb{R}^n)}.$$

Finally, we recall some results on the square root of L. Let L be defined as in (1.1). It is known that L is one-to-one and m-ω-accretive. By [6, p. 8], we know that L has a bounded holomorphic functional calculus in $L^2(\mathbb{R}^n)$. Thus, its square root $L^{1/2}$ is well defined on $L^2(\mathbb{R}^n)$.

Auscher, Hofmann, McIntosh and Tchamitchian proved the following result on Kato's square root problem of $L^{1/2}$ (see [4, Theorem 1.1]).

Proposition 2.6 ([4]). *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *The square root of L has a domain equal to the Sobolev space* $W^{m,2}(\mathbb{R}^n)$ *defined as in* (2.2). *Moreover, there exists a positive constant C such that, for all* $f \in W^{m,2}(\mathbb{R}^n)$,

$$\frac{1}{C}\|\sqrt{L}(f)\|_{L^2(\mathbb{R}^n)} \le \|\nabla^m f\|_{L^2(\mathbb{R}^n)} \le C\|\sqrt{L}(f)\|_{L^2(\mathbb{R}^n)}.$$

Proposition 2.6 implies immediately that the Riesz transform $\nabla^m L^{-1/2}$ associated with L is bounded on $L^2(\mathbb{R}^n)$. Moreover, Auscher proved the following boundedness of $\nabla^m L^{-1/2}$ on $L^p(\mathbb{R}^n)$ (see [2, p. 68]).

Proposition 2.7 ([2]). *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Then, for all p in* $(q_-(L), q_+(L))$, $\nabla^m L^{-1/2}$ *is bounded on* $L^p(\mathbb{R}^n)$.

We also refer the reader to [9, Theorem 1.2] for a related result on the boundedness of $\nabla^m L^{-1/2}$.

2.2 The Hardy space $H_L^p(\mathbb{R}^n)$

Let L be the homogeneous higher order operator defined as in (1.1) that satisfies the ellipticity condition $(\mathcal{E}_0)$. Let $H_L^p(\mathbb{R}^n)$ be the Hardy space associated with L defined as in Definition 1.1. In this subsection, we give some real-variable properties of $H_L^p(\mathbb{R}^n)$ for $p \in (0, \infty)$. Our first result is the following complex interpolation of $H_L^p(\mathbb{R}^n)$. Recall ([20]) that, for all $p \in (0, \infty)$, a function F on $\mathbb{R}_+^{n+1}$ is said to be in the *tent space* $T^p(\mathbb{R}_+^{n+1})$, if $\|F\|_{T^p(\mathbb{R}_+^{n+1})} =: \|\mathcal{A}(F)\|_{L^p(\mathbb{R}^n)} < \infty$, where

$$\mathcal{A}(F)(x) := \left\{ \iint_{\Gamma(x)} |F(y,t)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}}, \tag{2.3}$$

with $\Gamma(x)$ for all $x \in \mathbb{R}^n$ as in (1.3), denotes the $\mathcal{A}$-functional of F (see [20] for more properties of tent spaces).

Proposition 2.8. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Then, for each $\theta \in (0,1)$ and $0 < p_1 < p_2 < \infty$,*

$$[H_L^{p_1}(\mathbb{R}^n), H_L^{p_2}(\mathbb{R}^n)]_\theta = H_L^p(\mathbb{R}^n),$$

where p satisfies $\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$ and $[\,\cdot\,,\cdot\,]_\theta$ denotes the complex interpolation (see, for example, [33, Section 7]*).*

Proof. The proof of Proposition 2.8 is a consequence of the complex interpolation of tent spaces $T^p(\mathbb{R}_+^{n+1})$ and the fact that $H_L^p(\mathbb{R}^n)$ is a retract of $T^p(\mathbb{R}_+^{n+1})$ (see [29, Lemma 4.20] for more details in the case when $m = 1$), the details being omitted. □

For $H_L^p(\mathbb{R}^n)$ with $p \in (0,1]$, one of its most useful properties is its molecular characterization. To state it, we first recall the following notion of $(p, 2, M, q)_L$-molecules.

Definition 2.9. Let L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$, $p \in (0,1]$, $\epsilon \in (0, \infty)$ and $M \in \mathbb{N}$. A function $\alpha \in L^2(\mathbb{R}^n)$ is called a $(p, 2, M, \epsilon)_L$*-molecule* if there exists a ball $B \subset \mathbb{R}^n$ such that, for each $\ell \in \{0, \dots, M\}$, α belongs to the range of L^ℓ in $L^2(\mathbb{R}^n)$ and, for all $i \in \mathbb{Z}_+$ and $\ell \in \{0, \dots, M\}$,

$$\|(r_B^{2m} L)^{-\ell}(\alpha)\|_{L^2(S_i(B))} \le 2^{-i\epsilon} |2^i B|^{\frac{1}{2}-\frac{1}{p}}.$$

Assume that $\{\alpha_j\}_j$ is a sequence of $(p, 2, M, \epsilon)_L$-molecules and $\{\lambda_j\}_j \in l^p$. For any $f \in L^2(\mathbb{R}^n)$, if $f = \sum_j \lambda_j \alpha_j$ in $L^2(\mathbb{R}^n)$, then $\sum_j \lambda_j \alpha_j$ is called a *molecular* $(p, 2, M, \epsilon)_L$*-representation* of f.

Definition 2.10. Let L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$, and let $p \in (0,1]$, $\epsilon \in (0,\infty)$ and $M \in \mathbb{N}$. The *molecular Hardy space* $H^p_{L,\mathrm{mol},M,\epsilon}(\mathbb{R}^n)$ is defined as the completion of the space

$$\mathbb{H}^p_{L,\mathrm{mol},M,\epsilon}(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n) : f \text{ has a molecular } (p,2,M,\epsilon)_L\text{-representation}\}$$

with respect to the quasi-norm

$$\|f\|_{H^p_{L,\mathrm{mol},M,\epsilon}(\mathbb{R}^n)} := \inf\left\{ \left(\sum_j |\lambda_j|^p \right)^{\frac{1}{p}} : f = \sum_j \lambda_j \alpha_j \text{ is a molecular } (p,2,M,\epsilon)_L\text{-representation} \right\},$$

where the infimum is taken over all the molecular $(p, 2, M, \epsilon)_L$-representations of f as above.

Theorem 2.11 ([15, 21]). *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, *and let $p \in (0,1]$, $\epsilon \in (0,\infty)$ and $M \in \mathbb{N}$ such that $M > \frac{n}{2m}(\frac{1}{p} - \frac{1}{2})$. Then, $H_L^p(\mathbb{R}^n) = H^p_{L,\mathrm{mol},M,\epsilon}(\mathbb{R}^n)$ with equivalent quasi-norms.*

For more characterizations of $H_L^p(\mathbb{R}^n)$ with $p \in (0,1]$, we refer the reader to [15, 21].

We now study the relationship between $H_L^p(\mathbb{R}^n)$ and the Lebesgue space $L^p(\mathbb{R}^n)$.

Lemma 2.12. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, *and let $p_-(L)$ and $p_+(L)$ be as in Proposition 2.5. Then, for all $p \in (p_-(L), p_+(L))$, $H_L^p(\mathbb{R}^n) = L^p(\mathbb{R}^n)$ with equivalent norms.*

Proof. We prove Lemma 2.12 by borrowing some ideas from the proof of [29, Proposition 9.1 (v)]. First, from [12, Propositions 2.10 and 2.13], it follows that, for all $p \in (p_-(L), p_+(L))$, S_L is bounded on $L^p(\mathbb{R}^n)$. This, together with Definition 1.1, shows that, for all $p \in (p_-(L), 2]$ and $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$,

$$\|f\|_{H_L^p(\mathbb{R}^n)} := \|S_L(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)},$$

which immediately implies that, for all $p \in (p_-(L), 2]$,

$$[L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)] \subset [L^2(\mathbb{R}^n) \cap H_L^p(\mathbb{R}^n)].$$

On the other hand, recall the following Calderón reproducing formula for L (since L has a bounded holomorphic functional calculus in $L^2(\mathbb{R}^n)$): for all $g \in L^2(\mathbb{R}^n)$,

$$g = \tilde{C} \int_0^\infty (t^{2m}L)^{M+2} e^{-2t^{2m}L}(g)\, \frac{dt}{t} =: \tilde{C}\pi_{L,M} \circ Q_{L,1,t}(g) \tag{2.4}$$

holds true in $L^2(\mathbb{R}^n)$, where $\tilde{C}$ is a positive constant such that

$$\tilde{C} \int_0^\infty t^{2m(M+2)} e^{-2t^{2m}}\, \frac{dt}{t} = 1,$$

$M \in \mathbb{N}$ is sufficiently large,

$$\pi_{L,M} := \int_0^\infty (t^{2m}L)^{M+1} e^{-t^{2m}L}\, \frac{dt}{t}$$

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

$$Q_{L,k,t} := (t^{2m}L)^k e^{-t^{2m}L}.$$

Thus, for $p \in (p_-(L), 2]$, if $f \in L^2(\mathbb{R}^n) \cap H_L^p(\mathbb{R}^n)$, then, for all $g \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n)$, by (2.4), the duality between $T^p(\mathbb{R}_+^{n+1})$ and $T^{p'}(\mathbb{R}_+^{n+1})$ with $1/p + 1/p' = 1$ and Hölder's inequality, we see that

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} f(x)\overline{g(x)}\, dx \right| &= \tilde{C} \left| \int_{\mathbb{R}^n} \pi_{L,M} \circ Q_{L,1,t}(f)(x)\overline{g(x)}\, dx \right| \\
&= \tilde{C} \left| \iint_{\mathbb{R}_+^{n+1}} Q_{L,1,t}(f)(x)\overline{Q_{L^*,M+1,t}(g)(x)}\, \frac{dx\,dt}{t} \right| \\
&\lesssim \|Q_{L,1,t}(f)\|_{T^p(\mathbb{R}_+^{n+1})} \|Q_{L^*,M+1,t}(g)\|_{T^{p'}(\mathbb{R}_+^{n+1})} \\
&\sim \|f\|_{H_L^p(\mathbb{R}^n)} \left\| \left\{ \iint_{\Gamma(\cdot)} |(t^{2m}L^*)^{M+1} e^{-t^{2m}L^*}(g)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}} \right\|_{L^{p'}(\mathbb{R}^n)},
\end{aligned}$$

where $T^p(\mathbb{R}_+^{n+1})$ denotes the tent space and L^* the adjoint operator of L in $L^2(\mathbb{R}^n)$. Since $p' \in [2, p_+(L^*))$, similar to the boundedness of S_L on $L^{p'}(\mathbb{R}^n)$, for all $p' \in (p_-(L), p_+(L))$, we have

$$\|Q_{L^*,M+1,t}(g)\|_{T^{p'}(\mathbb{R}_+^{n+1})} \lesssim \|g\|_{L^{p'}(\mathbb{R}^n)},$$

which, combined with the arbitrariness of g, implies that $f \in L^p(\mathbb{R}^n)$ and

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{H_L^p(\mathbb{R}^n)},$$

and hence $[L^2(\mathbb{R}^n) \cap H_L^p(\mathbb{R}^n)] \subset [L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)]$. By the density, this finishes the proof of Lemma 2.12 for $p \in (p_-(L), 2]$. The case $p \in [2, p_+(L))$ follows from Definition 1.1 and a dual argument, the details being omitted. This finishes the proof of Lemma 2.12. □

Combining Lemma 2.12, Propositions 2.7 and 2.8, together with the fact that $\nabla^m L^{-1/2}$ is bounded from $H_L^p(\mathbb{R}^n)$ to the classical Hardy space $H^p(\mathbb{R}^n)$ for all $p \in (\frac{n}{n+m}, 1]$ (see [15, Theorem 6.2]), we conclude the following proposition, the details being omitted.

Proposition 2.13. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Then, for all* $p \in (\frac{n}{n+m}, q_+(L))$, $\nabla^m L^{-1/2}$ *is bounded from* $H_L^p(\mathbb{R}^n)$ *to* $H^p(\mathbb{R}^n)$.

Now, we establish the generalized square function characterization of $H_L^p(\mathbb{R}^n)$, which is available in [29] for $m = 1$ and $p \in (0, \infty)$, and in [15] for $m \in \mathbb{N}$ and $p \in (0, 1]$. Let $p \in (0, \infty)$, $\omega \in [0, \pi/2)$ be the type of L, $\alpha \in (0, \infty)$, $\beta \in (\frac{n}{2m}(\max\{\frac{1}{p}, 1\} - \frac{1}{2}), \infty)$ and $\psi \in \Psi_{\alpha,\beta}(S_\mu^0)$ with $\mu \in (\omega, \pi/2)$, where

$$S_\mu^0 := \{z \in \mathbb{C} \setminus \{0\} : |\arg z| < \mu\}$$

and

$$\Psi_{\alpha,\beta}(S_\mu^0) := \{f \text{ is analytic on } S_\mu^0 : \text{there is a constant } C > 0 \text{ such that } |f(\xi)| \le C \min\{|\xi|^\alpha, |\xi|^{-\beta}\} \text{ for all } \xi \in S_\mu^0\}.$$

For all $f \in L^2(\mathbb{R}^n)$ and $(x, t) \in \mathbb{R}_+^{n+1}$, define the operator $Q_{\psi,L}(f)$ by

$$Q_{\psi,L}(f)(x, t) := \psi(t^{2m}L)(f)(x).$$

Definition 2.14. Let $p \in (0, \infty)$, L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$, $\alpha \in (0, \infty)$, $\beta \in (\frac{n}{2m}(\max\{\frac{1}{p}, 1\} - \frac{1}{2}), \infty)$, $\mu \in (\omega, \pi/2)$ and

$$\psi \in \begin{cases} \Psi_{\alpha,\beta}(S_\mu^0) & \text{if } p \in (0, 2], \\ \Psi_{\beta,\alpha}(S_\mu^0) & \text{if } p \in (2, \infty). \end{cases}$$

The *generalized square function Hardy space* $H_{\psi,L}^p(\mathbb{R}^n)$ is defined as the completion of the space

$$\mathbb{H}_{\psi,L}^p(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n) : Q_{\psi,L}(f) \in T^p(\mathbb{R}_+^{n+1})\}$$

with respect to the quasi-norm $\|f\|_{H_{\psi,L}^p(\mathbb{R}^n)} := \|Q_{\psi,L}(f)\|_{T^p(\mathbb{R}_+^{n+1})}$.

The following result establishes the generalized square function characterization of $H_L^p(\mathbb{R}^n)$ for $p \in (0, \infty)$.

Proposition 2.15. *Let $p \in (0, \infty)$, L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, $\alpha \in (0, \infty)$, $\beta \in (\frac{n}{2m}(\max\{\frac{1}{p}, 1\} - \frac{1}{2}), \infty)$, $\mu \in (\omega, \pi/2)$ *and*

$$\psi \in \begin{cases} \Psi_{\alpha,\beta}(S_\mu^0) & \text{if } p \in (0, 2], \\ \Psi_{\beta,\alpha}(S_\mu^0) & \text{if } p \in (2, \infty). \end{cases}$$

Then, the Hardy space $H_L^p(\mathbb{R}^n) = H_{\psi,L}^p(\mathbb{R}^n)$ with equivalent quasi-norms.

Proof. If $p \in (0, 1]$, Proposition 2.15 is just [15, Theorem 5.2], where $\beta \in (\frac{n}{2m}(\frac{1}{p} - \frac{1}{2}), \infty)$ is needed to guarantee $H_L^p(\mathbb{R}^n) \subset H_{\psi,L}^p(\mathbb{R}^n)$, via an application of the Calderón reproducing formula.

If $p \in (1, \infty)$ and $m = 1$, Proposition 2.15 is just [29, Corollary 4.17], where $\beta \in (\frac{n}{4}, \infty)$ is used to guarantee $H_L^p(\mathbb{R}^n) \subset H_{\psi,L}^p(\mathbb{R}^n)$, via an application of the Calderón reproducing formula. If $p \in (1, \infty)$ and $m \in \mathbb{N} \cap [2, \infty)$, an argument similar to that used in the proof of [29, Corollary 4.17], together with an application of the Calderón reproducing formula, also gives us the desired conclusion of Proposition 2.15, where we need $\beta \in (\frac{n}{4m}, \infty)$ to guarantee $H_L^p(\mathbb{R}^n) \subset H_{\psi,L}^p(\mathbb{R}^n)$, the details being omitted, which completes the proof of Proposition 2.15. □

3 Proofs of Theorems 1.4, 1.5 and 1.8, and Proposition 1.6

In this section, we give the proofs of Theorems 1.4, 1.5 and 1.8, and Proposition 1.6. To this end, we first establish the following parabolic Caccioppoli's inequalities, resonating with [28, Lemma 2.8] and, in a different way, with [5, Proposition 40].

Proposition 3.1. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, *and let* $f \in L^2(\mathbb{R}^n)$, $t \in (0, \infty)$ *and* $u(x, t) := e^{-t^{2m}L}(f)(x)$ *for all* $x \in \mathbb{R}^n$. *For all* $\epsilon \in (0, \infty)$, *there exist positive constants* $C_{(\epsilon)}$, *depending on* ϵ, *and C such that, for all $x_0 \in \mathbb{R}^n$, $r \in (0,\infty)$ and $t_0 \in (3r,\infty)$,*

$$\int_{t_0-r}^{t_0+r}\int_{B(x_0,r)} |\nabla^m u(x,t)|^2\,dx\,dt \le \epsilon \int_{t_0-2r}^{t_0+2r}\int_{B(x_0,2r)} |\nabla^m u(x,t)|^2\,dx\,dt + \frac{C_{(\epsilon)}}{r^{2m}}\int_{t_0-2r}^{t_0+2r}\int_{B(x_0,2r)} |u(x,t)|^2\,dx\,dt \tag{3.1}$$

and

$$\int_{t_0-r}^{t_0+r}\int_{B(x_0,r)} |\nabla^m u(x,t)|^2\,dx\,dt \le \sum_{j=0}^{m-1}\frac{C}{r^{2(m-j)}}\int_{t_0-2r}^{t_0+2r}\int_{B(x_0,2r)} |\nabla^j u(x,t)|^2\,dx\,dt, \tag{3.2}$$

where $C_{(\epsilon)}$ and C are independent of f.

Proof. We first prove (3.2). To this end, we introduce two smooth cut-off functions. Let $\eta \in C_c^\infty(B(x_0,2r))$ satisfy $0 \le \eta \le 1$, $\eta \equiv 1$ on $B(x_0,r)$ and, for all $k \in \{0,\dots,m\}$,

$$\|\nabla^k \eta\|_{L^\infty(\mathbb{R}^n)} \lesssim r^{-k}.$$

Let $\gamma \in C_c^\infty(t_0-2r, t_0+2r)$ satisfy $0 \le \gamma \le 1$, $\gamma \equiv 1$ on (t_0-r, t_0+r) and

$$\|\partial_t \gamma\|_{L^\infty(\mathbb{R})} \lesssim \frac{1}{r}.$$

By the properties of η and γ, and the ellipticity condition $(\mathcal{E}_0)$, we first write

$$\begin{aligned}\int_{t_0-r}^{t_0+r}\int_{B(x_0,r)} |\nabla^m u(x,t)|^2\,dx\,dt &\le \int_{t_0-2r}^{t_0+2r}\int_{\mathbb{R}^n} |\nabla^m(u\eta^m)(x,t)|^2\,dx\,\gamma(t)\,dt\\ &\le \frac{1}{\lambda_0}\operatorname{Re}\left\{\int_{t_0-2r}^{t_0+2r}\int_{\mathbb{R}^n} A(x)\nabla^m(u\eta^m)(x,t)\overline{\nabla^m(u\eta^m)}(x,t)\,dx\,\gamma(t)\,dt\right\}\\ &=: \frac{1}{\lambda_0}\mathcal{A},\end{aligned} \tag{3.3}$$

where

$$A(x) := \{a_{\alpha,\beta}(x)\}_{|\alpha|=m=|\beta|} \quad \text{for all } x \in \mathbb{R}^n \tag{3.4}$$

is a (properly arranged) coefficient matrix of L so that, for all $f, g \in \dot{W}^{m,2}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$A(x)\nabla^m f(x)\overline{\nabla^m g(x)} := \sum_{|\alpha|=m=|\beta|} a_{\alpha,\beta}(x)\partial^\beta f(x)\overline{\partial^\alpha g(x)}.$$

To bound $\mathcal{A}$, let

$$\mathcal{B} := \operatorname{Re}\left\{\int_{t_0-2r}^{t_0+2r}\int_{\mathbb{R}^n} A(x)\nabla^m(u)(x,t)\overline{\nabla^m(u\eta^{2m})}(x,t)\,dx\,\gamma(t)\,dt\right\}.$$

We first bound $\mathcal{B}$. For all $(x,\tilde t) \in \mathbb{R}^{n+1}_+$, let

$$F(x,\tilde t) := e^{-\tilde tL}(f)(x)\overline{e^{-\tilde tL}(f)(x)}[\eta(x)]^{2m}.$$

Using $\partial_{\tilde t} e^{-\tilde tL} = -Le^{-\tilde tL}$ and

$$\overline{Le^{-\tilde tL}(f)(x)\overline{e^{-\tilde tL}(f)(x)}[\eta(x)]^{2m}} = e^{-\tilde tL}(f)(x)\overline{Le^{-\tilde tL}(f)(x)}[\eta(x)]^{2m},$$

we know that

$$\begin{aligned}\partial_{\tilde t}F(x,\tilde t) &= -Le^{-\tilde tL}(f)(x)\overline{e^{-\tilde tL}(f)(x)}[\eta(x)]^{2m} - e^{-\tilde tL}(f)(x)\overline{Le^{-\tilde tL}(f)(x)}[\eta(x)]^{2m}\\ &= -2\operatorname{Re}\left\{Le^{-\tilde tL}(f)(x)\overline{e^{-\tilde tL}(f)(x)}[\eta(x)]^{2m}\right\},\end{aligned}$$

which, together with integration by parts and the definition of the cut-off function γ, shows that

$$
\begin{aligned}
\mathcal{B} &= \frac{1}{\lambda_1}\operatorname{Re}\Bigg[\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} Le^{-t^{2m}L}(f)(x)\overline{e^{-t^{2m}L}(f)(x)[\eta(x)]^{2m}}\,dx\Bigg\}\gamma(t)\,dt\Bigg]\\
&= \frac{1}{2m\lambda_1}\operatorname{Re}\Bigg\{\int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\Bigg[\int_{\mathbb{R}^n} Le^{-\tilde{t}L}(f)(x)\overline{e^{-\tilde{t}L}(f)(x)[\eta(x)]^{2m}}\,dx\Bigg]\gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-1}\,d\tilde{t}\Bigg\}\\
&= -\frac{1}{4m\lambda_1}\int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\partial_{\tilde{t}}\Bigg(\int_{B(x_0,2r)} F(x,\tilde{t})\,dx\Bigg)\gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-1}\,d\tilde{t}\\
&= \frac{1}{4m\lambda_1}\int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\Bigg[\int_{B(x_0,2r)} F(x,\tilde{t})\,dx\Bigg]\partial_{\tilde{t}}(\gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-1})\,d\tilde{t}\\
&\qquad - \frac{1}{4m\lambda_1}\int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\partial_{\tilde{t}}\Bigg(\Bigg[\int_{B(x_0,2r)} F(x,\tilde{t})\,dx\Bigg]\gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-1}\Bigg)\,d\tilde{t}\\
&= \frac{1}{4m\lambda_1}\int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\Bigg[\int_{B(x_0,2r)} F(x,\tilde{t})\,dx\Bigg]\partial_{\tilde{t}}(\gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-1})\,d\tilde{t}.
\end{aligned}
\tag{3.5}
$$

This, combined with the change of variables, the size condition of γ and $t_0 \in (3r,\infty)$, implies that

$$
\begin{aligned}
|\mathcal{B}| &\lesssim \int_{(t_0-2r)^{2m}}^{(t_0+2r)^{2m}}\Bigg\{\int_{B(x_0,2r)} e^{-\tilde{t}L}(f)(x)\overline{e^{-\tilde{t}L}(f)(x)}[\eta(x)]^{2m}\,dx\Bigg\}\Big[|\partial_{\tilde{t}}\gamma(\tilde{t}^{\frac{1}{2m}})|\tilde{t}^{2(\frac{1}{2m}-1)} + \gamma(\tilde{t}^{\frac{1}{2m}})\tilde{t}^{\frac{1}{2m}-2}\Big]\,d\tilde{t}\\
&\lesssim \frac{1}{r^{2m}}\int_{t_0-2r}^{t_0+2r}\int_{B(x_0,2r)} |u(x,t)|^2\,dx\,dt,
\end{aligned}
\tag{3.6}
$$

which is desired.

On the other hand, by the definition of $\mathcal{B}$ and Leibniz's rule, we know that

$$
\begin{aligned}
\mathcal{B} &= \operatorname{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta u(x,t)\overline{\partial^\alpha(u\eta^{2m})}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg)\\
&= \operatorname{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta u(x,t)\eta^m\overline{\partial^\alpha(u\eta^{m})}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg)\\
&\qquad + \operatorname{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta u(x,t)\overline{\Bigg[\sum_{\theta\le\xi<\alpha} C_{(\alpha,\xi)}\partial^\xi(u\eta^m)\partial^{\alpha-\xi}(\eta^m)\Bigg]}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg),
\end{aligned}
\tag{3.7}
$$

where $\theta := (0,\dots,0) \in (\mathbb{Z}_+)^n$ and, for each α and ξ, $C_{(\alpha,\xi)}$ is a positive constant depending on α and ξ, and $\xi < \alpha$ means that each component of ξ is not larger than the corresponding component of α and $|\xi| < |\alpha|$.

Before going further, we make the following observation. For all multi-indices γ with $0 \le |\gamma| \le m$, by Leibniz's rule, we see that there exists a smooth function η_γ on $\mathbb{R}^n$ such that $\partial^\gamma(\eta^m) = \eta^{m-|\gamma|}\eta_\gamma$ and

$$\|\eta_\gamma\|_{L^\infty(\mathbb{R}^n)} \lesssim \frac{1}{r^{|\gamma|}}. \tag{3.8}$$

Indeed, if $\gamma \equiv (2,0,\dots,0)$, then we have

$$
\begin{aligned}
\partial^\gamma(\eta^m) &= \partial_{x_1}^2(\eta^m) = \partial_{x_1}(m\eta^{m-1}\partial_{x_1}\eta) = m(m-1)\eta^{m-2}(\partial_{x_1}\eta)^2 + m\eta^{m-1}\partial_{x_1}^2\eta\\
&= \eta^{m-2}[m(m-1)(\partial_{x_1}\eta)^2 + m\eta\partial_{x_1}^2\eta] =: \eta^{m-2}\eta_\gamma,
\end{aligned}
$$

where the fact that η_γ satisfies (3.8) is an easy consequence of the properties of η. The general cases follow from a similar calculation, the details being omitted. From this fact and Leibniz's rule again, we deduce that, for each α and β as in (3.7),

$$\begin{aligned} C_{(\alpha,\xi)}\partial^\xi(u\eta^m)\partial^{\alpha-\xi}(\eta^m) &= C_{(\alpha,\xi)}\sum_{\theta\le\zeta\le\xi} C_{(\xi,\zeta)}\partial^{\xi-\zeta}(\eta^m)\partial^\zeta u\partial^{\alpha-\xi}(\eta^m)\\ &= C_{(\alpha,\xi)}\sum_{\theta\le\zeta\le\xi} C_{(\xi,\zeta)}\eta^{m-|\xi-\zeta|}\eta_{\xi-\zeta}\partial^\zeta u\eta^{m-|\alpha-\xi|}\eta_{\alpha-\xi}\\ &= \eta^m C_{(\alpha,\xi)}\sum_{\theta\le\zeta\le\xi} C_{(\xi,\zeta)}\eta^{m-|\xi-\zeta|-|\alpha-\xi|}\eta_{\xi-\zeta}\eta_{\alpha-\xi}\partial^\zeta u\\ &=: \eta^m\sum_{\theta\le\zeta\le\xi}\eta_{\alpha,\zeta}\partial^\zeta u, \end{aligned} \tag{3.9}$$

where $\xi\le\alpha$ means that each component of ξ is not larger than the corresponding component of α and, by (3.8), we see that

$$\|\eta_{\alpha,\zeta}\|_{L^\infty(\mathbb{R}^n)} \lesssim \frac{1}{r^{|\alpha-\zeta|}}. \tag{3.10}$$

Combining (3.7) and (3.9), we conclude that

$$\begin{aligned} \mathcal{B} &= \mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta u(x,t)\eta^m\overline{\partial^\alpha(u\eta^m)}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg)\\ &\quad+\mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta u(x,t)\overline{\Bigg[\sum_{\theta\le\xi<\alpha}\eta^m\sum_{\theta\le\zeta\le\xi}\eta_{\alpha,\zeta}\partial^\zeta u\Bigg]}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg), \end{aligned}$$

which further implies that

$$\begin{aligned} \mathcal{B} &= \mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta(u\eta^m)(x,t)\overline{\partial^\alpha(u\eta^m)}(x,t)\,dx\Bigg]\gamma(t)\,dt\Bigg)\\ &\quad-\mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\sum_{\theta\le\xi<\beta}C_{(\beta,\xi)}\partial^\xi u(x,t)\partial^{\beta-\xi}(\eta^m)\overline{\partial^\alpha(u\eta^m)}(x,t)\,dx\Bigg]\gamma(t)\,dt\Bigg)\\ &\quad+\mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\partial^\beta(u\eta^m)(x,t)\overline{\Bigg[\sum_{\theta\le\xi<\alpha}\sum_{\theta\le\zeta\le\xi}\eta_{\alpha,\zeta}\partial^\zeta u\Bigg]}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg)\\ &\quad-\mathrm{Re}\Bigg(\sum_{|\alpha|=m=|\beta|}\int_{t_0-2r}^{t_0+2r}\Bigg\{\int_{\mathbb{R}^n} a_{\alpha,\beta}(x)\sum_{\theta\le\xi<\beta}C_{(\beta,\xi)}\partial^\xi u(x,t)\partial^{\beta-\xi}(\eta^m)\overline{\Bigg[\sum_{\theta\le\xi<\alpha}\sum_{\theta\le\zeta\le\xi}\eta_{\alpha,\zeta}\partial^\zeta u\Bigg]}(x,t)\,dx\Bigg\}\gamma(t)\,dt\Bigg). \end{aligned}$$

This, together with the definition of $\mathcal{A}$, Hölder's inequality, (3.8), (3.10), (3.6), Hölder's inequality with ϵ and the ellipticity condition $(\mathcal{E}_0)$, implies that, for all $\epsilon\in(0,\infty)$, there exists a positive constant $C_{(\epsilon)}$ such that

$$\begin{aligned} \mathcal{A} &\lesssim \mathcal{B}+\Bigg\{\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n}|\nabla^m(u\eta^m)(x,t)|^2\,dx\Bigg]\gamma(t)\,dt\Bigg\}^{\frac12}\Bigg\{\sum_{j=0}^{m-1}\frac{1}{r^{2(m-j)}}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n}|\nabla^j u(x,t)|^2\,dx\Bigg]\gamma(t)\,dt\Bigg\}^{\frac12}\\ &\quad+\sum_{j=0}^{m-1}\frac{1}{r^{2(m-j)}}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n}|\nabla^j u(x,t)|^2\,dx\Bigg]\gamma(t)\,dt\\ &\lesssim \mathcal{B}+\epsilon\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{\mathbb{R}^n}|\nabla^m(u\eta^m)(x,t)|^2\,dx\Bigg]\gamma(t)\,dt+C_{(\epsilon)}\sum_{j=0}^{m-1}\frac{1}{r^{2(m-j)}}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{B(x_0,2r)}|\nabla^j u(x,t)|^2\,dx\Bigg]\gamma(t)\,dt\\ &\lesssim \sum_{j=0}^{m-1}\frac{1}{r^{2(m-j)}}\int_{t_0-2r}^{t_0+2r}\Bigg[\int_{B(x_0,2r)}|\nabla^j u(x,t)|^2\,dx\Bigg]\gamma(t)\,dt+\frac{\epsilon}{\lambda_0}\mathcal{A}, \end{aligned}$$

which, combined with (3.3), shows that (3.2) holds true.

We now turn to the proof of (3.1). Note that, from [1, Theorem 5.2 (3)] (with some slight modifications), we easily deduce that there exists a positive constant $C_{(n,m)}$, depending only on n and m, such that, for all balls B, $f \in W^{m,p}(B)$ and $k \in \{0, \dots, m\}$,

$$\|\nabla^k f\|_{L^2(B)} \le C_{(n,m)} \|\nabla^m f\|_{L^2(B)}^{k/m} \|f\|_{L^2(B)}^{1-k/m}, \tag{3.11}$$

which, together with (3.2), the interpolation inequality and Young's inequality with ϵ, immediately implies that (3.1) holds true. This finishes the proof of Proposition 3.1. □

The following proposition improves Proposition 3.1 by removing all the terms with gradients on the right-hand side of Caccioppoli's inequality (3.2), which is motivated by a recent result of Barton [7].

Proposition 3.2. *Let L be as in (1.1) and satisfying the ellipticity condition $(\mathcal{E}_0)$, and let $f \in L^2(\mathbb{R}^n)$, $t \in (0,\infty)$ and $u(x,t) := e^{-t^{2m}L}(f)(x)$ for all $x \in \mathbb{R}^n$. Then, there exists a positive constant C such that, for all $x_0 \in \mathbb{R}^n$, $r \in (0,\infty)$ and $t_0 \in (3r,\infty)$,*

$$\int_{t_0-r}^{t_0+r} \int_{B(x_0,r)} |\nabla^m u(x,t)|^2 \,dx\,dt \le \frac{C}{r^{2m}} \int_{t_0-2r}^{t_0+2r} \int_{B(x_0,2r)} |u(x,t)|^2 \,dx\,dt. \tag{3.12}$$

Proof. We prove this proposition by borrowing some ideas from the proof of [7, Theorem 3.10]. We first make the following claim that, to finish the proof of Proposition 3.2, we only need to show that, for all $j \in \{1, \dots, m\}$ and $0 < \zeta < \xi \le 2r$,

$$\int_{t_0-\zeta}^{t_0+\zeta} \int_{B(x_0,\zeta)} |\nabla^j u(x,t)|^2 \,dx\,dt \lesssim \sum_{k=0}^{j-1} \frac{1}{(\xi-\zeta)^{2(j-k)}} \int_{t_0-\xi}^{t_0+\xi} \int_{B(x_0,\xi)} |\nabla^k u(x,t)|^2 \,dx\,dt. \tag{3.13}$$

Indeed, if (3.13) holds true for all $j \in \{1, \dots, m\}$ and $0 < \zeta < \xi < \infty$, then let

$$r = r_0 < r_1 < r_2 < \cdots < r_m = 2r$$

be an average decomposition of $(r, 2r)$ and

$$A_{s,l} := \int_{t_0-s}^{t_0+s} \int_{B(x_0,s)} |\nabla^l u(x,t)|^2 \,dx\,dt \tag{3.14}$$

with $l \in \{0, \dots, m\}$ and $s \in (0,\infty)$. By repetitively using (3.13) with $j \in \{m, \dots, 1\}$, we have that

$$\begin{aligned}
\int_{t_0-r}^{t_0+r} \int_{B(x_0,r)} |\nabla^m u(x,t)|^2 \,dx\,dt = A_{r_0,m} &\lesssim \sum_{j=0}^{m-1} \frac{1}{r^{2(m-j)}} A_{r_1,j} \\
&\lesssim \frac{1}{r^{2m}} A_{r_1,0} + \sum_{j=1}^{m-1} \frac{1}{r^{2(m-j)}} \sum_{k=0}^{j-1} \frac{1}{r^{2(j-k)}} A_{r_2,k} \\
&\sim \sum_{k=0}^{m-2} \frac{1}{r^{2(m-k)}} A_{r_2,k} \lesssim \frac{1}{r^{2m}} A_{2r,0},
\end{aligned}$$

which immediately implies that Proposition 3.2 holds true.

We now turn to the proof of (3.13). Observe that, if $j = m$, then (3.13) can be proved by using the same argument as the proof of the parabolic Caccioppoli inequality (3.2) with r and $2r$ replaced, respectively, by ζ and ξ, noticing that the assumption $0 < \zeta < \xi \le 2r$, together with $t_0 \in (3r,\infty)$, implies that, for all $t \in (t_0-\xi, t_0+\xi)$,

$$\frac{1}{t} < \frac{1}{t_0-\xi} < \frac{1}{r} < \frac{2}{\xi-\zeta}.$$

Thus, by induction, to finish the proof of (3.13), it remains to show that, if (3.13) holds true for some $j+1$, then (3.13) also holds true for j.

Now, for all $i \in \mathbb{Z}_+$, let $\{\rho_i\}_{i\in\mathbb{Z}_+}$ be a sequence of increasing numbers satisfying

$$\zeta = \rho_0 < \rho_1 < \cdots < \xi,$$

$\delta_i := \rho_{i+1} - \rho_i$ and $\tilde{\rho}_i := \rho_i + \delta_i/2$, where the exact value of ρ_i will be determined later. Let $\varphi_i \in C_c^\infty(B(x_0, \tilde{\rho}_i))$ satisfy $\operatorname{supp} \varphi_i \subset B(x_0, \tilde{\rho}_i)$, $\varphi_i \equiv 1$ on $B(x_0, \rho_i)$, $\|\nabla\varphi_i\|_{L^\infty(\mathbb{R}^n)} \lesssim 1/\delta_i$ and $\|\nabla^2\varphi_i\|_{L^\infty(\mathbb{R}^n)} \lesssim 1/\delta_i^2$. By the properties of φ_i, the Fourier transform and Hölder's inequality, we know that, for all $i \in \mathbb{Z}_+$,

$$\begin{aligned}
\int_{t_0-\rho_i}^{t_0+\rho_i}\int_{B(x_0,\rho_i)} |\nabla^j u(x,t)|^2\,dx\,dt &\le \int_{t_0-\tilde{\rho}_i}^{t_0+\tilde{\rho}_i}\int_{B(x_0,\tilde{\rho}_i)} |\nabla(\varphi_i\nabla^{j-1}u)(x,t)|^2\,dx\,dt \\
&\lesssim \left\{\int_{t_0-\tilde{\rho}_i}^{t_0+\tilde{\rho}_i}\int_{B(x_0,\tilde{\rho}_i)} |\nabla^{j-1}u(x,t)|^2\,dx\,dt\right\}^{\frac12}\left\{\int_{t_0-\tilde{\rho}_i}^{t_0+\tilde{\rho}_i}\int_{B(x_0,\tilde{\rho}_i)} |\nabla^2(\varphi_i\nabla^{j-1}u)(x,t)|^2\,dx\,dt\right\}^{\frac12} \\
&\lesssim \left\{\int_{t_0-\tilde{\rho}_i}^{t_0+\tilde{\rho}_i}\int_{B(x_0,\tilde{\rho}_i)} |\nabla^{j-1}u(x,t)|^2\,dx\,dt\right\}^{\frac12} \\
&\quad\times\left\{\int_{t_0-\tilde{\rho}_i}^{t_0+\tilde{\rho}_i}\int_{B(x_0,\tilde{\rho}_i)} \left|\left(|\nabla^{j+1}u| + \frac{1}{\delta_i}|\nabla^j u| + \frac{1}{\delta_i^2}|\nabla^{j-1}u|\right)(x,t)\right|^2\,dx\,dt\right\}^{\frac12} \\
&\lesssim [A_{\tilde{\rho}_i,j-1}]^{\frac12}\left[A_{\tilde{\rho}_i,j+1} + \frac{1}{\delta_i^2}A_{\tilde{\rho}_i,j} + \frac{1}{\delta_i^4}A_{\tilde{\rho}_i,j-1}\right]^{\frac12},
\end{aligned}$$

where $A_{\tilde{\rho}_i,j+1}$, $A_{\tilde{\rho}_i,j}$ and $A_{\tilde{\rho}_i,j-1}$ are defined as in (3.14).

From the assumption that (3.13) holds true for $j+1$ and Hölder's inequality with ϵ, we further deduce that, for any $\epsilon \in (0,\infty)$, there exists a positive constant $C_{(\epsilon)}$ such that

$$\begin{aligned}
\int_{t_0-\rho_i}^{t_0+\rho_i}\int_{B(x_0,\rho_i)} |\nabla^j u(x,t)|^2\,dx\,dt &\lesssim [A_{\rho_{i+1},j-1}]^{\frac12}\left[\sum_{k=0}^{j}\frac{1}{\delta_i^{2(j+1-k)}}A_{\rho_{i+1},k} + \frac{1}{\delta_i^2}A_{\tilde{\rho}_i,j} + \frac{1}{\delta_i^4}A_{\tilde{\rho}_i,j-1}\right]^{\frac12} \\
&\lesssim \frac{C_{(\epsilon)}}{\delta_i^2}A_{\rho_{i+1},j-1} + \epsilon\sum_{k=0}^{j}\frac{1}{\delta_i^{2(j-k)}}A_{\rho_{i+1},k} + \epsilon A_{\rho_{i+1},j}.
\end{aligned}$$

By letting ϵ be small enough, we conclude that there exists a positive constant $\tilde{C}$ such that, for all $i \in \mathbb{Z}_+$,

$$A_{\rho_i,j} = \int_{t_0-\rho_i}^{t_0+\rho_i}\int_{B(x_0,\rho_i)} |\nabla^j u(x,t)|^2\,dx\,dt \le \tilde{C}\sum_{k=0}^{j-1}\frac{1}{\delta_i^{2(j-k)}}A_{\rho_{i+1},k} + \frac12 A_{\rho_{i+1},j} =: \tilde{C}B_{i+1,j} + \frac12 A_{\rho_{i+1},j}, \tag{3.15}$$

which immediately implies that

$$A_{\rho_0,j} \le \tilde{C}B_{1,j} + \frac12 A_{\rho_{1,j}} \le \tilde{C}B_{1,j} + \frac12\left[\tilde{C}B_{2,j} + \frac12 A_{\rho_{2,j}}\right] \le \tilde{C}\sum_{i=0}^{\infty}2^{-i}B_{i+1,j}. \tag{3.16}$$

Moreover, for all $i \in \mathbb{Z}_+$, take $\tau \in (2^{-1/(2m)}, 1)$ and $\rho_i := \zeta + (\xi-\zeta)(1-\tau)\sum_{s=1}^{i}\tau^s$. Then, we have that $\delta_i = (\xi-\zeta)(1-\tau)\tau^{i+1}$ and

$$\begin{aligned}
\sum_{i=0}^{\infty}2^{-i}B_{i+1,j} &= \sum_{i=0}^{\infty}2^{-i}\tilde{C}\sum_{k=0}^{j-1}\frac{1}{\delta_i^{2(j-k)}}A_{\rho_{i+1},k} \\
&\lesssim \sum_{i=0}^{\infty}2^{-i}\tilde{C}\sum_{k=0}^{j-1}\frac{1}{[(\xi-\zeta)(1-\tau)\tau^i]^{2(j-k)}}A_{\rho_{i+1},k} \\
&\lesssim \sum_{k=0}^{j-1}\sum_{i=0}^{\infty}\frac{1}{[2\tau^{2(j-k)}]^i}\frac{1}{(\xi-\zeta)^{2(j-k)}}A_{\rho_{i+1},k} \lesssim \sum_{k=0}^{j-1}\frac{1}{(\xi-\zeta)^{2(j-k)}}A_{\xi,k},
\end{aligned} \tag{3.17}$$

where the implicit positive constants depend on m and τ, but are independent of j, ξ and ζ.

Combining the estimates (3.16) and (3.17), and using the definition of $B_{i+1,j}$ in (3.15), we conclude that

$$\int_{t_0-\zeta}^{t_0+\zeta}\int_{B(x_0,\zeta)} |\nabla^j u(x,t)|^2\,dx\,dt = A_{\rho_0,j} \lesssim \sum_{i=0}^{\infty} 2^{-i}B_{i+1,j} \lesssim \sum_{k=0}^{j-1}\frac{1}{(\xi-\zeta)^{2(j-k)}}A_{\xi,k},$$

which immediately implies that (3.13) holds for j. Thus, by induction, (3.13) holds true for all $j \in \{1,\dots,m\}$, which completes the proof of Proposition 3.2. □

Now, for $p \in (0, p_+(L))$, we want to control the $H^p_L(\mathbb{R}^n)$ quasi-norm by the $L^p(\mathbb{R}^n)$ quasi-norm of $S_{h,L}$ in (1.18), via the parabolic Caccioppoli inequality (3.12). Before going further, we point out that, in the remainder of this section – including the proofs of Propositions 3.3 and 3.9, and Theorem 1.4 – we borrow some ideas from the corresponding parts of [28], in which the authors considered the case when $m = 1$ and $p = 1$.

We first need the following notation. For all $\lambda \in (0,\infty)$, $k \in \mathbb{Z}_+$ and $f \in L^2(\mathbb{R}^n)$, let $S^\lambda_{L,k}(f)$ and $S^\lambda_{h,L,k}(f)$ be the same, respectively, as in (1.17) and (1.18). For any $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$, let $\Gamma^{\epsilon,R,\lambda}(x)$ be the *truncated cone* defined by setting

$$\Gamma^{\epsilon,R,\lambda}(x) := \{(y,t) \in \mathbb{R}^n \times (\epsilon,R) : |x-y| < \lambda t\}. \tag{3.18}$$

We write $S^\lambda_{L,k}(f)(x)$ and $S^\lambda_{L,h,k}(f)(x)$, respectively, by $S^{\epsilon,R,\lambda}_{L,k}(f)(x)$ and $S^{\epsilon,R,\lambda}_{L,h,k}(f)(x)$ when the cone $\Gamma^\lambda(x)$, in (1.17) and (1.18), is replaced by $\Gamma^{\epsilon,R,\lambda}(x)$.

From [20, Proposition 4], it follows that, for all $k \in \mathbb{Z}_+$, $\lambda \in (0,\infty)$, $p \in (0,\infty)$ and $f \in L^2(\mathbb{R}^n)$,

$$\|S^\lambda_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)} \sim \|S_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)} \tag{3.19}$$

and

$$\|S^\lambda_{L,k}(f)\|_{L^p(\mathbb{R}^n)} \sim \|S_{L,k}(f)\|_{L^p(\mathbb{R}^n)}, \tag{3.20}$$

where the implicit positive constants are independent of f.

For $p \in (0,\infty)$, we can control the $L^p(\mathbb{R}^n)$ quasi-norm of S_L in (1.17) by that of $S_{h,L}$ in (1.18) as follows.

Proposition 3.3. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$, *and let* $p \in (0,\infty)$. *Then, there exists a positive constant C such that, for all* $f \in L^2(\mathbb{R}^n)$,

$$\|S_L(f)\|_{L^p(\mathbb{R}^n)} \le C\|S_{h,L}(f)\|_{L^p(\mathbb{R}^n)}.$$

In what follows, for all $k \in \mathbb{Z}_+$, $x \in \mathbb{R}^n$ and a suitable function H on $\mathbb{R}^{n+1}_+$, we let

$$\mathcal{A}_k(H)(x) := \left\{\iint_{\Gamma^{2^k}(x)} |H(y,t)|^2\,\frac{dy\,dt}{t^{n+1}}\right\}^{\frac12}. \tag{3.21}$$

Observe that $\mathcal{A}_0(H)$ is just the $\mathcal{A}$-functional $\mathcal{A}(H)$ defined as in (2.3) with F replaced by H.

To prove Proposition 3.3, we need the following technical lemma, which is due to Steve Hofmann (a personal communication with the second author).

Lemma 3.4. *Let* $F, G \in T^2(\mathbb{R}^{n+1}_+)$. *If there exists a positive constant* C_0 *such that, for all* $k \in \mathbb{Z}_+$ *and almost every* $x \in \mathbb{R}^n$,

$$\mathcal{A}_k(F)(x) \le C_0[\mathcal{A}_{k+1}(G)(x)]^{\frac12}[\mathcal{A}_{k+1}(F)(x)]^{\frac12}, \tag{3.22}$$

then, for all $p \in (0,\infty)$, *there exists a positive constant* C_1, *independent of F and G, such that*

$$\|F\|_{T^p(\mathbb{R}^{n+1}_+)} \le C_1\|G\|_{T^p(\mathbb{R}^{n+1}_+)}.$$

Proof. From [3, Theorem 1.1], it follows that there exists a positive constant $C_{(n,p)} \in [1,\infty)$, depending on n and p, but being independent of F, such that, for all $F \in T^2(\mathbb{R}^{n+1}_+)$,

$$\|\mathcal{A}_k(F)\|_{L^p(\mathbb{R}^n)} \le C_{(n,p)}\|\mathcal{A}_{k-1}(F)\|_{L^p(\mathbb{R}^n)} \le [C_{(n,p)}]^k\|\mathcal{A}(F)\|_{L^p(\mathbb{R}^n)}. \tag{3.23}$$

Moreover, $C_{(n,p)} \ge C_{(n,2)}$ for all $p \in (0,\infty)$.

Let $R \in (2C_{(n,p)}, \infty)$ and $\mathcal{A}_*(F) := \sum_{k=0}^{\infty} \frac{1}{R^k} \mathcal{A}_k(F)$. By (3.23), we know that

$$\|\mathcal{A}_*(F)\|_{L^2(\mathbb{R}^n)} \le \sum_{k=0}^{\infty} \frac{1}{R^k} \|\mathcal{A}_k(F)\|_{L^2(\mathbb{R}^n)} \le \sum_{k=0}^{\infty} \left[\frac{C_{(n,2)}}{R}\right]^k \|\mathcal{A}(F)\|_{L^2(\mathbb{R}^n)} = 2\|F\|_{T^2(\mathbb{R}^{n+1}_+)} < \infty,$$

which immediately implies that $\mathcal{A}_*(F)(x) < \infty$ almost everywhere in $\mathbb{R}^n$.

On the other hand, using (3.22) and Cauchy's inequality, we find that, for all $k \in \mathbb{Z}_+$ and almost every $x \in \mathbb{R}^n$,

$$\mathcal{A}_k(F)(x) \le \frac{1}{2} C_0^2 R \mathcal{A}_{k+1}(G)(x) + \frac{1}{2R} \mathcal{A}_{k+1}(F)(x),$$

which, together with the definition of $\mathcal{A}_*$, shows that

$$\mathcal{A}_*(F)(x) \le \frac{C_0^2 R^2}{2} \mathcal{A}_*(G)(x) + \frac{1}{2} \mathcal{A}_*(F)(x).$$

This, combined with the fact $\mathcal{A}_*(F)(x) < \infty$ almost everywhere in $\mathbb{R}^n$, further implies that

$$\mathcal{A}_*(F)(x) \le C_0^2 R^2 \mathcal{A}_*(G)(x).$$

Thus, by the fact that $\sum_{k=0}^{\infty} \frac{1}{R^k} = \frac{R}{R-1}$, the definition of the $\mathcal{A}$-functional in (2.3) and (3.23), we have

$$\begin{aligned}
\|F\|_{T^p(\mathbb{R}^{n+1}_+)} = \|\mathcal{A}(F)\|_{L^p(\mathbb{R}^n)} &= \frac{R-1}{R} \left\| \sum_{k=0}^{\infty} \frac{1}{R^k} \mathcal{A}(F) \right\|_{L^p(\mathbb{R}^n)} \le \frac{R-1}{R} \left\| \sum_{k=0}^{\infty} \frac{1}{R^k} \mathcal{A}_k(F) \right\|_{L^p(\mathbb{R}^n)} \le \frac{R-1}{R} \|\mathcal{A}_*(F)\|_{L^p(\mathbb{R}^n)} \\
&\le \frac{R-1}{R} \|C_0^2 R^2 \mathcal{A}_*(G)\|_{L^p(\mathbb{R}^n)} \le \frac{R-1}{R} \left\| \sum_{k=0}^{\infty} \frac{C_0^2 R^2}{R^k} \mathcal{A}_k(G) \right\|_{L^p(\mathbb{R}^n)} \le \frac{(R-1)C_0^2 R^2}{R} \sum_{k=0}^{\infty} \frac{1}{2^k} \|\mathcal{A}(G)\|_{L^p(\mathbb{R}^n)} \\
&\le \frac{2(R-1)C_0^2 R^2}{R} \|\mathcal{A}(G)\|_{L^p(\mathbb{R}^n)} = \frac{2(R-1)C_0^2 R^2}{R} \|G\|_{T^p(\mathbb{R}^{n+1}_+)},
\end{aligned}$$

which completes the proof of Lemma 3.4. □

With the help of Lemma 3.4, we now prove Proposition 3.3.

Proof of Proposition 3.3. We begin the proof of this proposition by first introducing some smooth cut-off functions supported in truncated cones. For all $0 < \epsilon \ll R < \infty$, $\lambda \in (0, \infty)$ and $x \in \mathbb{R}^n$, let $\Gamma^{\epsilon,R,\lambda}(x)$ be the truncated cone defined as in (3.18).

Let $\eta \in C_c^{\infty}(\Gamma^{\epsilon/2,2R,3/2}(x))$ satisfy $\eta \equiv 1$ on $\Gamma^{\epsilon,R,1}(x)$, $0 \le \eta \le 1$ and

$$|\nabla^k \eta(y, t)| \lesssim \frac{1}{t^k}$$

for all $k \in \mathbb{N}$ with $k \le m$ and $(y, t) \in \Gamma^{\epsilon/2,2R,3/2}(x)$. From the definition of L and Minkowski's inequality, we deduce that, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}
\left\{ \iint_{\Gamma^{\epsilon,R,1}(x)} |t^{2m} L e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{2}} &\le \left| \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} t^{2m} L e^{-t^{2m}L}(f)(y) \overline{t^{2m} L e^{-t^{2m}L}(f)(y)} \eta(y,t) \frac{dy\,dt}{t^{n+1}} \right|^{\frac{1}{2}} \\
&= \left| \sum_{|\alpha|=m=|\beta|} \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y) t^m \partial^{\beta}(e^{-t^{2m}L}(f))(y) \right. \\
&\qquad \left. \times \overline{t^m \partial^{\alpha}(t^{2m} L e^{-t^{2m}L}(f)\eta)(y,t)} \frac{dy\,dt}{t^{n+1}} \right|^{\frac{1}{2}} \\
&\lesssim \sum_{k=0}^{m} \left| \sum_{|\alpha|=m=|\beta|} \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y) t^m \partial^{\beta}(e^{-t^{2m}L}(f))(y) \right. \\
&\qquad \left. \times \left[t^m \sum_{|\tilde{\alpha}|=k, \tilde{\alpha} \le \alpha} C_{(\alpha,\tilde{\alpha})} \overline{\partial^{\tilde{\alpha}}(t^{2m} L e^{-t^{2m}L}(f))(y) \partial^{\alpha-\tilde{\alpha}} \eta(y,t)} \right] \frac{dy\,dt}{t^{n+1}} \right|^{\frac{1}{2}} \\
&=: \sum_{k=0}^{m} \mathrm{I}_k. \qquad (3.24)
\end{aligned}$$

We first bound I_0. By Hölder's inequality, the size condition of η and the ellipticity condition $(\mathcal{E}_0)$, we see that, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}\mathrm{I}_0 &\lesssim \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^m\nabla^m(e^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14} \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^{2m}Le^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14}\\ &\lesssim [S_{h,L}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12}[S_L^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12},\end{aligned} \tag{3.25}$$

where $S_{h,L}^{\epsilon/2,2R,3/2}(f)$ and $S_L^{\epsilon/2,2R,3/2}(f)$ are defined, respectively, similar to $S_{h,L}(f)$ in (1.18) and $S_L(f)$ in (1.17), with $\Gamma(x)$ replaced by $\Gamma^{\epsilon/2,2R,3/2}(x)$.

To bound I_m, similar to (3.25), for all $x \in \mathbb{R}^n$, we have

$$\begin{aligned}\mathrm{I}_m &\lesssim \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^m\nabla^m(e^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14} \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^m\nabla^m(t^{2m}Le^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14}\\ &\sim [S_{h,L}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12}[S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12}.\end{aligned} \tag{3.26}$$

To bound $S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)$, let $Q(z,2r)$ be the cube with center z and sidelength $2r$ in $\mathbb{R}_+^{n+1}$. Write $z := (z^*, t)$ with $z^* \in \mathbb{R}^n$ and $t \in (0,\infty)$. Assume that $\{Q(z_j, 2r_j)\}_{j\in\mathbb{N}}$ is a covering of $\Gamma^{\epsilon/2,2R,3/2}(x)$ satisfying

$$\Gamma^{\epsilon/2,2R,3/2}(x) \subset \bigcup_{j\in\mathbb{N}} Q(z_j, 2r_j) \subset \bigcup_{j\in\mathbb{N}} Q(z_j, 4\sqrt{n}r_j) \subset \Gamma^{\epsilon/4,3R,2}(x),$$

$$d\Big(z_j, (\Gamma^{\epsilon/4,3R,2}(x))^\complement\Big) \sim r_j \sim d(z_j, \{t=0\}), \quad j \in \mathbb{N},$$

and the collection $\{B(z_j^*, \sqrt{n}r_j) \times (t_j - \sqrt{n}r_j, t_j + \sqrt{n}r_j)\}_{j\in\mathbb{N}}$ has a bounded overlap, where, for all $j \in \mathbb{N}$, $z_j := (z_j^*, t_j)$. This kind of covering is based on Whitney's decomposition; see [28, equation (5.26)] for a covering of similar nature consisting of balls.

It is easy to see that

$$\begin{aligned}\Gamma^{\epsilon/2,2R,3/2}(x) &\subset \bigcup_{j\in\mathbb{N}} Q(z_j, 2r_j) \subset \bigcup_{j\in\mathbb{N}} B(z_j^*, \sqrt{n}r_j) \times (t_j - \sqrt{n}r_j, t_j + \sqrt{n}r_j)\\ &\subset \bigcup_{j\in\mathbb{N}} B(z_j^*, 2\sqrt{n}r_j) \times (t_j - 2\sqrt{n}r_j, t_j + 2\sqrt{n}r_j)\\ &\subset \bigcup_{j\in\mathbb{N}} Q(z_j, 4\sqrt{n}r_j)\\ &\subset \Gamma^{\epsilon/4,3R,2}(x).\end{aligned}$$

From these and the parabolic Caccioppoli inequality (3.12), we deduce that, for all $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}[S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)]^2 &\lesssim \sum_{j\in\mathbb{N}} \int_{t_j-\sqrt{n}r_j}^{t_j+\sqrt{n}r_j} \int_{B(z_j^*,\sqrt{n}r_j)} |t^m\nabla^m(t^{2m}Le^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}}\\ &\lesssim \sum_{j\in\mathbb{N}} \frac{1}{r_j^{2m}} \int_{t_j-2\sqrt{n}r_j}^{t_j+2\sqrt{n}r_j} \int_{B(z_j^*,2\sqrt{n}r_j)} |t^{3m}Le^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}}\\ &\lesssim \iint_{\Gamma^{\epsilon/4,3R,2}(x)} |t^{2m}Le^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}}\\ &\sim [S_L^{\epsilon/4,3R,2}(f)(x)]^2.\end{aligned} \tag{3.27}$$

Moreover, by (3.26) and (3.27), we conclude that, for all $x \in \mathbb{R}^n$,

$$\mathrm{I}_m \lesssim [S_{h,L}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12}[S_L^{\epsilon/4,3R,2}(f)(x)]^{\frac12}. \tag{3.28}$$

We now turn to the estimates of I_k for all $k \in \{1, \dots, m-1\}$. Again, by Hölder's inequality, the ellipticity condition $(\mathcal{E}_0)$ and the size condition of η, we see that

$$\mathrm{I}_k \lesssim \left\{ \iint\limits_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^m \nabla^m (e^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac14} \left\{ \iint\limits_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^k \nabla^k (t^{2m} L e^{-t^{2m}L}(f))(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac14}$$
$$=: [S_{h,L}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12} \times (\mathrm{II}_k)^{\frac14}.$$

To bound II_k, using again the interpolation inequality (3.11) and Hölder's inequality, we conclude that

$$\begin{aligned}
\mathrm{II}_k &\sim \int_{\epsilon/2}^{2R} \|t^k \nabla^k (t^{2m} L e^{-t^{2m}L}(f))\|_{L^2(B(x,(3/2)t))}^2 \frac{dt}{t^{n+1}} \\
&\lesssim \int_{\epsilon/2}^{2R} \|t^m \nabla^m (t^{2m} L e^{-t^{2m}L}(f))\|_{L^2(B(x,(3/2)t))}^{2k/m} \|t^{2m} L e^{-t^{2m}L}(f)\|_{L^2(B(x,(3/2)t))}^{2(m-k)/m} \frac{dt}{t^{n+1}} \\
&\lesssim \left\{ \int_{\epsilon/2}^{2R} \|t^m \nabla^m (t^{2m} L e^{-t^{2m}L}(f))\|_{L^2(B(x,(3/2)t))}^2 \frac{dt}{t^{n+1}} \right\}^{\frac{k}{m}} \left\{ \int_{\epsilon/2}^{2R} \|t^{2m} L e^{-t^{2m}L}(f)\|_{L^2(B(x,(3/2)t))}^2 \frac{dt}{t^{n+1}} \right\}^{\frac{m-k}{m}} \\
&\sim [S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac{2k}{m}} [S_L^{\epsilon/2,2R,3/2}(f)(x)]^{\frac{2(m-k)}{m}}.
\end{aligned}$$

From this and (3.27), we deduce that

$$\mathrm{II}_k \lesssim [S_L^{\epsilon/4,3R,2}(f)(x)]^{\frac{2k}{m}} [S_L^{\epsilon/2,2R,3/2}(f)(x)]^{\frac{2(m-k)}{m}}.$$

Thus, it holds true that

$$\mathrm{I}_k \lesssim [S_{h,L}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12} [S_L^{\epsilon/4,3R,2}(f)(x)]^{\frac{k}{2m}} [S_L^{\epsilon/2,2R,3/2}(f)(x)]^{\frac{m-k}{2m}}. \tag{3.29}$$

Combining the estimates (3.24), (3.25), (3.28) and (3.29), the fact that $f \in L^2(\mathbb{R}^n)$ and letting $\epsilon \to 0$ and $R \to \infty$, we conclude that, for almost every $x \in \mathbb{R}^n$,

$$\begin{aligned}
S_L(f)(x) \lesssim{}& [S_{h,L}^{3/2}(f)(x)]^{\frac12} [S_L^{3/2}(f)(x)]^{\frac12} + [S_{h,L}^{3/2}(f)(x)]^{\frac12} [S_L^2(f)(x)]^{\frac12} \\
&+ \sum_{k=1}^{m-1} [S_{h,L}^{3/2}(f)(x)]^{\frac12} [S_L^2(f)(x)]^{\frac{k}{2m}} [S_L^{3/2}(f)(x)]^{\frac{m-k}{2m}},
\end{aligned}$$

which immediately shows that there exists a positive constant C_0 such that, for almost every $x \in \mathbb{R}^n$,

$$S_L(f)(x) \le C_0 [S_{h,L}^2(f)(x)]^{\frac12} [S_L^2(f)(x)]^{\frac12}. \tag{3.30}$$

Similarly, by following the same line of the proof of (3.30), we conclude that, for all $k \in \mathbb{Z}_+$ and almost every $x \in \mathbb{R}^n$,

$$S_L^{2^k}(f)(x) \le C_0 [S_{h,L}^{2^{k+1}}(f)(x)]^{\frac12} [S_L^{2^{k+1}}(f)(x)]^{\frac12},$$

which, combined with the definition of $\mathcal{A}_k$ in (3.21), implies that, for all $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$\mathcal{A}_k(F)(x) \le C_0 [\mathcal{A}_{k+1}(G)(x)]^{\frac12} [\mathcal{A}_{k+1}(F)(x)]^{\frac12},$$

where $F := t^{2m} L e^{-t^{2m}L}(f)$ and $G := (t\nabla)^m e^{-t^{2m}L}(f)$. Moreover, since $f \in L^2(\mathbb{R}^n)$, we know that $F, G \in T^2(\mathbb{R}^{n+1}_+)$. Thus, by Lemma 3.4, we conclude that, for all $p \in (0, \infty)$ and $f \in L^2(\mathbb{R}^n)$, $\|F\|_{T^p(\mathbb{R}^{n+1}_+)} \lesssim \|G\|_{T^p(\mathbb{R}^{n+1}_+)}$, which implies that, for all $p \in (0, \infty)$ and $f \in L^2(\mathbb{R}^n)$, $\|S_L(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|S_{h,L}(f)\|_{L^p(\mathbb{R}^n)}$ and hence completes the proof of Proposition 3.3. □

We also need the boundedness of $S_{L,k}$ and $S_{h,L,k}$ in $L^q(\mathbb{R}^n)$ as follows, which, when $k = 1$, was pointed out in [2, p. 68] without any details.

Lemma 3.5. *Let L be as in* (1.1) *and satisfying the ellipticity condition* ($\mathcal{E}_0$). *Let $S_{L,k}$ and $S_{h,L,k}$ be the same, respectively, as in* (1.17) *and* (1.18). *Then, the following hold:*

(i) *For all $k \in \mathbb{N}$ and $q \in (p_-(L), p_+(L))$, there exists a positive constant C such that, for all $f \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$,*

$$\|S_{L,k}(f)\|_{L^q(\mathbb{R}^n)} \leq C\|f\|_{L^q(\mathbb{R}^n)}.$$

(ii) *For all $k \in \mathbb{Z}_+$ and $q \in (q_-(L), q_+(L))$, there exists a positive constant C such that, for all $f \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$,*

$$\|S_{h,L,k}(f)\|_{L^q(\mathbb{R}^n)} \leq C\|f\|_{L^q(\mathbb{R}^n)}.$$

Proof. Observe that, by Proposition 2.5 (iii), we know that L satisfies all the assumptions of [10, Theorem 2.13] and, as a consequence, we obtain (i) of Lemma 3.5.

The proof of (ii) of this lemma is similar to that of (i). We only need to replace the m-L^p-L^q off-diagonal estimates from Proposition 2.5 (iii), in the proof of [10, Theorem 2.13], by the corresponding m-L^p-L^q off-diagonal estimates of the gradient semigroups from Proposition 2.5 (iv), the details being omitted. This finishes the proof of Lemma 3.5. □

The next proposition presents an equivalence between the $H_L^p(\mathbb{R}^n)$ norm, defined via the square function S_L, and the $L^p(\mathbb{R}^n)$ norm when $p \in (p_+(L), p_+(L))$. Recall that this conclusion was pointed out in [2, p. 68] without giving any details.

Proposition 3.6. *Let L be as in* (1.1) *and satisfying the ellipticity condition* ($\mathcal{E}_0$), *$p \in (p_-(L), p_+(L))$ with $p_-(L)$ and $p_+(L)$ as in Proposition 2.5, and S_L be as in* (1.17) *with $k = 1$ and $\lambda = 1$. Then, there exists a positive constant C such that, for all $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$,*

$$\frac{1}{C}\|f\|_{L^p(\mathbb{R}^n)} \leq \|S_L(f)\|_{L^p(\mathbb{R}^n)} \leq C\|f\|_{L^p(\mathbb{R}^n)}. \tag{3.31}$$

Proof. It is easy to see that the second inequality of (3.31) is a direct consequence of Lemma 3.5 (i) in the case when $k = 1$.

We now prove the first inequality of (3.31). Let $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. For all $p \in (p_-(L), p_+(L))$ and $g \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n)$ satisfying $\|g\|_{L^{p'}(\mathbb{R}^n)} \leq 1$ with $1/p + 1/p' = 1$, by the Calderón reproducing formula (2.4) with $M = 0$, duality, Fubini's theorem, Hölder's inequality and Lemma 3.5 (i), we have that

$$\begin{aligned} |\langle f, g\rangle_{L^2(\mathbb{R}^n)}| &\sim \left|\left\langle \int_0^\infty (t^{2m}L)^2 e^{-2t^{2m}L}(f)\,\frac{dt}{t}, g\right\rangle_{L^2(\mathbb{R}^n)}\right| \\ &\sim \left|\int_0^\infty \langle t^{2m}Le^{-t^{2m}L}(f), t^{2m}L^* e^{-t^{2m}L^*}(g)\rangle_{L^2(\mathbb{R}^n)}\,\frac{dt}{t}\right| \\ &\sim \|S_L(f)\|_{L^p(\mathbb{R}^n)}\|S_{L^*}(g)\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|S_L(f)\|_{L^p(\mathbb{R}^n)} \end{aligned} \tag{3.32}$$

and hence

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \|S_L(f)\|_{L^p(\mathbb{R}^n)}.$$

This finishes the proof of the first inequality of (3.31) and hence Proposition 3.6. □

Remark 3.7. For all $p \in (0, \infty)$, let $\mathbb{H}_L^p(\mathbb{R}^n)$ be the space defined as in Definition 1.1. From Proposition 3.6 and the definition of $\mathbb{H}_L^p(\mathbb{R}^n)$, it follows that, for all $p \in (p_-(L), p_+(L))$,

$$\mathbb{H}_L^p(\mathbb{R}^n) = L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n).$$

With the help of Propositions 3.3 and 3.6, we obtain the following corollary.

Corollary 3.8. *Let L be as in* (1.1) *and satisfying the ellipticity condition* ($\mathcal{E}_0$). *Then, for all $p \in (0, p_+(L))$, there exists a positive constant $C_{(p)}$, depending on p, such that, for all $f \in L^2(\mathbb{R}^n)$,*

$$\|f\|_{H_L^p(\mathbb{R}^n)} \leq C_{(p)}\|S_{h,L}(f)\|_{L^p(\mathbb{R}^n)}. \tag{3.33}$$

Proof. If $p \in (0, 2]$, Corollary 3.8 is an immediately consequence of Proposition 3.3 and Definition 1.1.

If $p \in (2, p_+(L))$, Corollary 3.8 follows from Lemma 2.12 and Propositions 3.3 and 3.6. This finishes the proof of Corollary 3.8. □

The next proposition shows that the $L^p(\mathbb{R}^n)$ quasi-norm of $S_{h,L}$, as in (1.18) with $\lambda = 1$, can be controlled by that of the non-tangential maximal function $\mathcal{N}^{\gamma}_{h,L}$ as in (1.7).

Proposition 3.9. *Let L be as in* (1.1) *and satisfying the strong ellipticity condition* $(\mathcal{E}_1)$*, and let $p \in (0,2)$. Then, there exist positive constants γ and C such that, for all $f \in L^2(\mathbb{R}^n)$,*

$$\|S_{h,L}(f)\|_{L^p(\mathbb{R}^n)} \le C\|\mathcal{N}^{\gamma}_{h,L}(f)\|_{L^p(\mathbb{R}^n)}.$$

Proof. To prove Proposition 3.9, we first introduce some notation. Let $\sigma \in (0,\infty)$. Assume that $\gamma \in (0,\infty)$, whose exact value will be determined later. Let

$$E := \{x \in \mathbb{R}^n : \mathcal{N}^{\gamma}_{h,L}(f)(x) \le \sigma\}.$$

Its subset E^* *of global* $1/2$ *density* is defined by

$$E^* := \Big\{x \in \mathbb{R}^n : \text{for all balls } B(x,r) \in \mathbb{R}^n,\ \frac{|E \cap B(x,r)|}{|B(x,r)|} \ge \frac{1}{2}\Big\}.$$

For all $0 < \epsilon \ll R < \infty$, let $\mathcal{R}^{\epsilon,R,\gamma}(E^*) := \bigcup_{x\in E^*} \Gamma^{\epsilon,R,\gamma}(x)$ be the *sawtooth region based on* E^* and $\mathcal{B}^{\epsilon,R,\gamma}(E^*)$ the boundary of $\mathcal{R}^{\epsilon,R,\gamma}(E^*)$. Moreover, for all $y \in \mathbb{R}^n$ and $t \in (0,\infty)$, let $u(y,t) := e^{-t^{2m}L}(f)(y)$. By Fubini's theorem, we find that

$$\int_{E^*} [S^{\epsilon,R,1/2}_{h,L}(f)(x)]^2\,dx \sim \iint_{\mathcal{R}^{\epsilon,R,1/2}(E^*)} t^{2m}|\nabla^m u(y,t)|^2\,\frac{dy\,dt}{t},$$

where $S^{\epsilon,R,1/2}_{h,L}(f)(x)$ is defined as in (3.26).

Now, let $\eta \in C^\infty_c(\mathcal{R}^{\epsilon/2,2R,3/2}(E^*))$ be a smooth cut-off function satisfying $0 \le \eta \le 1$, $\eta \equiv 1$ on $\mathcal{R}^{\epsilon,R,1/2}(E^*)$ and, for all $k \in \mathbb{N}$ with $k \le m$ and $(x,t) \in \mathcal{R}^{\epsilon/2,2R,3/2}(E^*)$,

$$|\nabla^k_x \eta(x,t)| \lesssim \frac{1}{t^k} \quad\text{and}\quad |\partial_t \eta(x,t)| \lesssim \frac{1}{t}.$$

These assumptions, together with the strong ellipticity condition $(\mathcal{E}_1)$, imply that

$$\begin{aligned}\iint_{\mathcal{R}^{\epsilon,R,1/2}(E^*)} t^{2m}|\nabla^m u(y,t)|^2\,\frac{dy\,dt}{t} &\lesssim \mathrm{Re}\Big\{\iint_{\mathcal{R}^{\epsilon,R,1/2}(E^*)} \Big[t^{2m}\sum_{|\alpha|=m=|\beta|} a_{\alpha,\beta}(y)\partial^\beta u(y,t)\overline{\partial^\alpha u(y,t)}\Big]\frac{dy\,dt}{t}\Big\}\\ &\lesssim \mathrm{Re}\Big\{\iint_{\mathbb{R}^{n+1}_+} \Big[t^{2m}\sum_{|\alpha|=m=|\beta|} a_{\alpha,\beta}(y)\partial^\beta u(y,t)\overline{\partial^\alpha u(y,t)}\eta(y,t)\Big]\frac{dy\,dt}{t}\Big\}.\end{aligned}$$

From this, we further deduce that

$$\begin{aligned}\iint_{\mathcal{R}^{\epsilon,R,1/2}(E^*)} t^{2m}|\nabla^m u(y,t)|^2\,\frac{dy\,dt}{t} &\lesssim \Big|\mathrm{Re}\Big\{\iint_{\mathbb{R}^{n+1}_+} \Big[t^{2m}\sum_{|\alpha|=m=|\beta|}(-1)^m\partial^\alpha(\eta a_{\alpha,\beta}\partial^\beta u)(y,t)\overline{u(y,t)}\Big]\frac{dy\,dt}{t}\Big\}\Big|\\ &\lesssim \sum_{k=0}^m \Big|\mathrm{Re}\Big\{\iint_{\mathbb{R}^{n+1}_+}\Big[t^{2m}\sum_{|\alpha|=m=|\beta|}\ \sum_{|\tilde\alpha|=k,\,\tilde\alpha\le\alpha} C_{(\alpha,\tilde\alpha)}(-1)^m\partial^{\tilde\alpha}\eta(y,t)\\ &\qquad\times \partial^{\alpha-\tilde\alpha}(a_{\alpha,\beta}\partial^\beta u)(y,t)\overline{u(y,t)}\Big]\frac{dy\,dt}{t}\Big\}\Big|\\ &=: \sum_{k=0}^m \mathrm{J}_k,\end{aligned} \tag{3.34}$$

where, for any multi-indices α and $\tilde\alpha$ as above, $C_{(\alpha,\tilde\alpha)}$ denotes a positive constant depending on α and $\tilde\alpha$.

We first bound J_0. Since, for all $(y,t) \in \mathbb{R}^{n+1}_+$, $\frac{\partial}{\partial t}u(y,t) = -2mt^{2m-1}L(u)(y,t)$, we have that

$$\begin{aligned}\frac{\partial}{\partial t}|u(y,t)|^2 &= -2mt^{2m-1}L(u)(y,t)\overline{u(y,t)} - 2mt^{2m-1}u(y,t)\overline{L(u)(y,t)}\\ &= -4mt^{2m-1}\,\mathrm{Re}\{L(u)(y,t)\overline{u(y,t)}\},\end{aligned}$$

which, together with integration by parts and the properties of the cut-off function η, shows that

$$\mathrm{J}_0 \sim \left|\mathrm{Re}\left\{\iint_{\mathbb{R}^{n+1}_+} t^{2m-1}\eta(y,t)L(u)(y,t)\overline{u(y,t)}\,dy\,dt\right\}\right|$$
$$\sim \left|\iint_{\mathbb{R}^{n+1}_+} \eta(y,t)\frac{\partial}{\partial t}|u(y,t)|^2\,dy\,dt\right| \lesssim \iint_{\mathcal{R}^{\epsilon/2,2R,3/2}(E^*)\setminus\mathcal{R}^{\epsilon,R,1/2}(E^*)} |u(y,t)|^2\,\frac{dy\,dt}{t}.$$

To estimate the last term in the above formulae, we let

$$\tilde{\mathcal{B}}^{\epsilon}(E^*) := \left\{(x,t)\in\mathbb{R}^n\times\left(\frac{\epsilon}{2},\epsilon\right) : d(x,E^*) < \frac{3}{2}t\right\}, \tag{3.35}$$
$$\tilde{\mathcal{B}}^{R}(E^*) := \left\{(x,t)\in\mathbb{R}^n\times(R,2R) : d(x,E^*) < \frac{3}{2}t\right\} \tag{3.36}$$

and

$$\tilde{\mathcal{B}}_0(E^*) := \left\{(x,t)\in\mathbb{R}^n\times\left(\frac{\epsilon}{2},2R\right) : \frac{1}{2}t \le d(x,E^*) < \frac{3}{2}t\right\}. \tag{3.37}$$

It is easy to see that $[\mathcal{R}^{\epsilon/2,2R,3/2}(E^*)\setminus\mathcal{R}^{\epsilon,R,1/2}(E^*)] \subset [\tilde{\mathcal{B}}^{\epsilon}(E^*)\cup\tilde{\mathcal{B}}^{R}(E^*)\cup\tilde{\mathcal{B}}_0(E^*)]$. For any $(y,t)\in\tilde{\mathcal{B}}^{\epsilon}(E^*)$, we find that there exists $x\in E^*$ such that $|x-y| < \frac{3}{2}t$. Moreover, from the definition of E^*, it follows that, for all $t\in(0,\infty)$,

$$|E\cap B(x,t)| \ge \frac{1}{2}|B(x,t)| = \frac{1}{2}\omega_n t^n,$$

where $\omega_n := |B(x,1)| = |B(0,1)|$. Thus, $|E\cap B(y,3t)| \ge \frac{1}{2}\omega_n t^n$, which, combined with Fubini's theorem, implies that

$$\begin{aligned}
\iint_{\tilde{\mathcal{B}}^{\epsilon}(E^*)} |u(y,t)|^2\,\frac{dy\,dt}{t} &\lesssim \iint_{\tilde{\mathcal{B}}^{\epsilon}(E^*)} \left[\int_{E\cap B(y,3t)} |u(y,t)|^2\,dz\right]\frac{dy\,dt}{t^{n+1}}\\
&\lesssim \int_{\epsilon/2}^{\epsilon}\int_E \left[\frac{1}{t^n}\int_{B(z,3t)} |e^{-t^{2m}L}(f)(y)|^2\,dy\right]\frac{dz\,dt}{t}\\
&\lesssim \int_{\epsilon/2}^{\epsilon}\int_E \left|\sup_{(x,t)\in\Gamma^3(z)}\left\{\frac{1}{\omega_n(3t)^n}\int_{B(x,3t)} |e^{-t^{2m}L}(f)(y)|^2\,dy\right\}^{\frac{1}{2}}\right|^2\frac{dz\,dt}{t}\\
&\sim \int_{\epsilon/2}^{\epsilon}\int_E |\mathcal{N}^3_{h,L}(f)(z)|^2\,\frac{dz\,dt}{t}\\
&\sim \int_E |\mathcal{N}^3_{h,L}(f)(z)|^2\,dz.
\end{aligned} \tag{3.38}$$

Similarly, we have

$$\iint_{\tilde{\mathcal{B}}^{R}(E^*)} |u(y,t)|^2\,\frac{dy\,dt}{t} \lesssim \int_E |\mathcal{N}^3_{h,L}(f)(z)|^2\,dz. \tag{3.39}$$

To estimate the integrand on the region $\tilde{\mathcal{B}}_0(E^*)$, let $\{B(x_k,r_k)\}_k$ be Whitney's covering of B^*, where $B^* := (E^*)^{\complement}$. Then, the following hold:

(i) We have that $\bigcup_k B(x_k,r_k) = B^*$.

(ii) There exist positive constants C_1 and $C_2\in(0,1)$ such that, for all k,

$$C_1 d(x_k,E^*) \le r_k \le C_2 d(x_k,E^*).$$

(iii) There exists a positive constant C_3 such that, for all $x\in B^*$,

$$\sum_k \chi_{B(x_k,r_k)}(x) \le C_3.$$

From these, we deduce that

$$\iint_{\tilde{\mathcal{B}}_0(E^*)} |u(y,t)|^2 \frac{dy\,dt}{t} \lesssim \sum_k \int_{2r_k/3(1/C_2-1)}^{2r_k(1/C_1+1)} \int_{B(x_k,r_k)} |u(y,t)|^2 \frac{dy\,dt}{t}$$
$$\lesssim \sum_k \int_{2r_k/3(1/C_2-1)}^{2r_k(1/C_1+1)} r_k^n \bigg[\frac{1}{t^n} \int_{B(x_k,r_k)} |u(y,t)|^2\,dy\bigg] \frac{dt}{t}.$$

By the fact $E^* \subset E$, we know that $d(x_k, E) \le d(x_k, E^*) \le \frac{C_2}{(1-C_2)C_1} t$. Thus, by taking $\gamma \in (\frac{C_2}{(1-C_2)C_1}, \infty)$, we conclude that

$$\iint_{\tilde{\mathcal{B}}_0(E^*)} |u(y,t)|^2 \frac{dy\,dt}{t} \lesssim \sum_k r_k^n \Big[\sup_{z\in E} \mathcal{N}_{h,L}^{\gamma}(f)(z)\Big]^2 \lesssim |B^*| \Big[\sup_{z\in E} \mathcal{N}_{h,L}^{\gamma}(f)(z)\Big]^2. \tag{3.40}$$

Combining the estimates (3.38), (3.39) and (3.40), we see that

$$\mathrm{J}_0 \lesssim \int_E |\mathcal{N}_{h,L}^3(f)(z)|^2\,dz + |B^*| \Big[\sup_{z\in E} \mathcal{N}_{h,L}^{\gamma}(f)(z)\Big]^2. \tag{3.41}$$

Now, we turn to the estimates of J_k for all $k \in \{1, \dots, m\}$. Using integration by parts and Hölder's inequality, we write

$$\mathrm{J}_k \sim \bigg|\mathrm{Re}\bigg\{\iint_{\mathbb{R}_+^{n+1}} t^{2m} \sum_{|\alpha|=m=|\beta|} \sum_{|\tilde{\alpha}|=k,\,\tilde{\alpha}\le\alpha} \tilde{C}_{(\alpha,\tilde{\alpha},m)} a_{\alpha,\beta}(y) \partial^{\beta} u(y,t) \overline{\partial^{\alpha-\tilde{\alpha}}((\partial^{\tilde{\alpha}}\eta)u)(y,t)} \frac{dy\,dt}{t}\bigg\}\bigg|$$
$$\lesssim \sum_{|\alpha|=m=|\beta|} \sum_{|\tilde{\alpha}|=k,\,\tilde{\alpha}\le\alpha} \bigg\{\iint_{\mathcal{R}^{\epsilon/2,2R,3/2}(E^*)\setminus\mathcal{R}^{\epsilon,R,1/2}(E^*)} |t^m \partial^{\beta} u(y,t)|^2 \frac{dy\,dt}{t}\bigg\}^{\frac{1}{2}}$$
$$\times \bigg\{\int_{\mathcal{R}^{\epsilon/2,2R,3/2}(E^*)\setminus\mathcal{R}^{\epsilon,R,1/2}(E^*)} |t^{m-k} \partial^{\alpha-\tilde{\alpha}}(t^k[\partial^{\tilde{\alpha}}\eta]u)(y,t)|^2 \frac{dy\,dt}{t}\bigg\}^{\frac{1}{2}}$$
$$=: \sum_{|\alpha|=m=|\beta|} \sum_{|\tilde{\alpha}|=k,\,\tilde{\alpha}\le\alpha} \mathrm{J}_{\alpha,\tilde{\alpha},\beta,1} \times \mathrm{J}_{\alpha,\tilde{\alpha},\beta,2}, \tag{3.42}$$

where, for any $m \in \mathbb{N}$ and any multi-indices α and $\tilde{\alpha}$ as above, $\tilde{C}_{(\alpha,\tilde{\alpha},m)}$ denotes a constant depending on α, $\tilde{\alpha}$ and m.

We first control $\mathrm{J}_{\alpha,\tilde{\alpha},\beta,1}$. Let $\tilde{\mathcal{B}}^{\epsilon}(E^*)$, $\tilde{\mathcal{B}}^R(E^*)$ and $\tilde{\mathcal{B}}_0(E^*)$ be, respectively, as in (3.35), (3.36) and (3.37). Similar to (3.38), we obtain

$$\iint_{\tilde{\mathcal{B}}^{\epsilon}(E^*)} |t^m \nabla^m u(y,t)|^2 \frac{dy\,dt}{t} \lesssim \int_E \bigg[\int_{\epsilon/2}^{\epsilon} \frac{1}{t^n} \int_{B(z,3t)} |t^m \nabla^m u(y,t)|^2 \frac{dy\,dt}{t}\bigg]\,dz.$$

This, together with the parabolic Caccioppoli inequality (3.12), implies that

$$\iint_{\tilde{\mathcal{B}}^{\epsilon}(E^*)} |t^m \nabla^m u(y,t)|^2 \frac{dy\,dt}{t} \lesssim \int_E \int_{\epsilon/4}^{2\epsilon} \frac{1}{t^{n+2m}} \int_{B(z,6t)} [|t^m u(y,t)|^2] \frac{dy\,dt\,dz}{t}$$
$$\lesssim \int_E [\mathcal{N}_{h,L}^6(f)(y)]^2\,dy. \tag{3.43}$$

Similarly, resting on estimates (3.39), (3.40), (3.43) and the parabolic Caccioppoli inequality (3.12), we conclude that there exists a positive constant $\gamma \in (0,\infty)$ large enough such that

$$\iint_{\tilde{\mathcal{B}}^R(E^*)\cup\tilde{\mathcal{B}}_0(E^*)} |t^m \nabla^m u(y,t)|^2 \frac{dy\,dt}{t} \lesssim \int_E [\mathcal{N}_{h,L}^{\gamma}(f)(y)]^2\,dy + |B^*| \Big[\sup_{x\in E} \mathcal{N}_{h,L}^{\gamma}(f)(x)\Big]^2.$$

This, combined with (3.43), implies that

$$\mathrm{J}_{\alpha,\tilde{\alpha},\beta,1} \lesssim \left\{\int_E [\mathcal{N}^{\gamma}_{h,L}(f)(y)]^2\,dy + |B^*|\Big[\sup_{x\in E}\mathcal{N}^{\gamma}_{h,L}(f)(x)\Big]^2\right\}^{\frac{1}{2}}. \tag{3.44}$$

The estimate of $\mathrm{J}_{\alpha,\tilde{\alpha},\beta,2}$ can be obtained by using the definition of η, the interpolation inequality (3.11) and the estimates of J_0 and $\mathrm{J}_{\alpha,\tilde{\alpha},\beta,1}$. By the estimate of $\mathrm{J}_{\alpha,\tilde{\alpha},\beta,2}$, (3.34) and the estimates (3.41), (3.42) and (3.44), we see that

$$\int_{\mathcal{R}^{2\epsilon,R,1/2}(E^*)} t^{2m}|\nabla^m u(y,t)|^2\,\frac{dy\,dt}{t} \lesssim \int_E [\mathcal{N}^{\gamma}_{h,L}(f)(y)]^2\,dy + |B^*|\Big[\sup_{x\in E}\mathcal{N}^{\gamma}_{h,L}(f)(x)\Big]^2,$$

where $\gamma\in(0,\infty)$ is a sufficient large constant. By this and an argument similar to that used in [28, equations (6.31)–(6.37)], we complete the proof of Proposition 3.9. □

We are now in a position to prove our main result of this article.

Proof of Theorem 1.4. The inclusion that $H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)\subset H^p_L(\mathbb{R}^n)$, for all $p\in(0,2)$, is a direct consequence of Propositions 3.3 and 3.9, and Corollary 3.8. We now turn to the proof of the inclusion $H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)\subset H^p_L(\mathbb{R}^n)$ for all $p\in[2,p_+(L))$. Using Lemma 2.12, we are reduced to proving that, for all $p\in[2,p_+(L))$ and $f\in\mathbb{H}^p_L(\mathbb{R}^n)$,

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \|\mathcal{N}_{h,L}(f)\|_{L^p(\mathbb{R}^n)}. \tag{3.45}$$

To show (3.45), for $p\in[2,p_+(L))$, let $\psi\in L^{p'}(\mathbb{R}^n)$ satisfying that $\|\psi\|_{L^{p'}(\mathbb{R}^n)}\le 1$ and

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \left|\int_{\mathbb{R}^n} f(y)\overline{\psi(y)}\,dy\right|. \tag{3.46}$$

Here and hereafter, $1/p+1/p'=1$. We first claim that

$$\left|\int_{\mathbb{R}^n} f(y)\overline{\psi(y)}\,dy\right| = \lim_{t\to 0^+}\left|\int_{\mathbb{R}^n} e^{-t^{2m}L}(f)(y)\left[\frac{1}{t^n}\int_{B(y,t)}\overline{\psi(x)}\,dx\right]dy\right|. \tag{3.47}$$

Here and hereafter, "$t\to 0^+$" means that "$t>0$ and $t\to 0$". Indeed, if claim (3.47) holds true, then, from Fubini's theorem, Hölder's inequality and Remark 1.3 (ii), we deduce that

$$\begin{aligned}
\left|\int_{\mathbb{R}^n} f(y)\overline{\psi(y)}\,dy\right| &\le \sup_{t>0}\left\{\int_{\mathbb{R}^n}|\psi(x)|\left[\frac{1}{t^n}\int_{B(x,t)}|e^{-t^{2m}L}(f)(y)|\,dy\right]dx\right\}\\
&\lesssim \|\psi\|_{L^{p'}(\mathbb{R}^n)}\sup_{t>0}\left\{\int_{\mathbb{R}^n}\left[\frac{1}{t^n}\int_{B(x,t)}|e^{-t^{2m}L}(f)(y)|^2\,dy\right]^{\frac{p}{2}}dx\right\}^{\frac{1}{p}}\\
&\lesssim \left\{\int_{\mathbb{R}^n}\left(\sup_{t>0}\left[\frac{1}{t^n}\int_{B(x,t)}|e^{-t^{2m}L}(f)(y)|^2\,dy\right]^{\frac{1}{2}}\right)^p dx\right\}^{\frac{1}{p}}\\
&\sim \|\mathcal{R}_{h,L}(f)\|_{L^p(\mathbb{R}^n)} \sim \|\mathcal{N}_{h,L}(f)\|_{L^p(\mathbb{R}^n)},
\end{aligned}$$

which, together with (3.46), implies that, for all $p\in[2,p_+(L))$, (3.45) holds true.

Thus, to finish the proof of (3.45), it remains to show the claim (3.47). By Hölder's inequality and an elementary calculation, we see that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n} f(y)\overline{\psi(y)}\,dy - \int_{\mathbb{R}^n} e^{-t^{2m}L}(f)(y)\left[\frac{1}{t^n}\int_{B(y,t)}\overline{\psi(x)}\,dx\right]dy\right|\\
&\quad\le \left|\int_{\mathbb{R}^n}\{f(y)-e^{-t^{2m}L}(f)(y)\}\overline{\psi(y)}\,dy\right| + \left|\int_{\mathbb{R}^n} e^{-t^{2m}L}(f)(y)\left\{\overline{\psi(y)}-\frac{1}{t^n}\int_{B(y,t)}\overline{\psi(x)}\,dx\right\}dy\right|
\end{aligned}$$

$$\begin{aligned}&\leq \left\{ \int_{\mathbb{R}^n} |(e^{-t^{2m}L} - I)(f)(y)|^p \, dy \right\}^{\frac{1}{p}} \left\{ \int_{\mathbb{R}^n} |\psi(y)|^{p'} \, dy \right\}^{\frac{1}{p'}} \\ &\quad + \left\{ \int_{\mathbb{R}^n} |e^{-t^{2m}L}(f)(y)|^p \, dy \right\}^{\frac{1}{p}} \left\{ \int_{\mathbb{R}^n} \left| \psi(y) - \frac{1}{t^n} \int_{B(y,t)} \psi(x) \, dx \right|^{p'} dy \right\}^{\frac{1}{p'}} \\ &=: \mathrm{A}_t + \mathrm{B}_t.\end{aligned}$$

Notice that, by the fact that $\lim_{t\to 0^+} e^{-tz} = 1$ for all complex numbers z and the fact that L has a bounded functional calculus in $L^q(\mathbb{R}^n)$ with $q \in (p_-(L), p_+(L))$ (which is a simple corollary of Proposition 2.5 (iii) and [8, Theorem 1.2]), we know that $\{e^{-tL}\}_{t>0}$ has the strong continuity in $L^q(\mathbb{R}^n)$ for all $q \in (p_-(L), p_+(L))$. Letting $t \to 0^+$, and using the strong continuity of the semigroup $\{e^{-tL}\}_{t>0}$ in $L^p(\mathbb{R}^n)$ for $p \in [2, p_+(L))$ and $\|\psi\|_{L^{p'}(\mathbb{R}^n)} \leq 1$, we have that

$$\lim_{t\to 0^+} A_t = 0.$$

In what follows, for any locally integrable function f, let $\mathcal{M}(f)$ be the *Hardy–Littlewood maximal function* defined by setting, for all $x \in \mathbb{R}^n$,

$$\mathcal{M}(f)(x) := \sup_{B\ni x} \frac{1}{|B|} \int_B |f(x)| \, dx,$$

where the supremum is taken over all the balls B in $\mathbb{R}^n$ containing x. Observe that, for all $y \in \mathbb{R}^n$,

$$\left| \psi(y) - \frac{1}{t^n} \int_{B(y,t)} \psi(x) \, dx \right| \leq 2\mathcal{M}(\psi)(y)$$

and $\mathcal{M}(\psi) \in L^{p'}(\mathbb{R}^n)$. From this, the Lebesgue dominated convergence theorem and the Lebesgue differentiation theorem, together with the uniformly boundedness of the semigroup $\{e^{-tL}\}_{t>0}$ in $L^p(\mathbb{R}^n)$ for $p \in [2, p_+(L))$, we deduce that

$$\lim_{t\to 0^+} B_t = 0,$$

which completes the proof of claim (3.47). Thus, $H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n) \subset H^p_L(\mathbb{R}^n)$ for all $p \in (0, p_+(L))$.

To prove the inclusion $H^p_L(\mathbb{R}^n) \subset H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)$, we consider two cases. If $p \in (0, 1]$, by Theorem 2.11 and Remark 1.3, we see that it suffices to show that, for all $(p, 2, M, \epsilon)_L$-molecules α,

$$\|\mathcal{R}_{h,L}(\alpha)\|_{L^p(\mathbb{R}^n)} \lesssim 1,$$

where $\mathcal{R}_{h,L}$ is the radial heat maximal function defined as in (1.6). The latter estimate can be obtained by using the same method as that used in the proof of [28, Theorem 6.3], the details being omitted here. This finishes the proof of Theorem 1.4 for $p \in (0, 1]$.

If $p \in (1, p_+(L))$, let $\mathcal{R}_{h,L}$ be the radial maximal function defined as in (1.6) and, for any ball B and $j \in \mathbb{N}$, let $S_j(B) := 2^jB \setminus (2^{j-1}B)$ and $S_0(B) := B$. Then, for any $q \in (2, \infty)$, using Minkowski's inequality, Proposition 2.5 and the boundedness of $\mathcal{M}$ on $L^{q/2}(\mathbb{R}^n)$, we know that there exists a positive constant η such that

$$\begin{aligned}\|\mathcal{R}_{h,L}(f)\|_{L^q(\mathbb{R}^n)} &= \left\| \sup_{t\in(0,\infty)} \left\{ \frac{1}{t^n} \int_{B(\cdot,t)} \left| e^{-t^{2m}L}\left(\sum_{j\in\mathbb{Z}_+} \chi_{S_j(B(\cdot,t))} f \right)(y) \right|^2 dy \right\}^{\frac{1}{2}} \right\|_{L^q(\mathbb{R}^n)} \\ &\lesssim \sum_{j\in\mathbb{Z}_+} \left\| \sup_{t\in(0,\infty)} \left[\frac{1}{t^{\frac{n}{2}}} \exp\left\{ -\frac{[d(B(\cdot,t), S_j(B(\cdot,t)))]^{2m/(2m-1)}}{t^{2m/(2m-1)}} \right\} \|f\|_{L^2(S_j(B(\cdot,t)))} \right] \right\|_{L^q(\mathbb{R}^n)} \\ &\lesssim \sum_{j\in\mathbb{Z}_+} 2^{-j\eta} \left\| \sup_{t\in(0,\infty)} \left[\frac{1}{(2^jt)^n} \int_{2^jB(\cdot,t)} |f(x)|^2 \, dx \right]^{\frac{1}{2}} \right\|_{L^q(\mathbb{R}^n)} \\ &\lesssim \left\| [\mathcal{M}(|f|^2)]^{\frac{1}{2}} \right\|_{L^q(\mathbb{R}^n)} \lesssim \|f\|_{L^q(\mathbb{R}^n)},\end{aligned}$$

which implies that $\mathcal{R}_{h,L}$ is bounded on $L^q(\mathbb{R}^n)$. This, together with Remark 1.3, further shows that the non-tangential maximal function $\mathcal{N}_{h,L}$ is also bounded on $L^q(\mathbb{R}^n)$ for all $q \in (2, \infty)$. For the case $p \in (0, 1]$, from Lemma 2.12 and the complex interpolation of $H_L^p(\mathbb{R}^n)$ (see Proposition 2.8 together with [30, p. 52, Theorem]), we see that, for all $p \in (0, p_+(L))$, $\mathcal{N}_{h,L}$ is bounded from $H_L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$. This implies the inclusion $H_L^p(\mathbb{R}^n) \subset H^p_{\mathcal{N}_{h,L}}(\mathbb{R}^n)$ for all $p \in (0, p_+(L))$ and hence finishes the proof of Theorem 1.4. □

With the help of Theorem 1.4, we are able to prove Theorem 1.5.

Proof of Theorem 1.5. We first point out that, similar to Remark 1.3, we have

$$H^p_{\tilde{\mathcal{N}}_{h,L}}(\mathbb{R}^n) = H^p_{\tilde{\mathcal{R}}_{h,L}}(\mathbb{R}^n)$$

with equivalent quasi-norms. Moreover, by (1.6), (1.7), (1.14) and (1.15), we immediately see that, for all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\mathcal{N}_{h,L}(x) \leq \tilde{\mathcal{N}}_{h,L}(f)(x),$$

which, together with Theorem 1.4, implies the inclusion $H^p_{\tilde{\mathcal{N}}_{h,L}}(\mathbb{R}^n) \subset H_L^p(\mathbb{R}^n)$. The proof of the inclusion $H_L^p(\mathbb{R}^n) \subset H^p_{\tilde{\mathcal{R}}_{h,L}}(\mathbb{R}^n)$ is similar to the corresponding part of the proof of Theorem 1.4. Here, we only need to use the L^2 off-diagonal estimates of the gradient semigroup $\{\sqrt{t}\nabla^m e^{-t^{2m}L}\}_{t>0}$ to replace the L^2 off-diagonal estimates of the semigroup $\{e^{-t^{2m}L}\}_{t>0}$ therein, which completes the proof of Theorem 1.5. □

Now we turn to the proof of Proposition 1.6.

Proof of Proposition 1.6. For any $(x, t) \in \mathbb{R}^{n+1}_+$, following [1, pp. 59–60, Section 3.2 (c)], let $W_0^{m,2}(B(x, 2t))$ be the Sobolev space over $B(x, 2t)$, defined as the completion of $C_c^\infty(B(x, 2t))$ with respect to the *norm* $\|\cdot\|_{W^{m,2}(B(x,2t))}$, where, for all $\varphi \in C_c^\infty(B(x, 2t))$,

$$\|\varphi\|_{W^{m,2}(B(x,2t))} := \left[\sum_{0\leq|\alpha|\leq m} \|\partial^\alpha \varphi\|^2_{L^2(B(x,2t))}\right]^{\frac{1}{2}}.$$

Since $f \in L^2(\mathbb{R}^n)$, it follows that $e^{-t^{2m}L}(f) \in W^{m,2}(\mathbb{R}^n)$. Thus, we have

$$\psi_{x,t}e^{-t^{2m}L}(f) \in W_0^{m,2}(B(x, 2t)) \subset W_0^{m-1,2}(B(x, 2t)).$$

Recalling Poincaré's inequality from [35, Theorem 3.2.1], for all $k \in \{0, \dots, m-1\}$, $v \in W_0^{m-1,2}(B(x, 2t))$ and for all $(x, t) \in \mathbb{R}^{n+1}_+$,

$$\int_{B(x,2t)} |\nabla^k v(y)|^2\, dy \leq 2^{k-m+1}(2t)^{(m-1-k)2} \int_{B(x,2t)} |\nabla^{m-1} v(y)|^2\, dy.$$

Thus, by this and the properties of ψ, we see that, for all $(x, t) \in \mathbb{R}^{n+1}_+$,

$$\begin{aligned}\left\{\frac{1}{t^n}\int_{B(x,t)} |e^{-t^{2m}L}(f)(z)|^2\, dz\right\}^{\frac{1}{2}} &\leq \left\{\frac{1}{t^n}\int_{B(x,2t)} |\psi_{x,t}(z)e^{-t^{2m}L}(f)(z)|^2\, dz\right\}^{\frac{1}{2}}\\ &\lesssim \left\{\frac{1}{t^n}\int_{B(x,2t)} \left|(t\nabla)^{m-1}(\psi_{x,t}e^{-t^{2m}L}(f))(z)\right|^2 dz\right\}^{\frac{1}{2}},\end{aligned} \tag{3.48}$$

which, combined with (3.48), further shows that, for all $x \in \mathbb{R}^n$,

$$\mathcal{N}_{h,L}(f)(x) \lesssim \mathcal{N}^2_{h,\psi,L}(f)(x). \tag{3.49}$$

As for the converse direction, by Leibniz's rule and the properties of ψ, for all $(x, t) \in \mathbb{R}^{n+1}_+$, we have

$$\left\{\frac{1}{t^n}\int_{B(x,2t)} \left|(t\nabla)^{m-1}(\psi_{x,t}e^{-t^{2m}L}(f))(z)\right|^2 dz\right\}^{\frac{1}{2}} \lesssim \left\{\frac{1}{t^n}\int_{B(x,2t)} \sum_{k=0}^{m-1} \left|(t\nabla)^{k}(e^{-t^{2m}L}(f))(z)\right|^2 dz\right\}^{\frac{1}{2}},$$

which implies that, for all $x \in \mathbb{R}^n$,

$$\mathcal{N}^2_{h,\psi,L}(f)(x) \lesssim \tilde{\mathcal{N}}^2_{h,L}(f)(x). \tag{3.50}$$

Combining (3.49), (3.50), and Theorems 1.4 and 1.5, we conclude that $H_L^p(\mathbb{R}^n) = H^p_{\mathcal{N}_{h,\psi,L}}(\mathbb{R}^n)$ with equivalent quasi-norms, which completes the proof of Proposition 1.6. □

Now, we prove Theorem 1.8. To this end, we need the following proposition.

Proposition 3.10. *Let $k \in \mathbb{N}$, L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Let also $S_{L,k}$ and $S_{h,L,k}$ be the same, respectively, as in* (1.17) *and* (1.18). *Then, for all $q \in (0, p_+(L))$, there exists a positive constant $C_{(k,q)}$ such that the following hold:*

(i) *For all $f \in L^2(\mathbb{R}^n)$,*

$$\|S_{L,k}(f)\|_{L^q(\mathbb{R}^n)} \le C_{(k,q)}\|S_{L,1}(f)\|_{L^q(\mathbb{R}^n)}.$$

(ii) *For all $f \in L^2(\mathbb{R}^n)$,*

$$\|S_{h,L,k}(f)\|_{L^q(\mathbb{R}^n)} \le C_{(k,q)}\|S_{L,1}(f)\|_{L^q(\mathbb{R}^n)}.$$

Proof. We prove Proposition 3.10 by mathematical induction.

If $k = 1$, Proposition 3.10 (i) automatically holds true. To prove Proposition 3.10 (ii), by (3.27), we know that, for all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$[S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)]^2 \le [S_{L,1}^{\epsilon/4,3R,2}(f)(x)]^2. \tag{3.51}$$

By this, together with letting $\epsilon \to 0$ and $R \to \infty$, (3.19) and (3.20), we further conclude that, for all $q \in (0, \infty)$ and $f \in L^2(\mathbb{R}^n)$,

$$\|S_{h,L,1}(f)\|_{L^q(\mathbb{R}^n)} \lesssim \|S_{L,1}(f)\|_{L^q(\mathbb{R}^n)}, \tag{3.52}$$

which implies that Proposition 3.10 (ii) holds true in this case.

If $k = 2$, we prove (i) by first establishing a desired estimate for $\|S_{L,2}(f)\|_{L^q(\mathbb{R}^n)}$ with $q \in (0, \infty)$ (see (3.59) below). To this end, for all $0 < \epsilon \ll R < \infty$, $\lambda \in (0, \infty)$ and $x \in \mathbb{R}^n$, let $\Gamma^{\epsilon,R,\lambda}(x)$ be the truncated cone as in (3.18). Also, let $\eta \in C_c^\infty(\Gamma^{\epsilon/2,2R,3/2}(x))$ satisfy $\eta \equiv 1$ on $\Gamma^{\epsilon,R,1}(x)$, $0 \le \eta \le 1$ and, for all $l \in \mathbb{N}$ with $l \le m$ and $(y, t) \in \Gamma^{\epsilon/2,2R,3/2}(x)$,

$$|\nabla^l \eta(y, t)| \lesssim \frac{1}{t^l}.$$

From the properties of η, the definition of L, Leibniz's rule and Minkowski's inequality, we deduce that, for all $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}
S_{L,2}^{\epsilon,R,1}(f)(x) &= \left[\iint_{\Gamma^{\epsilon,R,1}(x)} |(t^{2m}L)^2 e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}}\right]^{\frac12} \\
&\le \left[\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} (t^{2m}L)^2 e^{-t^{2m}L}(f)(y)\overline{(t^{2m}L)^2 e^{-t^{2m}L}(f)(y)}\eta(y, t) \frac{dy\,dt}{t^{n+1}}\right]^{\frac12} \\
&= \left|\sum_{|\alpha|=m=|\beta|} \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y) t^m \partial^\beta(t^{2m}Le^{-t^{2m}L}(f))(y)\overline{t^m\partial^\alpha((t^{2m}L)^2 e^{-t^{2m}L}(f)\eta)(y, t)} \frac{dy\,dt}{t^{n+1}}\right|^{\frac12} \\
&\lesssim \sum_{l=0}^m \left|\sum_{|\alpha|=m=|\beta|} \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y) t^m \partial^\beta(t^{2m}Le^{-t^{2m}L}(f))(y)\right. \\
&\quad\times \left.\left[t^m \sum_{|\tilde\alpha|=l,\ \tilde\alpha\le\alpha} C_{(\alpha,\tilde\alpha)} \overline{\partial^{\tilde\alpha}((t^{2m}L)^2 e^{-t^{2m}L}(f))(y)\, \partial^{\alpha-\tilde\alpha}\eta(y, t)}\right] \frac{dy\,dt}{t^{n+1}}\right|^{\frac12} =: \sum_{l=0}^m \mathrm{I}_l,
\end{aligned} \tag{3.53}$$

where, for all multi-indices α and $\tilde\alpha$, $C_{(\alpha,\tilde\alpha)}$ denotes a positive constant depending on α and $\tilde\alpha$.

For I_0, by Hölder's inequality, the size condition of η and the ellipticity condition $(\mathcal{E}_0)$, we see that, for all $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}
\mathrm{I}_0 &\lesssim \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} \left|t^m \nabla^m(t^{2m}Le^{-t^{2m}L}(f))(y)\right|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14} \left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |(t^{2m}L)^2 e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}}\right\}^{\frac14} \\
&\sim [S_{h,L,1}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12} [S_{L,2}^{\epsilon/2,2R,3/2}(f)(x)]^{\frac12},
\end{aligned} \tag{3.54}$$

where $S_{h,L,1}^{\epsilon/2,2R,3/2}(f)$ and $S_{L,2}^{\epsilon/2,2R,3/2}(f)$ are defined, respectively, similar to $S_{h,L,1}(f)$ in (1.18) and $S_{L,2}(f)$ in (1.17), with $\Gamma(x)$ in (1.3) replaced by $\Gamma^{\epsilon/2,2R,3/2}(x)$ in (3.18).

For I_m, using an argument similar to that used in the proof of (3.28), we have that, for all $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$,

$$\mathrm{I}_m \lesssim [S^{\epsilon/2,2R,3/2}_{h,L,1}(f)(x)]^{\frac12}[S^{\epsilon/2,2R,3/2}_{h,L,2}(f)(x)]^{\frac12}. \tag{3.55}$$

Moreover, by an argument similar to that used in the proof of (3.27) (see also (3.51)), we find that, for all $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$,

$$[S^{\epsilon/2,2R,3/2}_{h,L,2}(f)(x)]^2 \lesssim [S^{\epsilon/4,3R,2}_{L,2}(f)(x)]^2. \tag{3.56}$$

Also, similar to (3.29), we know that, for all $l \in \{1, \dots, m-1\}$, $0 < \epsilon \ll R < \infty$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}\mathrm{I}_l &\lesssim [S^{\epsilon/2,2R,3/2}_{h,L,1}(f)(x)]^{\frac12}\left\{\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |t^l\nabla^l([t^{2m}L]^2e^{-t^{2m}L}(f))(y)|^2\,\frac{dy\,dt}{t^{n+1}}\right\}^{\frac14}\\ &\lesssim [S^{\epsilon/2,2R,3/2}_{h,L,1}(f)(x)]^{\frac12}[S^{\epsilon/2,2R,3/2}_{h,L,2}(f)(x)]^{\frac{l}{2m}}[S^{\epsilon/2,2R,3/2}_{L,2}(f)(x)]^{\frac{m-l}{2m}}.\end{aligned} \tag{3.57}$$

By combining (3.53)–(3.57) and then letting $\epsilon \to 0$ and $R \to \infty$, we conclude that, there exists a positive constant C such that, for all $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$S_{L,2}(f)(x) \le C[S^2_{h,L,1}(f)(x)]^{\frac12}[S^2_{L,2}(f)(x)]^{\frac12}. \tag{3.58}$$

Similarly, by following the same line of the proof of (3.58), we conclude that, for all $k \in \mathbb{Z}_+$, $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$S^{2^k}_{L,2}(f)(x) \le C[S^{2^{k+1}}_{h,L,1}(f)(x)]^{\frac12}[S^{2^{k+1}}_{L,2}(f)(x)]^{\frac12},$$

which, combined with the definition of $\mathcal{A}_k$ in (3.21), implies that, for all $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$\mathcal{A}_k(F)(x) \le C[\mathcal{A}_{k+1}(G)(x)]^{\frac12}[\mathcal{A}_{k+1}(F)(x)]^{\frac12},$$

where $F := (t^{2m}L)^2e^{-t^{2m}L}(f)$ and $G := (t\nabla)^m t^{2m}Le^{-t^{2m}L}(f)$. Moreover, since $f \in L^2(\mathbb{R}^n)$, F and G are in $T^2(\mathbb{R}^{n+1}_+)$. Thus, by Lemma 3.4, we conclude that, for all $q \in (0,\infty)$ and $f \in L^2(\mathbb{R}^n)$, $\|F\|_{T^q(\mathbb{R}^{n+1}_+)} \lesssim \|G\|_{T^q(\mathbb{R}^{n+1}_+)}$, which implies that, for all $q \in (0,\infty)$ and $f \in L^2(\mathbb{R}^n)$,

$$\|S_{L,2}(f)\|_{L^q(\mathbb{R}^n)} \lesssim \|S_{h,L,1}(f)\|_{L^q(\mathbb{R}^n)}. \tag{3.59}$$

This, combined with (3.52), shows that Proposition 3.10 (i) holds true for $k = 2$.

Moreover, Proposition 3.10 (ii) when $k = 2$ follows from (3.56), (3.19), (3.20) and Proposition 3.10 (i) when $k = 2$.

Now, let $\tilde{k} \in \mathbb{N} \cap [3,\infty)$. Assume that Proposition 3.10 holds true for all $k \in \{1,\dots,\tilde{k}\}$. Thus, by mathematical induction, to finish the proof of Proposition 3.10, it suffices to show that Proposition 3.10 also holds true for $\tilde{k}+1$.

Similar to (3.53), for all $0 < \epsilon \ll R < \infty$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, we have

$$\begin{aligned}S^{\epsilon,R,1}_{L,\tilde{k}+1}(f)(x) &= \left[\iint_{\Gamma^{\epsilon,R,1}(x)} |(t^{2m}L)^{\tilde{k}+1}e^{-t^{2m}L}(f)(y)|^2\,\frac{dy\,dt}{t^{n+1}}\right]^{\frac12}\\ &\le \left[\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} (t^{2m}L)^{\tilde{k}+1}e^{-t^{2m}L}(f)(y)\overline{(t^{2m}L)^{\tilde{k}+1}e^{-t^{2m}L}(f)(y)}\eta(y,t)\,\frac{dy\,dt}{t^{n+1}}\right]^{\frac12}\\ &= \left|\sum_{|\alpha|=m=|\beta|}\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y)t^m\partial^\beta([t^{2m}L]^{\tilde{k}}e^{-t^{2m}L}(f))(y)\overline{t^m\partial^\alpha((t^{2m}L)^{\tilde{k}+1}e^{-t^{2m}L}(f)\eta)(y,t)}\,\frac{dy\,dt}{t^{n+1}}\right|^{\frac12}\\ &\lesssim \sum_{l=0}^m\left|\sum_{|\alpha|=m=|\beta|}\iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} a_{\alpha,\beta}(y)t^m\partial^\beta([t^{2m}L]^{\tilde{k}}e^{-t^{2m}L}(f))(y)\right.\\ &\qquad\left.\times\left[t^m\sum_{|\tilde{\alpha}|=l,\,\tilde{\alpha}\le\alpha}\overline{C_{(\alpha,\tilde{\alpha})}\partial^{\tilde{\alpha}}((t^{2m}L)^{\tilde{k}+1}e^{-t^{2m}L}(f))(y)\partial^{\alpha-\tilde{\alpha}}\eta(y,t)}\right]\frac{dy\,dt}{t^{n+1}}\right|^{\frac12} =: \sum_{l=0}^m \tilde{\mathrm{I}}_l.\end{aligned} \tag{3.60}$$

For $\tilde{\mathrm{I}}_0$, by Hölder's inequality, the size condition of η and the ellipticity condition $(\mathcal{E}_0)$, we see that, for all $0 < \epsilon \ll R < \infty$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}\tilde{\mathrm{I}}_0 &\lesssim \left\{ \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} \left|t^m \nabla^m([t^{2m}L]^{\tilde{k}} e^{-t^{2m}L}(f))(y)\right|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{4}} \left\{ \iint_{\Gamma^{\epsilon/2,2R,3/2}(x)} |(t^{2m}L)^{\tilde{k}+1} e^{-t^{2m}L}(f)(y)|^2 \frac{dy\,dt}{t^{n+1}} \right\}^{\frac{1}{4}} \\ &\sim [S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}}(f)(x)]^{\frac{1}{2}} [S^{\epsilon/2,2R,3/2}_{L,\tilde{k}+1}(f)(x)]^{\frac{1}{2}},\end{aligned} \tag{3.61}$$

where $S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}}(f)$ and $S^{\epsilon/2,2R,3/2}_{L,\tilde{k}+1}(f)$ are defined, respectively, similar to $S_{h,L,\tilde{k}}(f)$ in (1.18) and $S_{L,\tilde{k}+1}(f)$ in (1.17), with $\Gamma(x)$ in (1.3) replaced by $\Gamma^{\epsilon/2,2R,3/2}(x)$ in (3.18).

For $\tilde{\mathrm{I}}_m$, using an argument similar to that used in the proof of (3.55), we have that, for all $0 < \epsilon \ll R < \infty$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\tilde{\mathrm{I}}_m \lesssim [S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}}(f)(x)]^{\frac{1}{2}} [S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}+1}(f)(x)]^{\frac{1}{2}}. \tag{3.62}$$

By an argument similar to that of (3.51), we know that, for all $0 < \epsilon \ll R < \infty$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$[S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}+1}(f)]^2 \lesssim [S^{\epsilon/4,3R,2}_{L,\tilde{k}+1}(f)(x)]^2.$$

By this, together with letting $\epsilon \to 0$ and $R \to \infty$, (3.19) and (3.20), we further conclude that, for all $q \in (0, \infty)$ and $f \in L^2(\mathbb{R}^n)$,

$$\|S_{h,L,\tilde{k}+1}(f)\|_{L^q(\mathbb{R}^n)} \lesssim \|S_{L,\tilde{k}+1}(f)\|_{L^q(\mathbb{R}^n)}. \tag{3.63}$$

Also, similar to (3.57), we know that, for all $l \in \{1, \dots, m-1\}$, $0 < \epsilon \ll R < \infty$, $f \in L^2(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\tilde{\mathrm{I}}_l \lesssim [S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}}(f)(x)]^{\frac{1}{2}} [S^{\epsilon/2,2R,3/2}_{h,L,\tilde{k}+1}(f)(x)]^{\frac{l}{2m}} [S^{\epsilon/2,2R,3/2}_{L,\tilde{k}+1}(f)(x)]^{\frac{m-l}{2m}}. \tag{3.64}$$

Thus, combining (3.60)–(3.64), and then letting $\epsilon \to 0$ and $R \to \infty$, we conclude that there exists a positive constant C such that, for all $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$S_{L,\tilde{k}}(f)(x) \le C[S^2_{h,L,\tilde{k}-1}(f)(x)]^{\frac{1}{2}} [S^2_{L,\tilde{k}}(f)(x)]^{\frac{1}{2}}. \tag{3.65}$$

Similarly, by following the same line of the proof of (3.65), we conclude that, for all $l \in \mathbb{Z}_+$, $f \in L^2(\mathbb{R}^n)$ and almost every $x \in \mathbb{R}^n$,

$$S^{2^l}_{L,\tilde{k}}(f)(x) \le C[S^{2^{l+1}}_{h,L,\tilde{k}-1}(f)(x)]^{\frac{1}{2}} [S^{2^{l+1}}_{L,\tilde{k}}(f)(x)]^{\frac{1}{2}}.$$

This, together with Lemma 3.4 and the assumption that Proposition 3.10 holds true for all $k \in \{1, \dots, \tilde{k}\}$, implies that Proposition 3.10 (i) also holds true in the case $\tilde{k}+1$, the details being omitted.

Finally, we see that Proposition 3.10 (ii) in the case $\tilde{k}+1$ follows from (3.63) and Proposition 3.10 (i) in the case $\tilde{k}+1$, which completes the proof of Proposition 3.10. □

From the proof of Proposition 3.10, we immediately deduce the following conclusions.

Corollary 3.11. *Let $k \in \mathbb{N}$, L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Let $S_{L,k}$ and $S_{h,L,k}$ be the same, respectively, as in* (1.17) *and* (1.18). *Then, for all $p \in (0, p_+(L))$, there exists a positive constant $C_{(p,k)}$, depending on p and k, such that, for all $f \in L^2(\mathbb{R}^n)$,*

$$\frac{1}{C_{(p,k)}}\|S_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)} \le \|S_{L,k}(f)\|_{L^p(\mathbb{R}^n)} \le C_{(p,k)}\|S_{h,L,k-1}(f)\|_{L^p(\mathbb{R}^n)}. \tag{3.66}$$

Proof. The first inequality of (3.66) follows immediately from (3.63) with $\tilde{k}+1$ replaced by k in the proof of Proposition 3.10, while the second inequality of (3.66) is proved in the proof of Proposition 3.10, which completes the proof of Corollary 3.11. □

Combining Corollary 3.11 and Lemma 3.5 (i), we immediately obtain the following corollary, which improves Lemma 3.5 (ii) by extending the range of q from $(q_-(L), q_+(L))$ to $(p_-(L), p_+(L))$.

Corollary 3.12. *Let L be as in* (1.1) *and satisfying the ellipticity condition* $(\mathcal{E}_0)$. *Let* $k \in \mathbb{N}$, $q \in (p_-(L), p_+(L))$ *and* $S_{h,L,k}$ *be as in* (1.18). *Then, there exists a positive constant C such that, for all* $f \in L^2(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$,

$$\|S_{h,L,k}(f)\|_{L^q(\mathbb{R}^n)} \le C\|f\|_{L^q(\mathbb{R}^n)}.$$

We are now in a position to prove Theorem 1.8.

Proof of Theorem 1.8. We first prove Theorem 1.8 (i). If $p \in (0, 2]$, by Proposition 2.15, we know that

$$H^p_{S_{L,k}}(\mathbb{R}^n) = H^p_L(\mathbb{R}^n).$$

Thus, to finish the proof of Theorem 1.8 (i), it suffices to consider the case $p \in (2, p_+(L))$.

Moreover, by Lemma 3.5 (i) and a density argument, we see that, for all $p \in (2, p_+(L))$,

$$H^p_L(\mathbb{R}^n) \subset H^p_{S_{L,k}}(\mathbb{R}^n).$$

On the other hand, for all $p \in (2, p_+(L))$, by an argument similar to that used in the proof of (3.32), we know that, for all $k \in \mathbb{N}$ and $f \in L^2(\mathbb{R}^n)$,

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \|S_{L,k}(f)\|_{L^p(\mathbb{R}^n)},$$

which, combined with Lemma 2.12 and a density argument, implies that, for all $p \in (2, p_+(L))$,

$$H^p_{S_{L,k}}(\mathbb{R}^n) \subset H^p_L(\mathbb{R}^n).$$

This shows that Theorem 1.8 (i) holds true.

We now prove Theorem 1.8 (ii). To show the inclusion that

$$H^p_{S_{h,L,k}}(\mathbb{R}^n) \subset H^p_L(\mathbb{R}^n),$$

for all $p \in (0, p_+(L))$, by Theorem 1.8 (i) and Corollary 3.11, we conclude that, for all $f \in L^2(\mathbb{R}^n)$,

$$\|f\|_{H^p_L(\mathbb{R}^n)} \sim \|S_{L,k+1}(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|S_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)},$$

which, together with a density argument, implies that $H^p_{S_{h,L,k}}(\mathbb{R}^n) \subset H^p_L(\mathbb{R}^n)$.

For the converse inclusion, if $p \in (p_-(L), p_+(L))$, by Propositions 3.10 (ii) and 3.6, we see that, for all $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$,

$$\|S_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|S_L(f)\|_{L^p(\mathbb{R}^n)} \sim \|f\|_{L^p(\mathbb{R}^n)},$$

which, combined with Lemma 2.12 and a density argument, implies that $H^p_L(\mathbb{R}^n) \subset H^p_{S_{h,L,k}}(\mathbb{R}^n)$ holds true in the range $p \in (p_-(L), p_+(L))$.

If $p \in (0, 1]$, by considering the action of $S_{h,L,k}$ on each $(p, 2, M, \epsilon)_L$-molecule (see, for example, [15, equation (4.4)] for a proof of a similar conclusion) and Theorem 2.11, we know that, for all $p \in (0, 1]$ and $f \in H^p_L(\mathbb{R}^n)$,

$$\|S_{h,L,k}(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)},$$

which, together with the conclusion in the case $p \in (p_-(L), p_+(L))$, Lemma 2.12, the interpolation (see Proposition 2.8 together with [30, p. 52, Theorem]) and a density argument, implies that $H^p_L(\mathbb{R}^n) \subset H^p_{S_{h,L,k}}(\mathbb{R}^n)$ holds true for all $p \in (0, q_+(L))$. This finishes the proof of Theorem 1.8. □

Acknowledgment: The authors would like to thank the referee for his very carefully reading and many stimulating remarks which do improve the presentation of this article.

Funding: J. Cao is supported by the National Natural Science Foundation of China (grant no. 11501506) and the Natural Science Foundation of Zhejiang University of Technology (grant no. 2014XZ011). S. Mayboroda was partially supported by the NSF grants DMS 1220089 (CAREER), DMS 1344235 (INSPIRE), DMR 0212302 (UMN MRSEC Seed grant) and the Alfred P. Sloan Fellowship. D. Yang is supported by the National Natural Science Foundation of China (grant no. 11571039 and 11361020), the Specialized Research Fund for the Doctoral Program of Higher Education of China (grant no. 20120003110003) and the Fundamental Research Funds for the Central Universities of China (grant no. 2013YB60 and 2014KJJCA10).

References

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Pure Appl. Math. (Amst.) 140, Academic Press, Amsterdam, 2003.

[2] P. Auscher, On necessary and sufficient conditions for L^p-estimates of Riesz transforms associated to elliptic operators on $\mathbb{R}^n$ and related estimates, *Mem. Amer. Math. Soc.* **186** (2007), no. 871, DOI 10.1090/memo/0871.

[3] P. Auscher, Change of angle in tent spaces, *C. R. Math. Acad. Sci. Paris* **349** (2011), 297–301.

[4] P. Auscher, S. Hofmann, A. McIntosh and P. Tchamitchian, The Kato square root problem for higher order elliptic operators and systems on $\mathbb{R}^n$, *J. Evol. Equ.* **1** (2001), 361–385.

[5] P. Auscher and M. Qafsaoui, Equivalence between regularity theorems and heat kernel estimates for higher order elliptic operators and systems under divergence form, *J. Funct. Anal.* **177** (2000), 310–364.

[6] P. Auscher and P. Tchamitchian, *Square Root Problem for Divergence Operators and Related Topics*, Société Mathématique de France, Paris, 1988.

[7] A. Barton, Gradient estimates and the fundamental solution for higher-order elliptic systems with rough coefficients, preprint (2014), http://arxiv.org/abs/1409.7600.

[8] S. Blunck and P. C. Kunstmann, Calderón–Zygmund theory for non-integral operators and the H^∞ functional calculus, *Rev. Mat. Iberoam.* **19** (2003), 919–942.

[9] S. Blunck and P. C. Kunstmann, Weak type (p, p) estimates for Riesz transforms, *Math. Z.* **247** (2004), 137–148.

[10] T. A. Bui, J. Cao, L. D. Ky, D. Yang and S. Yang, Musielak–Orlicz–Hardy spaces associated with operators satisfying reinforced off-diagonal estimates, *Anal. Geom. Metr. Spaces* **1** (2013), 69–129.

[11] D. L. Burkholder, R. F. Gundy and M. L. Silverstein, A maximal function characterization of the class H^p, *Trans. Amer. Math. Soc.* **157** (1971), 137–153.

[12] J. Cao, D.-C. Chang, D. Yang and S. Yang, Boundedness of generalized Riesz transforms on Orlicz–Hardy spaces associated to operators, *Integral Equations Operator Theory* **76** (2013), 225–283.

[13] J. Cao, D.-C. Chang, D. Yang and S. Yang, Weighted local Orlicz–Hardy spaces on domains and their applications in inhomogeneous Dirichlet and Neumann problems, *Trans. Amer. Math. Soc.* **365** (2013), 4729–4809.

[14] J. Cao, S. Mayboroda and D. Yang, Local Hardy spaces associated with inhomogeneous higher order elliptic operators, *Anal. Appl. (Singap.)* **2015** (2015), DOI 10.1142/S0219530515500189.

[15] J. Cao and D. Yang, Hardy spaces $H_L^p(\mathbb{R}^n)$ associated to operators satisfying k-Davies–Gaffney estimates, *Sci. China Math.* **55** (2012), 1403–1440.

[16] D.-C. Chang, G. Dafni and E. M. Stein, Hardy spaces, BMO and boundary value problems for the Laplacian on a smooth domain in $\mathbb{R}^n$, *Trans. Amer. Math. Soc.* **351** (1999), 1605–1661.

[17] D.-C. Chang, S. G. Krantz and E. M. Stein, Hardy spaces and elliptic boundary value problems, in: *The Madison Symposium on Complex Analysis* (Madison 1991), Contemp. Math. 137, American Mathematical Society, Providence (1992), 119–131.

[18] D.-C. Chang, S. G. Krantz and E. M. Stein, H^p theory on a smooth domain in $\mathbb{R}^N$ and elliptic boundary value problems, *J. Funct. Anal.* **114** (1993), 286–347.

[19] R. R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes, Compensated compactness and Hardy spaces, *J. Math. Pures Appl. (9)* **72** (1993), 247–286.

[20] R. R. Coifman, Y. Meyer and E. M. Stein, Some new function spaces and their applications to harmonic analysis, *J. Funct. Anal.* **62** (1985), 304–335.

[21] Q. Deng, Y. Ding and X. Yao, Characterizations of Hardy spaces associated to higher order elliptic operators, *J. Funct. Anal.* **263** (2012), 604–674.

[22] X. T. Duong, S. Hofmann, D. Mitrea, M. Mitrea and L. Yan, Hardy spaces and regularity for the inhomogeneous Dirichlet and Neumann problems, *Rev. Mat. Iberoam.* **29** (2013), 183–236.

[23] X. T. Duong and J. Li, Hardy spaces associated to operators satisfying Davies–Gaffney estimates and bounded holomorphic functional calculus, *J. Funct. Anal.* **264** (2013), 1409–1437.

[24] X. T. Duong, J. Xiao and L. Yan, Old and new Morrey spaces with heat kernel bounds, *J. Fourier Anal. Appl.* **13** (2007), 87–111.

[25] X. T. Duong and L. Yan, New function spaces of BMO type, the John–Nirenberg inequality, interpolation, and applications, *Comm. Pure Appl. Math.* **58** (2005), 1375–1420.

[26] C. Fefferman and E. M. Stein, H^p spaces of several variables, *Acta Math.* **129** (1972), 137–193.

[27] M. Haase, *The Functional Calculus for Sectorial Operators*, Oper. Theory Adv. Appl. 169, Birkhäuser, Basel, 2006.

[28] S. Hofmann and S. Mayboroda, Hardy and BMO spaces associated to divergence form elliptic operators, *Math. Ann.* **344** (2009), 37–116.

[29] S. Hofmann, S. Mayboroda and A. McIntosh, Second order elliptic operators with complex bounded measurable coefficients in L^p, Sobolev and Hardy spaces, *Ann. Sci. Éc. Norm. Supér. (4)* **44** (2011), 723–800.

[30] S. Janson, On the interpolation of sublinear operators, *Studia Math.* **75** (1982), 51–53.

[31] R. Jiang and D. Yang, New Orlicz–Hardy spaces associated with divergence form elliptic operators, *J. Funct. Anal.* **258** (2010), 1167–1224.

[32] R. Jiang and D. Yang, Orlicz–Hardy spaces associated with operators satisfying Davies–Gaffney estimates, *Commun. Contemp. Math.* **13** (2011), 331–373.
[33] N. Kalton, S. Mayboroda and M. Mitrea, Interpolation of Hardy–Sobolev–Besov–Triebel–Lizorkin spaces and applications to problems in partial differential equations, in: *Interpolation Theory and Applications*, Contemp. Math. 445, American Mathematical Society, Providence (2007), 121–177.
[34] S. Mayboroda and M. Mitrea, Sharp estimates for Green potentials on non-smooth domains, *Math. Res. Lett.* **11** (2004), 481–492.
[35] C. B. Morrey, Jr., *Multiple Integrals in the Calculus of Variations*, Classics Math., Springer, Berlin, 2008.
[36] E. M. Ouhabaz, *Analysis the of Heat Equations on Domains*, Princeton University Press, Princeton, 2005.