SCIENCE CHINA
Mathematics

• ARTICLES •

December 2010 Vol. 53 No. 12: 3163–3178
doi: 10.1007/s11425-010-4108-2

Anisotropic singular integrals in product spaces

LI BaoDe[1], BOWNIK Marcin[2], YANG DaChun[3,*] & ZHOU Yuan[3]

[1]*School of Mathematics and System Sciences, Xinjiang University, Urumqi* 830046*, China;*
[2]*Department of Mathematics, University of Oregon, Eugene, OR* 97403-1222*, USA;*
[3]*School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing* 100875*, China*
*Email: baodeli*1981*@sina.com, mbownik@uoregon.edu, dcyang@bnu.edu.cn, yuanzhou@mail.bnu.edu.cn*

Received January 28, 2010; accepted May 11, 2010

Abstract In this paper, the authors introduce a class of product anisotropic singular integral operators, whose kernels are adapted to the action of a pair $\vec{A} := (A_1, A_2)$ of expansive dilations on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. This class is a generalization of product singular integrals with convolution kernels introduced in the isotropic setting by Fefferman and Stein. The authors establish the boundedness of these operators in weighted Lebesgue and Hardy spaces with weights in product A_∞ Muckenhoupt weights on $\mathbb{R}^n \times \mathbb{R}^m$. These results are new even in the unweighted setting for product anisotropic Hardy spaces.

Keywords expansive dilation, Muckenhoupt weight, product space, Hardy space, bump function, singular integral

MSC(2000): 42B30, 42B20, 42B25, 42B35

Citation: Li B D, Bownik M, Yang D C, et al. Anisotropic singular integrals in product spaces. Sci China Math, 2010, 53(12): 3163–3178, doi: 10.1007/s11425-010-4108-2

1 Introduction

The theory of Hardy spaces and singular integrals plays an important role in harmonic analysis and partial differential equations; see, for example, [13, 17, 18, 30]. There have been several directions of extending Hardy and other function space theory from Euclidean spaces to other domains and non-isotropic settings; see, for example, [1, 5–7, 10, 15, 24, 29, 31–33]. A significant effort has been devoted to developing a theory of Hardy spaces and singular integrals on product domains. This direction was initiated by Gundy and Stein [19] with Fefferman, Nagel and Stein among its main contributors [11, 14, 15, 26]. In particular, Fefferman and Stein [14] introduced a class of product singular integrals with convolution kernels and established their boundedness in Lebesgue spaces. Fefferman further proved the boundedness of some singular integrals from product Hardy spaces to Lebesgue spaces in [11] and also established some weighted boundedness in [12].

The goal of this paper is to extend some of the existing isotropic product Hardy space theory to the non-isotropic setting associated with expansive dilations. Let A_1 and A_2 be expansive dilations, on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let w be a product A_∞ Muckenhoupt weight associated with a pair of dilations, $\vec{A} := (A_1, A_2)$. Recently, the authors [4] developed the theory of weighted anisotropic product Hardy spaces $H^p_w(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$ with $p \in (0, 1]$. In this paper, motivated by Bownik [1] and Nagel-Stein [26], we introduce a class of anisotropic singular integrals on $\mathbb{R}^n \times \mathbb{R}^m$, whose kernels are adapted to $\vec{A}$ in

*Corresponding author

the sense of Bownik and have vanishing moments, defined via bump functions in the sense of Stein. Then we establish the boundedness of these anisotropic singular integrals on weighted Lebesgue spaces $L^q_w(\mathbb{R}^n\times\mathbb{R}^m)$ with $q\in(1,\infty)$ and weighted Hardy spaces $H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ with $p\in(0,1]$. These results are new even in the unweighted setting $w=1$.

We point out that the vanishing moments of singular integrals defined via bump functions were originally introduced by Stein [30]. To obtain the estimates for solutions of the Kohn-Laplacian on some classes of model domains in $\mathbb{C}^N$, Nagel and Stein [26, 27] introduced a class of singular integrals including their product versions, whose vanishing moments are defined via bump functions. Such a theory of product singular integrals is also used in the analysis on Heisenberg-type groups; see [25].

To state our main results, we carefully define the class of product anisotropic singular integral operators adapted to the action of a pair $\vec{A}$ of expansive dilations.

Definition 1.1. *A real $n\times n$ matrix A is an expansive dilation, shortly a dilation, if all its eigenvalues λ satisfy $|\lambda|>1$. Throughout the whole paper, for convenience, we sometimes use $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$ to denote, $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. For expansive dilation A_i on $\mathbb{R}^{n_i}$, $i=1,2$, we always let $b_i:=|\det(A_i)|$ and $\vec{A}:=(A_1,A_2)$. We also let $B^{(i)}_k$, $k\in\mathbb{Z}$, be dilated balls and ρ_i the step homogeneous-norm associated with A_i as in Definition 2.1.*

Definition 1.2. *Let $N\in\mathbb{N}$. A function ψ on $\mathbb{R}^n$ is called an N-normalized bump function associated with the ball B_0, if $\operatorname{supp}\psi\subset B_0$, and $\|\partial^\alpha\psi\|_{L^\infty(\mathbb{R}^n)}\leqslant 1$ for all $\alpha\in\mathbb{Z}^n_+$ with $|\alpha|\leqslant N$. A function ψ on $\mathbb{R}^n$ is called an N-normalized bump function associated with the ball B_k with $k\in\mathbb{Z}$ if and only if $\psi(A^k\cdot)$ is an N-normalized bump function associated with the ball B_0.*

Let $\mathcal{D}(\mathbb{R}^n\times\mathbb{R}^m)$ be the space of all infinitely differentiable functions with compact supports endowed with the inductive limit topology and $\mathcal{D}'(\mathbb{R}^n\times\mathbb{R}^m)$ its topological dual space. Also, let $\Omega_{n\times m}:=(\mathbb{R}^n\times\mathbb{R}^m)\setminus\{(x_1,x_2):\ x_1=0 \text{ or } x_2=0\}$, $\mathbb{N}:=\{1,2,\ldots\}$ and $\mathbb{Z}_+:=\mathbb{N}\cup\{0\}$.

Definition 1.3. *Let $s_1,s_2\in\mathbb{Z}_+$. Let $T:\mathcal{D}(\mathbb{R}^n\times\mathbb{R}^m)\to\mathcal{D}'(\mathbb{R}^n\times\mathbb{R}^m)$ be a continuous linear mapping. Then, T is called a product anisotropic singular integral operator (PASIO) of order (s_1,s_2), if the following conditions are met:*

(K0) *T has a distribution kernel K, which is a continuous function on $\Omega_{n\times m}$, such that for all $\varphi=\varphi^{(1)}\otimes\varphi^{(2)}\in\mathcal{D}(\mathbb{R}^n\times\mathbb{R}^m)$ and $x_1\notin\operatorname{supp}\varphi^{(1)}$, $x_2\notin\operatorname{supp}\varphi^{(2)}$,*

$$T(\varphi)(x_1,x_2)=\int_{\mathbb{R}^n\times\mathbb{R}^m}K(x_1-y_1,\,x_2-y_2)\varphi^{(1)}(y_1)\varphi^{(2)}(y_2)\,dy_1\,dy_2;$$

(K1) *there exists a positive constant C_1 such that for all $(x_1,x_2)\in\Omega_{n\times m}$ with $\rho_i(x_i)=b_i^{\ell_i}$, and for all $\alpha_i\in\mathbb{Z}^{n_i}_+$ with $|\alpha_i|\leqslant s_i$, $i=1,2$,*

$$|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K(A_1^{\ell_1}\cdot,\,A_2^{\ell_2}\cdot)](A_1^{-\ell_1}x_1,\,A_2^{-\ell_2}x_2)|\leqslant C_1[\rho_1(x_1)]^{-1}[\rho_2(x_2)]^{-1};$$

(K2) *there exist $N_1,N_2\in\mathbb{N}$ such that for each N_1-normalized bump function $\psi^{(1)}$ associated with $B_0^{(1)}$ and N_2-normalized bump function $\psi^{(2)}$ associated with $B_0^{(2)}$, and all $k_1,k_2\in\mathbb{Z}$,*

$$|\langle K,\,\psi^{(1)}(A_1^{k_1}\cdot)\otimes\psi^{(2)}(A_2^{k_2}\cdot)\rangle|\leqslant C_1;$$

(K3) *for each N_2-normalized bump function $\psi^{(2)}$ associated with $B_0^{(2)}$ and $k_2\in\mathbb{Z}$, there exists a continuous linear operator $T^{\psi^{(2)},k_2}:\mathcal{D}(\mathbb{R}^n)\to\mathcal{D}'(\mathbb{R}^n)$ with a distribution kernel $K^{\psi^{(2)},k_2}$, which is a continuous function on $\mathbb{R}^n\setminus\{0\}$, such that for all $\varphi^{(1)}\in\mathcal{D}(\mathbb{R}^n)$ and $x_1\notin\operatorname{supp}\varphi^{(1)}$,*

$$T^{\psi^{(2)},k_2}(\varphi^{(1)})=T(\varphi^{(1)}\otimes[\psi^{(2)}(A_2^{k_2}\cdot)])=\int_{\mathbb{R}^n}K^{\psi^{(2)},k_2}(x_1-y_1)\varphi^{(1)}(y_1)dy_1.$$

Furthermore, for all $x_1\neq 0$ with $\rho_1(x_1)=b_1^{\ell_1}$ and for all $\alpha_1\in\mathbb{Z}^n_+$ with $|\alpha_1|=s_1$,

$$|\partial_1^{\alpha_1}[K^{\psi^{(2)},k_2}(A_1^{\ell_1}\cdot)](A_1^{-\ell_1}x_1)|\leqslant C_1[\rho_1(x_1)]^{-1}.$$

(K3) *also holds with the roles of x_1 and x_2 interchanged.*

In the case when less regularity is desired, one can weaken conditions (K1) and (K3) on the derivatives to more familiar conditions on differences, as in the work of Han and Yang [22] (see also [23]).

Definition 1.4. *In what follows, let σ_i for $i=1,2$ be as in* (2.1) *associated with A_i. We say that T is a product anisotropic singular integral operator of order* 0*, if it satisfies Definition* 1.3 *with $s_1=s_2=0$. Moreover, there exist $\epsilon_1,\epsilon_2>0$ such that for all $(x_1,x_2)\in\Omega_{n\times m}$ with $\rho_i(x_i)=b_i^{\ell_i}$ and $h_i\in\mathbb{R}^{n_i}$ with $\rho_i(h_i)\leqslant b_i^{-2\sigma_i}\rho_i(x_i)$, we have*

$$\begin{aligned}|\Delta_{h_1}^{(1)}K(x_1,x_2)|&\leqslant C_1\frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}}\frac{1}{\rho_2(x_2)},\\|\Delta_{h_1}^{(1)}\Delta_{h_2}^{(2)}K(x_1,x_2)|&\leqslant C_1\frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}}\frac{[\rho_2(h_2)]^{\epsilon_2}}{[\rho_2(x_2)]^{1+\epsilon_2}},\\|\Delta_{h_1}^{(1)}K^{\psi^{(2)},k_2}(x_1)|&\leqslant C_1\frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}}.\end{aligned}$$

Here, we used difference operators $\Delta_{h_1}^{(1)}K(x_1,x_2):=K(x_1+h_1,x_2)-K(x_1,x_2)$ and $\Delta_{h_2}^{(2)}K(x_1,x_2):=K(x_1,x_2+h_2)-K(x_1,x_2)$. The above estimates must also hold with the roles of x_1 and x_2 interchanged.

Finally, we are ready to formulate the two main results of this paper. Theorem 1.5 is a generalization of a result of Fefferman and Stein [14] from the classical isotropic setting to the non-isotropic setting. Likewise, Theorem 1.6 is a generalization of a result of Han and Yang [22] to the setting of weighted anisotropic product Hardy spaces.

Theorem 1.5. *Let $w\in\mathcal{A}_p(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ with $p\in(1,\infty)$. Then, a PASIO T of order* 0 *uniquely extends to a bounded operator on $L_w^p(\mathbb{R}^n\times\mathbb{R}^m)$.*

Theorem 1.6. *Let $w\in\mathcal{A}_\infty(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ and q_w be its critical index as in* (2.4)*. Let $s_1,s_2\in\mathbb{Z}_+$ and $p\in(0,1]$. If*

$$s_i>(q_w/p-1)\log_{|\lambda_{i,1}|}b_i\qquad\text{for } i=1,2,\tag{1.1}$$

where $\lambda_{i,1}$ is the smallest eigenvalue of A_i in absolute value, then a PASIO T of order (s_1+1,s_2+1) extends uniquely to a bounded operator on $H_w^p(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$. Moreover, T admits another unique bounded extension to an operator $H_w^p(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})\to L_w^p(\mathbb{R}^n\times\mathbb{R}^m)$.

Remark 1.7. Consider the classical case corresponding to the choice of dyadic dilations $A_1=2I_{n_1}$, $A_2=2I_{n_2}$ and weight $w=1$. Then, $q_w=1$, $\rho_i(x)=|x|^{n_i}$, and $\log_{|\lambda_{i,1}|}b_i=n_i$ for $i=1,2$. In this case, if $p\in(1,\infty)$ and $\epsilon_i\in(0,1/n_i]$, the boundedness on $L^p(\mathbb{R}^n\times\mathbb{R}^m)$ of product singular integrals, as in Definition 1.4, follows from results of Nagel and Stein [26]. On the other hand, if $\max\{n_1/(n_1+\epsilon_1),\,n_2/(n_2+\epsilon_2)\}<p\leqslant1$, then the boundedness in $H^p(\mathbb{R}^n\times\mathbb{R}^m)$ of such product singular integrals was established by Han and Yang [22, Theorem 2].

This paper is organized as follows. In Section 2, we recall some notation and known notions. The proofs of Theorems 1.5 and 1.6 are presented in Sections 3 and 4, respectively. The methods used in these proofs borrow some ideas from [22] and [26]; see also [23] and [20]. However, unlike [22], [23] and [20], the discrete Calderón reproduction formula with kernel having compact support and the g-function characterization of the product anisotropic Hardy spaces are not available. Instead, we use the Lusin-area characterization with the kernels having no compact support. To overcome these additional difficulties, we invoke a decomposition of kernels technique used by Nagel and Stein; see [26, Lemma 3.5.1] and Lemma 3.1 below. Moreover, to prove Theorem 1.6, we use a variant of a key boundedness criterion established in [4, Corollary 6.1], which reduces the boundedness of the considered singular integrals to their behaviors on rectangular atoms; see Lemma 4.3 below and also [8, Corollary 1.1] for the corresponding result on $H^p(\mathbb{R}^n\times\mathbb{R}^m)$.

Finally we make some conventions. Throughout this paper, 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. Constants with subscripts do not change through the whole paper. We use the symbol $f\lesssim g$ to denote $f\leqslant Cg$, and if $f\lesssim g\lesssim f$, we write $f\sim g$. For all $x\in\mathbb{R}$, we denote by $\lfloor x\rfloor$ the *maximal integer no greater than x*.

2 Preliminaries

In this section, we recall basic facts about product Hardy spaces associated with expansive dilations.

By [1, Lemma 2.2], for a given expansive dilation A, there exist an open ellipsoid Δ and $r \in (1, \infty)$ such that $\Delta \subset r\Delta \subset A\Delta$. Moreover, $|\Delta| = 1$, where $|\Delta|$ denotes the n-dimensional Lebesgue measure of the set Δ. Throughout the whole paper, we

$$\text{set } B_k := A^k\Delta \text{ for } k \in \mathbb{Z} \quad \text{and} \quad \text{let } \sigma \text{ be the } \textit{minimum integer} \text{ such that } 2B_0 \subset A^\sigma B_0. \tag{2.1}$$

Then B_k is open, $B_k \subset rB_k \subset B_{k+1}$ and $|B_k| = b^k$. Obviously, $\sigma \geqslant 1$. For any subset E of $\mathbb{R}^n$, let $E^{\complement} := \mathbb{R}^n \setminus E$. Then it is easy to prove (see [1, p. 8]) that for all $k, \ell \in \mathbb{Z}$,

$$B_k + B_\ell \subset B_{\max\{k,\,\ell\}+\sigma}, \tag{2.2}$$

$$B_k + (B_{k+\sigma})^{\complement} \subset (B_k)^{\complement}, \tag{2.3}$$

where $E + F$ denotes the algebraic sums $\{x + y : x \in E, y \in F\}$ of sets $E, F \subset \mathbb{R}^n$.

Recall that the homogeneous quasi-norm associated with A was introduced in [1, Definition 2.3] as follows. For a fixed dilation A, we let $b := |\det A|$.

Definition 2.1. *A homogeneous quasi-norm associated with an expansive dilation A is a measurable mapping $\rho : \mathbb{R}^n \to [0, \infty)$ such that*

(i) $\rho(x) = 0$ *if and only if* $x = 0$;

(ii) $\rho(Ax) = b\rho(x)$ *for all* $x \in \mathbb{R}^n$;

(iii) $\rho(x + y) \leqslant H[\rho(x) + \rho(y)]$ *for all* $x, y \in \mathbb{R}^n$, *where H is a constant no less than* 1.

Define the step homogeneous quasi-norm ρ associated with A and Δ by setting, for all $x \in \mathbb{R}^n$, $\rho(x) = b^k$ if $x \in B_{k+1} \setminus B_k$, or else 0 if $x = 0$.

It has been proved that all homogeneous quasi-norms associated with a given dilation A are equivalent (see [1, Lemma 2.4]). Therefore, in what follows, for a given expansive dilation A, for convenience, we always use the step homogeneous quasi-norm ρ. Moreover, from (2.2) and (2.3), it follows that for all $x, y \in \mathbb{R}^n$,

$$\rho(x + y) \leqslant b^\sigma \max\{\rho(x), \rho(y)\} \leqslant b^\sigma[\rho(x) + \rho(y)].$$

The class of Muckenhoupt weights associated with A was introduced in [2]. For more details about weights, see [3, 16–18, 31].

Definition 2.2. *Let $p \in [1, \infty)$, A be a dilation and w a nonnegative measurable function on $\mathbb{R}^n$. The function w is said to belong to the weight class of Muckenhoupt $\mathcal{A}_p(\mathbb{R}^n; A)$, if there exists a positive constant C such that when $p > 1$,*

$$\sup_{x\in\mathbb{R}^n,\, k\in\mathbb{Z}} \left\{\frac{1}{|B_k|}\int_{x+B_k} w(y)\,dy\right\}\left\{\frac{1}{|B_k|}\int_{x+B_k} [w(y)]^{-1/(p-1)}\,dy\right\}^{p-1} \leqslant C,$$

$$\sup_{x\in\mathbb{R}^n,\, k\in\mathbb{Z}} \left\{\frac{1}{|B_k|}\int_{x+B_k} w(y)\,dy\right\}\left\{\operatorname*{ess\,sup}_{y\in x+B_k} [w(y)]^{-1}\right\} \leqslant C, \qquad \textit{when } p = 1.$$

Moreover, the minimal constant C as above is denoted by $C_A(w)$.

Define $\mathcal{A}_\infty(\mathbb{R}^n; A) := \bigcup_{1\leqslant p<\infty} \mathcal{A}_p(\mathbb{R}^n; A)$.

Product Muckenhoupt weights were first studied by Fefferman [11]; see also [28]. Among several equivalent ways of introducing product weights [16, Theorem VI.6.2], we adopt the following definition as in [4].

Definition 2.3. *Let $\vec{A} = (A_1, A_2)$ be a pair of expansive dilations, on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Let $p \in (1, \infty)$ and w be a nonnegative measurable function on $\mathbb{R}^n \times \mathbb{R}^m$. The function w is said to be in the weight class of Muckenhoupt $\mathcal{A}_p(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$, if $w(x_1, \cdot) \in \mathcal{A}_p(\mathbb{R}^m; A_2)$ for almost every*

$x_1 \in \mathbb{R}^n$ *and* $\operatorname{ess\,sup}_{x_1\in\mathbb{R}^n} C_{A_2}(w(x_1,\,\cdot)) < \infty$, *and* $w(\cdot,\,x_2) \in \mathcal{A}_p(\mathbb{R}^n;\,A_1)$ *for almost every* $x_2 \in \mathbb{R}^m$ *and* $\operatorname{ess\,sup}_{x_2\in\mathbb{R}^m} C_{A_1}(w(\cdot,\,x_2)) < \infty$. *In what follows, let*

$$C_{\vec{A}}(w) := \max\Big\{ \operatorname*{ess\,sup}_{x_1\in\mathbb{R}^n} C_{A_2}(w(x_1,\,\cdot)),\ \operatorname*{ess\,sup}_{x_2\in\mathbb{R}^m} C_{A_1}(w(\cdot,\,x_2))\Big\}.$$

Define $\mathcal{A}_\infty(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A}) := \bigcup_{1<p<\infty}\mathcal{A}_p(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A})$.

For any $w \in \mathcal{A}_\infty(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A})$, define the critical index of w by

$$q_w := \inf\{q \in (1,\,\infty) : w \in \mathcal{A}_q(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A})\}. \tag{2.4}$$

Let $\mathcal{S}(\mathbb{R}^n)$ be the *space of Schwartz functions on* $\mathbb{R}^n$. For $\alpha \in \mathbb{Z}_+^n$ and $m \in \mathbb{Z}_+$, define seminorms $\|\varphi\|_{\alpha,\,m} := \sup_{x\in\mathbb{R}^n}[\rho(x)]^m|\partial^\alpha\varphi(x)| < \infty$. It is well known that $\mathcal{S}(\mathbb{R}^n)$ forms a locally convex complete metric space endowed with the seminorms $\{\|\cdot\|_{\alpha,\,m}\}_{\alpha\in\mathbb{Z}_+^n,\,m\in\mathbb{Z}_+}$. The space $\mathcal{S}(\mathbb{R}^n)$ coincides with the classical space of Schwartz functions; see [1, p. 11]. The dual space of $\mathcal{S}(\mathbb{R}^n)$, namely, the space of tempered distributions on $\mathbb{R}^n$ is denoted by $\mathcal{S}'(\mathbb{R}^n)$. Moreover, let $\mathcal{S}_0(\mathbb{R}^n) := \big\{\psi \in \mathcal{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n}\psi(x)\,dx = 0\big\}$.

For functions φ on $\mathbb{R}^n$, ψ on $\mathbb{R}^n\times\mathbb{R}^m$, $k,\ k_1,\ k_2 \in \mathbb{Z}$, let $\varphi_k(x) := b^{-k}\varphi(A^{-k}x)$ and $\psi_{k_1,\,k_2}(x) := b_1^{-k_1}b_2^{-k_2}\psi(A_1^{-k_1}x_1,\,A_2^{-k_2}x_2)$.

Next, we introduce the product Lusin-area function and product Littlewood-Paley $\vec{g}$-function following [4].

Definition 2.4. *Let* $\varphi^{(1)} \in \mathcal{S}_0(\mathbb{R}^n)$ *and* $\varphi^{(2)} \in \mathcal{S}_0(\mathbb{R}^m)$. *Let* $\varphi := \varphi^{(1)}\otimes\varphi^{(2)}$, *where, for all* $x = (x_1,\,x_2) \in \mathbb{R}^n\times\mathbb{R}^m$, $\varphi(x) := \varphi^{(1)}(x_1)\varphi^{(2)}(x_2)$. *For all* $f \in \mathcal{S}'(\mathbb{R}^n\times\mathbb{R}^m)$ *and* $x \in \mathbb{R}^n\times\mathbb{R}^m$, *define the anisotropic product Lusin-area function of* f *by*

$$\vec{S}_\varphi(f)(x) := \Bigg\{\sum_{k_1,\,k_2\in\mathbb{Z}} b_1^{-k_1}b_2^{-k_2}\int_{B_{k_1}^{(1)}\times B_{k_2}^{(2)}}|\varphi_{k_1,\,k_2}*f(x-y)|^2\,dy\Bigg\}^{1/2}.$$

Define the anisotropic product Littlewood-Paley $\vec{g}$*-function of* f *by*

$$\vec{g}_\varphi(f)(x) := \Bigg\{\sum_{k_1,\,k_2\in\mathbb{Z}}|\varphi_{k_1,\,k_2}*f(x)|^2\Bigg\}^{1/2}.$$

A distribution $f \in \mathcal{S}'(\mathbb{R}^n\times\mathbb{R}^m)$ is said to *vanish weakly at infinity* if for any $\varphi^{(1)} \in \mathcal{S}(\mathbb{R}^n)$ and $\varphi^{(2)} \in \mathcal{S}(\mathbb{R}^m)$, $f*\varphi_{k_1,\,k_2} \to 0$ in $\mathcal{S}'(\mathbb{R}^n\times\mathbb{R}^m)$ as $k_1,\,k_2 \to \infty$. We denote by $\mathcal{S}'_\infty(\mathbb{R}^n\times\mathbb{R}^m)$ the *set of all* $f \in \mathcal{S}'(\mathbb{R}^n\times\mathbb{R}^m)$ *vanishing weakly at infinity.*

We shall need the following existence result for functions appearing in the Calderón formula; see [4, Propositions 2.14 and 2.16].

Proposition 2.5. *For* $i = 1,\,2$, *let* $s_i \in \mathbb{Z}_+$, A_i *be a dilation on* $\mathbb{R}^{n_i}$, *and* A_i^* *its transpose. Then, there exist* $\theta^{(i)},\,\psi^{(i)} \in \mathcal{S}(\mathbb{R}^{n_i})$ *such that*:

(i) $\operatorname{supp}\theta^{(i)} \subset B_0^{(i)}$, $\int_{\mathbb{R}^{n_i}} x_i^{\gamma_i}\theta^{(i)}(x_i)\,dx_i = 0$ *for all* $\gamma_i \in (\mathbb{Z}_+)^{n_i}$ *with* $|\gamma_i| \leqslant s_i$, $\widehat{\theta^{(i)}}(\xi_i) \geqslant C > 0$ *for* ξ_i *in some annulus*;

(ii) $\operatorname{supp}\widehat{\psi^{(i)}}$ *is compact and bounded away from the origin*;

(iii) $\sum_{j\in\mathbb{Z}}\widehat{\psi^{(i)}}((A_i^*)^j\xi_i)\widehat{\theta^{(i)}}((A_i^*)^j\xi_i) = 1$ *for all* $\xi_i \in \mathbb{R}^{n_i}\setminus\{0\}$;

(iv) $\psi^{(i)} = \phi^{(i)}*\phi^{(i)}$ *for some* $\phi^{(i)} \in \mathcal{S}(\mathbb{R}^{n_i})$.

Parts (i)–(iii) of Proposition 2.5 were proved in the course of the proof of [2, Theorem 5.8]. Part (iv) can be shown by a minor refinement of this argument leading to the existence of $\phi^{(i)} \in \mathcal{S}(\mathbb{R}^{n_i})$ such that $(\widehat{\phi^{(i)}})^2 = \widehat{\psi^{(i)}}$.

The following result says that the space $L^p_w(\mathbb{R}^n\times\mathbb{R}^m)$ can be characterized by the Lusin-area $\vec{S}$-function and the Littlewood-Paley $\vec{g}$-function. Proposition 2.6 is just [4, Theorem 3.2], which also holds for the $\vec{g}$-function by a similar proof.

Proposition 2.6. *Let* $\psi := \psi^{(1)}\otimes\psi^{(2)}$ *be as in Proposition* 2.5. *Then, the following are equivalent for* $p \in (1,\,\infty)$:

(i) $f \in L^p_w(\mathbb{R}^n \times \mathbb{R}^m)$;
(ii) $f \in \mathcal{S}'_\infty(\mathbb{R}^n \times \mathbb{R}^m)$ *and* $\vec{S}_\psi(f) \in L^p_w(\mathbb{R}^n \times \mathbb{R}^m)$;
(iii) $f \in \mathcal{S}'_\infty(\mathbb{R}^n \times \mathbb{R}^m)$ *and* $\vec{g}_\psi(f) \in L^p_w(\mathbb{R}^n \times \mathbb{R}^m)$.
Moreover, for all $f \in L^p_w(\mathbb{R}^n \times \mathbb{R}^m)$,

$$\|f\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \sim \|\vec{S}_\psi(f)\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \sim \|\vec{g}_\psi(f)\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)}.$$

Finally, we recall the definition of weighted anisotropic product Hardy spaces in [4].

Definition 2.7. *Let* $w \in \mathcal{A}_\infty(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$ *and* $p \in (0, 1]$. *Let* $\psi = \psi^{(1)} \otimes \psi^{(2)}$ *be as in Proposition* 2.5. *The weighted anisotropic product Hardy space is defined by*

$$H^p_w(\mathbb{R}^n \times \mathbb{R}^m; \vec{A}) := \{f \in \mathcal{S}'_\infty(\mathbb{R}^n \times \mathbb{R}^m) : \|f\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A})} := \|\vec{S}_\psi(f)\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} < \infty\}.$$

3 Proof of Theorem 1.5

To prove Theorem 1.5, we need the following decomposition of kernels technique, which adapts the methods established by Nagel and Stein [26, Lemma 3.5.1] to our setting. For the convenience of the reader, we present a detailed proof.

Lemma 3.1. *Let* $N \in \mathbb{N}$ *and* $\psi \in \mathcal{S}_0(\mathbb{R}^n)$. *For any* $M > 0$, *there exist a constant* $c > 0$ *and a decomposition* $\psi = \sum_{k=0}^\infty b^{-kM}\psi^{(k)}$ *such that each* $c\psi^{(k)} \in \mathcal{S}_0(\mathbb{R}^n)$ *is an* N*-normalized bump function associated with* B_k.

Proof. Let $\theta \in \mathcal{C}^\infty(\mathbb{R}^n)$ be a non-negative function such that $\operatorname{supp}\theta \subset B_0$, $\theta(x) = 1$ for all $x \in B_{-1}$, and $\|\partial^\alpha\theta\|_{L^\infty(\mathbb{R}^n)} \leqslant 1$ for $|\alpha| \leqslant N$. Obviously, θ is an N-normalized bump function associated with B_0. For all $x \in \mathbb{R}^n$ and $k \in \mathbb{N}$, set $D_0(x) := \psi(x)\theta(x)$ and $D_k(x) := \psi(x)[\theta(A^{-k}x) - \theta(A^{-(k-1)}x)]$. It is easy to check that $\psi(x) = \sum_{k=0}^\infty D_k(x)$ pointwise. For any $k \in \mathbb{Z}_+$, let $d_k := \int_{\mathbb{R}^n} D_k(x)\,dx$, $s_0 := 0$ and $s_k = \sum_{j=0}^{k-1} d_j$ for $k \geqslant 1$.

Notice that for any $k \in \mathbb{N}$, we have $\operatorname{supp} D_k \subset B_k \setminus B_{k-2}$. Fix $M > 0$. Since $D_k(x) \neq 0$ implies that $\rho(x) \sim b^k$, we have

$$|D_k(x)| \lesssim [\rho(x)]^{-M-1} \lesssim b^{-(M+1)k}, \tag{3.1}$$

because $\psi \in \mathcal{S}_0(\mathbb{R}^n)$ and $\|\theta\|_{L^\infty(\mathbb{R}^n)} \leqslant 1$. From this and $\operatorname{supp} D_k \subset B_k$, it follows that $\sum_{k=0}^\infty \int_{\mathbb{R}^n} |D_k(x)|dx \lesssim 1$. Using that $\psi = \sum_{k=0}^\infty D_k$ and $\psi \in \mathcal{S}_0(\mathbb{R}^n)$, we obtain $\sum_{k=0}^\infty d_k = \int_{\mathbb{R}^n} \psi(x)\,dx = 0$. Thus, we also have $s_k = -\sum_{j\geqslant k} d_j$. Moreover, from (3.1) and $\operatorname{supp} D_k \subset B_k$, it follows that $|d_k| \lesssim b^{-kM}$, and hence $|s_k| \lesssim b^{-kM}$.

For any $k \in \mathbb{Z}_+$ and $x \in \mathbb{R}^n$, we define

$$\widetilde{D}_k(x) := D_k(x) - d_k b^{-k}\widetilde{\theta}(A^{-k}x) + s_k[b^{-(k-1)}\widetilde{\theta}(A^{-(k-1)}x) - b^{-k}\widetilde{\theta}(A^{-k}x)],$$

where $\widetilde{\theta}(x) := \theta(x)/\|\theta\|_{L^1(\mathbb{R}^n)}$. We claim that $\psi^{(k)} := b^{Mk}\widetilde{D}_k \in \mathcal{S}_0(\mathbb{R}^n)$ is the desired constant multiple of an N-normalized bump function associated with B_k. Indeed, it is easy to check that $\widetilde{D}_k \in \mathcal{C}^\infty(\mathbb{R}^n)$ with $\operatorname{supp}\widetilde{D}_k \subset B_k$ and $\int_{\mathbb{R}^n} \widetilde{D}_k(x)\,dx = 0$. Using $\sum_{k=0}^\infty d_k = 0$ and $s_k = -\sum_{j\geqslant k} d_k$, by Abel's summation, we have

$$\sum_{k=0}^\infty s_k[b^{-(k-1)}\widetilde{\theta}(A^{-(k-1)}x) - b^{-k}\widetilde{\theta}(A^{-k}x)] = \sum_{k=0}^\infty d_k b^{-k}\widetilde{\theta}(A^{-k}x).$$

This together with $\psi = \sum_{k=0}^\infty D_k$ implies that $\psi = \sum_{k=0}^\infty \widetilde{D}_k = \sum_{k=0}^\infty b^{-Mk}\psi^{(k)}$.

Finally, it remains to show that $\|\partial^\alpha\widetilde{D}_k(A^k\cdot)\|_{L^\infty(\mathbb{R}^n)} \lesssim b^{-Mk}$ for any $k \in \mathbb{Z}_+$ and $|\alpha| \leqslant N$. Since $\|\partial^\alpha\theta\|_{L^\infty(\mathbb{R}^n)}$, $\|\partial^\alpha\theta(A\cdot)\|_{L^\infty(\mathbb{R}^n)} \lesssim 1$, and $|s_k|$, $|d_k| \lesssim b^{-Mk}$, it suffices to prove

$$\|\partial^\alpha D_k(A^k\cdot)\|_{L^\infty(\mathbb{R}^n)} \lesssim b^{-Mk}.$$

Recall that $\operatorname{supp} D_k \subset B_k \setminus B_{k-2}$ for $k \in \mathbb{N}$. Thus, we only need to check that $|\partial^\alpha D_k(A^k\cdot)(x)| \lesssim b^{-Mk}$ for all $x \in B_0 \setminus B_{-2}$ for all $|\alpha| \leqslant N$. Since $\psi \in \mathcal{S}(\mathbb{R}^n)$, for all $x \in B_0 \setminus B_{-2}$ and for all $|\alpha| \leqslant N$, we have

$$|\partial^\alpha D_k(A^k\cdot)(x)| = |\partial^\alpha[\psi(A^k\cdot)(\theta(\cdot) - \theta(A^{-1}\cdot))](x)|$$

$$\lesssim \sum_{|\beta|\leqslant|\alpha|} |\partial^\beta \psi(A^k\cdot)(x)| \lesssim \|A^k\|^{|\alpha|} \sum_{|\beta|\leqslant|\alpha|} |\partial^\beta \psi(A^k x)|$$
$$\lesssim \|A\|^{Nk}[\rho(A^k x)]^{-M'} \lesssim \|A\|^{Nk} b^{-kM'} \lesssim b^{-kM}. \tag{3.2}$$

This finishes the proof of the claim and hence of Lemma 3.1.

For $i = 1, 2$, let A_i be a dilation on $\mathbb{R}^{n_i}$ as in Definition 1.1. Let $\lambda_{i,-}$ and $\lambda_{i,+}$ be two *positive numbers* such that

$$1 < \lambda_{i,-} < \min\{|\lambda| : \lambda \in \sigma(A_i)\} \leqslant \max\{|\lambda| : \lambda \in \sigma(A_i)\} < \lambda_{i,+}.$$

In the case when A_i is diagonalizable over $\mathbb{C}$, we can even take $\lambda_{i,-} = \min$ and $\lambda_{i,+} = \max$ above. Otherwise, we need to choose them sufficiently close to these equalities according to what we need in our arguments. Let $\zeta_{i,\pm} = \log_{b_i} \lambda_{i,\pm}$. It is useful to make some remarks about Definitions 1.3 and 1.4.

Remark 3.2. (i) One can show that if T is a PASIO of order $(1, 1)$, then it is also a PASIO of order 0. This is a consequence of Lemma 3.3 below.

(ii) In Definition 1.4, the range of ϵ_i is effectively restricted to the interval $(0, \log_{b_i} |\lambda_{i,+}|]$, where $\lambda_{i,+}$ denotes the largest eigenvalue of A_i in absolute value. Let us explain this in the one parameter setting. Assume that $\epsilon > \log_b |\lambda_+|$ and $|K(x)| \leqslant C_1[\rho(x)]^{-1}$ and $|K(x + h) - K(x)| \leqslant C_1[\rho(h)]^\epsilon[\rho(x)]^{-1-\epsilon}$ for $\rho(h) \leqslant b^{-2\sigma}\rho(x)$ and $x \neq 0$. Choose λ such that $|\lambda_+| < \lambda < b^\epsilon$ and let $\zeta := \log_b \lambda$. For any $x \neq 0$, when $\rho(h) \leqslant \min\{1, b^{-2\sigma}\rho(x)\}$, by $[\rho(h)]^\zeta \leqslant C_1|h|$ (see [1, (3.3)]), we have

$$|K(x + h) - K(x)| \leqslant C_1[\rho(x)]^{-1-\epsilon}|h|^{\epsilon/\zeta} \leqslant C_1[\rho(x)]^{-1-\epsilon}|h|,$$

which implies that K is locally Lipschitz continuous away from 0. Moreover, for all $x \neq 0$,

$$\limsup_{h\to 0} \frac{1}{|h|}|K(x + h) - K(x)| \leqslant C_1[\rho(x)]^{-1-\epsilon} \limsup_{h\to 0}[\rho(h)]^{\epsilon-\zeta} = 0,$$

which implies that K is a constant function away from 0 and thus, by $|K(x)| \leqslant C_1[\rho(x)]^{-1}$, we further have $K(x) = 0$ for all $x \neq 0$.

Lemma 3.3. *Let K be the kernel of a PASIO of order $(s_1 + 1, s_2 + 1)$, where $s_1, s_2 \in \mathbb{Z}_+$. Then, there exist positive constants C and ϵ_i such that K has the following three additional properties:*

(K1)$'$ *for all $(x_1, x_2) \in \Omega_{n\times m}$ with $\rho_1(x_1) = b_1^{\ell_1}$ for some $\ell_1 \in \mathbb{Z}$, $h_1 \in \mathbb{R}^n$ with $\rho_1(h_1) \leqslant b_1^{-2\sigma_1}\rho_1(x_1)$ and $\alpha_1 \in \mathbb{Z}_+^n$ with $|\alpha_1| = s_1$,*

$$|\Delta^{(1)}_{A_1^{-\ell_1}h_1} \partial_1^{\alpha_1}[K(A_1^{\ell_1}\cdot, x_2)](A_1^{-\ell_1}x_1)| \leqslant C_1 \frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}} \frac{1}{\rho_2(x_2)}.$$

This also holds with the roles of x_1 and x_2 interchanged;

(K1)$''$ *for all $(x_1, x_2) \in \Omega_{n\times m}$ with $\rho_i(x_i) = b_i^{\ell_i}$ for some $\ell_i \in \mathbb{Z}$, $h_i \in \mathbb{R}^{n_i}$ with $\rho_i(h_i) \leqslant b_i^{-2\sigma_i}\rho_i(x_i)$ and $\alpha_i \in \mathbb{Z}_+^{n_i}$ with $|\alpha_i| = s_i$, $i = 1, 2$,*

$$|\Delta^{(1)}_{A_1^{-\ell_1}h_1} \Delta^{(2)}_{A_2^{-\ell_2}h_2} \partial_1^{\alpha_1}\partial_2^{\alpha_2}[K(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}x_1, A_2^{-\ell_2}x_2)| \leqslant C_1 \frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}} \frac{[\rho_2(h_2)]^{\epsilon_2}}{[\rho_2(x_2)]^{1+\epsilon_2}};$$

(K3)$'$ *the kernel $K^{\psi^{(2)}, k_2}$ as in* (K3) *satisfies that for all $x_1 \in \mathbb{R}^n \setminus \{0\}$, $h_1 \in \mathbb{R}^n$ with $\rho_1(h_1) \leqslant b_1^{-2\sigma_1}\rho_1(x_1)$, and $\alpha_1 \in \mathbb{Z}_+^n$ with $|\alpha_1| = s_1$,*

$$|\Delta^{(1)}_{A_1^{-\ell_1}h_1} \partial_1^{\alpha_1}[K^{\psi^{(2)}, k_2}(A_1^{\ell_1}\cdot)](A_1^{-\ell_1}x_1)| \leqslant C_1 \frac{[\rho_1(h_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}}.$$

This also holds with the roles of x_1 and x_2 interchanged.

Proof. We will only prove (K3)$'$. The other properties are shown in the same fashion. Let $K := K^{\psi^{(2)}, k_2}$ be the kernel as in (K3). Assume that $\rho(x_1) = b_1^{\ell_1}$ and $\rho(h_1) \leqslant b_1^{-2\sigma_1+\ell_1}$. Take any $\epsilon_1 \in (0, \log_{b_1} \lambda_{1,-})$. By (K3) and the Taylor's formula, for $|\alpha_1| = s_1$ we have

$$|\Delta_{A_1^{-\ell_1}h_1} \partial_1^{\alpha_1}[K(A^{\ell_1}\cdot)](A_1^{-\ell_1}x_1)| \leqslant |A_1^{-\ell_1}h_1| \sup_{\rho(z_1)\leqslant b_1^{-2\sigma_1+\ell_1}} |\nabla\partial_1^{\alpha_1}[K(A_1^{\ell_1}\cdot)](A_1^{-\ell_1}(x_1 + z_1))|$$

$$\lesssim |A_1^{-\ell_1}h_1|b_1^{-\ell_1} \lesssim [\rho(h_1)]^{\epsilon_1}[\rho(x_1)]^{-(1+\epsilon_1)},$$

which completes the proof.

The following lemma plays a key role in the proofs of Theorems 1.5 and 1.6 by generalizing [22, Lemma 1] to the anisotropic setting and to the higher order partial derivatives of corresponding kernels.

Lemma 3.4. *Let K be the kernel of a PASIO of order (s_1+1, s_2+1), where $s_1, s_2 \in \mathbb{Z}_+$. For $i=1, 2$, let $\varphi^{(i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$ be an (N_i+s_i+1)-normalized bump function associated with some dilated ball $B_{j_i}^{(i)}$, where $j_i \in \mathbb{Z}_+$ and $N_i \in \mathbb{N}$ is as in* (K2) *and* (K3) *of Definition* 1.3. *For all $k_1, k_2 \in \mathbb{Z}$, define $K_{k_1,k_2} := K * \varphi_{k_1,k_2}$, where $\varphi := \varphi^{(1)} \otimes \varphi^{(2)}$. Then, there exist positive constants C and ϵ_i such that: for all h_i, x_i, $y_i \in \mathbb{R}^{n_i}$ with $\rho_i(x_i) = b_i^{\ell_i}$ for some $\ell_i \in \mathbb{Z}$ and $\rho_i(x_i - y_i) < b_i^{k_i}$, and $\alpha_i \in \mathbb{Z}_+^{n_i}$ with $|\alpha_i| \leqslant s_i$, $i = 1, 2$,*

$$|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K_{k_1,k_2}(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)| \leqslant C\prod_{i=1}^{2}\frac{b_i^{k_i\epsilon_i}}{[b_i^{k_i}+b_i^{-j_i}\rho_i(x_i)]^{1+\epsilon_i}}.$$

Proof. To prove this lemma, we first present two basic facts. Let $i = 1, 2$. For any $\alpha_i \in \mathbb{Z}_+^{n_i}$, by (3.13) in [2] when $\ell_i - k_i < 0$, or a similar proof when $\ell_i - k_i \geqslant 0$, for all x_i, $z_i \in \mathbb{R}^{n_i}$, we have

$$\begin{aligned}\partial^{\alpha_i}[\varphi^{(i)}(A_i^{\ell_i-k_i}\cdot - A_i^{-k_i}z_i)](A_i^{-\ell_i}y_i) &= \partial^{\alpha_i}[\varphi^{(i)}(A_i^{\ell_i-k_i}\cdot)](A_i^{-\ell_i}(y_i - z_i))\\ &= \sum_{|\beta_i|=|\alpha_i|} a_{\beta_i}^{(i)}\partial^{\beta_i}[\varphi^{(i)}(A_i^{j_i}\cdot)](A_i^{-j_i-k_i}(y_i-z_i)),\end{aligned} \tag{3.3}$$

where

$$|a_{\beta_i}^{(i)}| \lesssim b_i^{(\ell_i-j_i-k_i)|\beta_i|\zeta_{i,-}} \quad \text{if } \ell_i - j_i - k_i \leqslant 0, \tag{3.4}$$

and

$$|a_{\beta_i}^{(i)}| \lesssim b_i^{(\ell_i-j_i-k_i)|\beta_i|\zeta_{i,+}} \quad \text{if } \ell_i - j_i - k_i > 0. \tag{3.5}$$

Moreover, for any fixed $x_i \in \mathbb{R}^{n_i}$ with $\rho_i(x_i) = b_i^{\ell_i}$, if $\ell_i \leqslant k_i + j_i + 4\sigma_i$, we claim that

$$\xi_{\beta_i}^{(i)}(z_i) := \partial^{\beta_i}[\varphi^{(i)}(A_i^{j_i}\cdot)](A_i^{-j_i-k_i}y_i - A_i^{6\sigma_i+1}z_i) \tag{3.6}$$

is an N_i-normalized bump function associated with $B_0^{(i)}$. Indeed, if $\xi_{\beta_i}^{(i)}(z_i) \neq 0$, then by $\operatorname{supp}(\partial^{\beta_i}\varphi^{(i)}) \subset B_{j_i}^{(i)}$, $x_i \in B_{\ell_i+1}^{(i)}$, $y_i \in x_i + B_{k_i+1}^{(i)}$, $\ell_i \leqslant k_i + j_i + 4\sigma_i$ and (2.2), we obtain

$$z_i \in A_i^{-k_i-j_i-6\sigma_i-1}y_i + B_{-6\sigma_i-1}^{(i)} \subset B_{-\sigma_i}^{(i)} + B_{-6\sigma_i-1}^{(i)} \subset B_0^{(i)}.$$

Moreover, since $\varphi^{(i)}$ is an (s_i+N_i+1)-normalized bump function associated with $B_{j_i}^{(i)}$, then for all $z_i \in \mathbb{R}^{n_i}$ and $\gamma_i \in \mathbb{Z}_+^{n_i}$ with $|\gamma_i| \leqslant N_i$, we have $|\partial^{\gamma_i}(\xi_{\beta_i}^{(i)})(z_i)| \lesssim 1$. Thus, the above claim holds.

We now show Lemma 3.4 by considering the following four cases. In Case (i) through Case (iv), we assume that $\rho_i(x_i) = b_i^{\ell_i}$ for some $\ell_i \in \mathbb{Z}$ and $\alpha_i \in \mathbb{Z}_+^{n_i}$ with $|\alpha_i| \leqslant s_i$, $i = 1, 2$.

Case (i). $\ell_1 \leqslant k_1 + j_1 + 4\sigma_1$ and $\ell_2 \leqslant k_2 + j_2 + 4\sigma_2$. In this case, by (3.3)–(3.6), (K2), $|\alpha_i| \leqslant s_i$, $\zeta_{i,+} = \log_{b_i}\lambda_{i,+} < 1$ and $j_i \geqslant 0$, we have

$$\begin{aligned}&|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K_{k_1,k_2}(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)|\\ &\quad= \left|b_1^{-k_1}b_2^{-k_2}\sum_{\substack{|\beta_1|\leqslant|\alpha_1|\\|\beta_2|\leqslant|\alpha_2|}} a_{\beta_1}^{(1)}a_{\beta_2}^{(2)}\left\langle K, \bigotimes_{i=1}^{2}\xi_{\beta_i}^{(i)}(A_i^{-j_i-k_i-6\sigma_i-1}\cdot)\right\rangle\right| \lesssim b_1^{-k_1}b_2^{-k_2},\end{aligned}$$

which is as desired. Here

$$\bigotimes_{i=1}^{2}\xi_{\beta_i}^{(i)}(A_i^{-j_i-k_i-6\sigma_i-1}\cdot) := \xi_{\beta_1}^{(1)}(A_1^{-j_1-k_1-6\sigma_1-1}\cdot)\bigotimes\xi_{\beta_2}^{(2)}(A_2^{-j_2-k_2-6\sigma_2-1}\cdot).$$

Case (ii). $\ell_1 \leqslant k_1 + j_1 + 4\sigma_1$ and $\ell_2 > k_2 + j_2 + 4\sigma_2$. In this case, if $z_2 \in B^{(2)}_{k_2+j_2}$, $\rho_2(x_2 - y_2) < b_2^{k_2}$ and $x_2 \in B^{(2)}_{\ell_2+1} \setminus B^{(2)}_{\ell_2}$ with $\ell_2 > k_2 + j_2 + 4\sigma_2$, then by Definition 2.1, it is easy to obtain that $\rho_2(y_2) \geqslant b^{\ell_2-\sigma_2}$ and $\rho_2(z_2) < b_2^{-3\sigma_2}\rho_2(y_2)$. Thus, by $\varphi^{(2)} \in \mathcal{S}_0(\mathbb{R}^m)$, (3.3)–(3.6) with $i = 1$, $\operatorname{supp}\varphi^{(2)}(A_2^{-k_2}\cdot) \subset B^{(2)}_{j_2+k_2}$ and (K3)$'$, we have

$$
\begin{aligned}
&|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K_{k_1,k_2}(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)|\\
&\quad = \left| b_1^{-k_1} \sum_{|\beta_1|\leqslant|\alpha_1|} a^{(1)}_{\beta_1} \int_{\mathbb{R}^m} \varphi^{(2)}_{k_2}(z_2)\Delta^{(2)}_{-A_2^{\ell_2}z_2}\partial_2^{\alpha_2}[K^{\xi^{(1)}_{\beta_1}, -j_1-k_1-6\sigma_1-1}(A_2^{\ell_2}\cdot)](A_2^{-\ell_2}y_2)\,dz_2 \right|\\
&\quad \lesssim b_1^{-k_1}\int_{B^{(2)}_{j_2+k_2}} \frac{[\rho_2(z_2)]^{\epsilon_2}}{[\rho_2(x_2)]^{1+\epsilon_2}}|\varphi^{(2)}_{k_2}(z_2)|\,dz_2 \lesssim b_1^{-k_1} b_2^{j_2+(j_2+k_2)\epsilon_2-\ell_2(1+\epsilon_2)},
\end{aligned}
$$

which is desired.

Case (iii). $\ell_1 > k_1 + j_1 + 4\sigma_1$ and $\ell_2 \leqslant k_2 + j_2 + 4\sigma_2$. In this case, by symmetry, similar to the estimate of Case (ii), we also have

$$
|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K_{k_1,k_2}(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)| \lesssim b_1^{j_1+(j_1+k_1)\epsilon_1-\ell_1(1+\epsilon_1)} b_2^{-k_2}.
$$

Case (iv). $\ell_1 > k_1 + j_1 + 4\sigma_1$ and $\ell_2 > k_2 + j_2 + 4\sigma_2$. In this case, for $i = 1, 2$, $z_i \in B^{(i)}_{j_i+k_i}$, $\rho_i(x_i - y_i) < b^{k_i}$ and $\rho_i(x_i) = b_i^{\ell_i}$, we have $\rho_i(y_i) \geqslant b_i^{\ell_i-\sigma_i}$ and $\rho_i(z_i) < b_i^{-3\sigma_i}\rho_i(y_i)$. By this, $\varphi^{(i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$, $\operatorname{supp}\varphi^{(i)}(A_i^{-k_i}\cdot) \subset B^{(i)}_{j_i+k_i}$ and (K1)$''$, we obtain

$$
\begin{aligned}
&|\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K_{k_1,k_2}(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)|\\
&\quad = \left| \int_{\mathbb{R}^n\times\mathbb{R}^m} \varphi^{(1)}_{k_1}(z_1)\varphi^{(2)}_{k_2}(z_2)\Delta^{(1)}_{-A_1^{-\ell_1}z_1}\Delta^{(2)}_{-A_2^{-\ell_2}z_2}\partial_1^{\alpha_1}\partial_2^{\alpha_2}[K(A_1^{\ell_1}\cdot, A_2^{\ell_2}\cdot)](A_1^{-\ell_1}y_1, A_2^{-\ell_2}y_2)\,dz \right|\\
&\quad \lesssim \int_{B^{(1)}_{j_1+k_1}\times B^{(2)}_{j_2+k_2}} |\varphi^{(1)}_{k_1}(z_1)\varphi^{(2)}_{k_2}(z_2)| \frac{[\rho_1(z_1)]^{\epsilon_1}}{[\rho_1(x_1)]^{1+\epsilon_1}} \frac{[\rho_2(z_2)]^{\epsilon_2}}{[\rho_2(x_2)]^{1+\epsilon_2}}\,dz\\
&\quad \lesssim \prod_{i=1}^{2} b_i^{j_i+\epsilon_i(j_i+k_i)-\ell_i(1+\epsilon_i)},
\end{aligned}
$$

which is desired.

Combining the above estimates then completes the proof of Lemma 3.4.

Remark 3.5. Notice that in the proof of Lemma 3.4, we have not used explicitly the bounds on the highest order derivatives of K. Instead, we used the difference properties (K1)$'$, (K1)$''$, and (K3)$'$ from Lemma 3.3. Thus, if K is merely a kernel of a PASIO of order 0, then the conclusions of Lemma 3.4 apply. In particular, there exists a positive constant C such that for all $h_i, x_i, y_i \in \mathbb{R}^{n_i}$ with $\rho_i(x_i) = b_i^{\ell_i}$ for some $\ell_i \in \mathbb{Z}$ and $\rho_i(x_i - y_i) < b_i^{k_i}$,

$$
|K_{k_1,k_2}(y_1, y_2)| \leqslant C \prod_{i=1}^{2} \frac{b_i^{k_i\epsilon_i}}{[b_i^{k_i} + b_i^{-j_i}\rho_i(x_i)]^{1+\epsilon_i}}. \tag{3.7}
$$

Proof of Theorem 1.5. Let $p \in (1, \infty)$ and $w \in \mathcal{A}_p(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$. Let T be a PASIO of order 0 with kernel K as in Definition 1.4. Let N_1 and N_2 be as in (K2) and (K3). Let $\psi := \psi^{(1)} \otimes \psi^{(2)}$ and $\phi := \phi^{(1)} \otimes \phi^{(2)}$ be as in Proposition 2.5. By part (iv) of this proposition, we have $\psi = \phi * \phi$. From Proposition 2.6, it follows that for all $f \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$,

$$
\|\vec{S}_{\phi*\phi}(f)\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \sim \|f\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)}. \tag{3.8}
$$

Since $\phi^{(i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$ for $i = 1, 2$, then by Lemma 3.1, we have $\phi^{(i)} = \sum_{j_i=0}^{\infty} b_i^{-4j_i}\phi^{(i,j_i)}$, where $\phi^{(i,j_i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$ is a constant multiple of an (N_1+1)-normalized bump function associated with $B^{(i)}_{j_i}$. For $j_1, j_2 \in \mathbb{Z}_+$ and $k_1, k_2 \in \mathbb{Z}$, let $\phi^{\{j_1,j_2\}} := \phi^{(1,j_1)} \otimes \phi^{(2,j_2)}$ and $K^{j_1,j_2}_{k_1,k_2} := K * \phi^{\{j_1,j_2\}}_{k_1,k_2}$. For any

$x,\, z \in \mathbb{R}^n \times \mathbb{R}^m$, $k_1,\, k_2 \in \mathbb{Z}$, $j_1,\, j_2 \in \mathbb{Z}_+$, $y \in \mathbb{R}^n \times \mathbb{R}^m$ with $\rho_1(y_1) < b_1^{k_1}$ and $\rho_2(y_2) < b_2^{k_2}$, and locally integrable function f on $\mathbb{R}^n \times \mathbb{R}^m$, by the estimate (3.7) in Remark 3.2 and $b_i^{k_i} + b_i^{-j_i}\rho_i(z_i) \sim b_i^{k_i} + b_i^{-j_i}\rho_i(z_i - y_i)$, $i = 1, 2$, we obtain

$$\begin{aligned}|f * K_{k_1,\, k_2}^{j_1,\, j_2}(x-y)| &\lesssim \int_{\mathbb{R}^n\times\mathbb{R}^m} |f(x-y-z)| \prod_{i=1}^{2} \frac{b_i^{k_i\epsilon_i}}{[b_i^{k_i} + b_i^{-j_i}\rho_i(z_i)]^{1+\epsilon_i}}\, dz \\ &\lesssim \int_{\mathbb{R}^n\times\mathbb{R}^m} |f(x-z)| \prod_{i=1}^{2} \frac{b_i^{k_i\epsilon_i}}{[b_i^{k_i} + b_i^{-j_i}\rho_i(z_i)]^{1+\epsilon_i}}\, dz \\ &\lesssim b_1^{j_1(1+\epsilon_1)} b_2^{j_2(1+\epsilon_2)} \mathcal{M}_s(f)(x), \end{aligned} \tag{3.9}$$

where and in what follows, $\mathcal{M}_s(f)$ denotes the *strong maximal function* which is defined by setting, for all $x \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\mathcal{M}_s(f)(x) := \sup_{k_1,\, k_2\in\mathbb{Z}} \sup_{x\in y+B_{k_1}^{(1)}\times B_{k_2}^{(2)}} \frac{1}{b_1^{k_1} b_2^{k_2}} \int_{y+B_{k_1}^{(1)}\times B_{k_2}^{(2)}} |f(z)|\, dz.$$

Thus, by (3.8), (3.9), the weighted vector-valued maximal inequality for $\mathcal{M}_s$ (see [4, Proposition 2.2]), and the $L_w^p(\mathbb{R}^n \times \mathbb{R}^m)$-boundedness of $\vec{g}_\phi$, which was proved in the proof of [4, Theorem 3.2], we have that for $f \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$,

$$\begin{aligned}\|Tf\|_{L_w^p(\mathbb{R}^n\times\mathbb{R}^m)} &\lesssim \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} b_1^{-4j_1} b_2^{-4j_2} \\ &\quad \times \left\| \left\{ \sum_{k_1,\, k_2\in\mathbb{Z}} \frac{1}{b_1^{k_1} b_2^{k_2}} \int_{B_{k_1}^{(1)}\times B_{k_2}^{(2)}} |K_{k_1,\, k_2}^{j_1,\, j_2} * f * \phi_{k_1,\, k_2}(\cdot - y)|^2\, dy \right\}^{1/2} \right\|_{L_w^p(\mathbb{R}^n\times\mathbb{R}^m)} \\ &\lesssim \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} b_1^{-2j_1} b_2^{-2j_2} \left\| \left\{ \sum_{k_1,\, k_2\in\mathbb{Z}} |\mathcal{M}_s\left(f * \phi_{k_1,\, k_2}\right)|^2 \right\}^{1/2} \right\|_{L_w^p(\mathbb{R}^n\times\mathbb{R}^m)} \\ &\lesssim \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} b_1^{-2j_1} b_2^{-2j_2} \|\vec{g}_\phi(f)\|_{L_w^p} \\ &\lesssim \|f\|_{L_w^p(\mathbb{R}^n\times\mathbb{R}^m)}. \end{aligned}$$

This combined with the density of $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$ in $L_w^p(\mathbb{R}^n \times \mathbb{R}^m)$ completes the proof of Theorem 1.5.

4 Proof of Theorem 1.6

To prove Theorem 1.6, we need to use a vector-valued variant of the boundedness criterion established in [4, Corollary 6.5]. We shall use an analogue of the grid of Euclidean dyadic cubes which is mainly due to Christ [9] and formulated as in [4, Lemma 2.2].

Lemma 4.1. *Let A be a dilation. There exists a collection $\mathcal{Q} := \{Q_\alpha^k \subset \mathbb{R}^n :\ k \in \mathbb{Z},\ \alpha \in \mathrm{I}_k\}$ of open subsets, where I_k is some index set, such that*

(i) $|\mathbb{R}^n \setminus \cup_\alpha Q_\alpha^k| = 0$ *for each fixed k and $Q_\alpha^k \cap Q_\beta^k = \emptyset$ if $\alpha \neq \beta$;*

(ii) *for any α, β, k, ℓ with $\ell \geqslant k$, either $Q_\alpha^k \cap Q_\beta^\ell = \emptyset$ or $Q_\alpha^\ell \subset Q_\beta^k$;*

(iii) *for each (ℓ, β) and each $k < \ell$, there exists a unique α such that $Q_\beta^\ell \subset Q_\alpha^k$;*

(iv) *there exist some negative integer v and positive integer u such that for all Q_α^k with $k \in \mathbb{Z}$ and $\alpha \in \mathrm{I}_k$, there exists $x_{Q_\alpha^k} \in Q_\alpha^k$ satisfying that for all $x \in Q_\alpha^k$, $x_{Q_\alpha^k} + B_{vk-u} \subset Q_\alpha^k \subset x + B_{vk+u}$.*

In what follows, for convenience, we call $\{Q_\alpha^k\}_{k\in\mathbb{Z},\, \alpha\in\mathrm{I}_k}$ dyadic cubes. Also for any dyadic cube Q_α^k with $k \in \mathbb{Z}$ and $\alpha \in \mathrm{I}_k$, we set its *level* as $\ell(Q_\alpha^k) := k$.

Let A_i be a dilation on $\mathbb{R}^{n_i}$, and $\mathcal{Q}^{(i)}$, $\ell(Q_i)$, v_i, u_i the same as in Lemma 4.1 corresponding to A_i for $i = 1, 2$. Let $\mathcal{R} := \mathcal{Q}^{(1)} \times \mathcal{Q}^{(2)}$. For $R \in \mathcal{R}$, we always write $R := R_1 \times R_2$ with $R_i \in \mathcal{Q}^{(i)}$ and call R a dyadic rectangle. We need the notion of rectangular atoms for anisotropic product Hardy spaces.

Definition 4.2. *Let $w \in \mathcal{A}_\infty(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$ and q_w be as in (2.4). The triplet $(p,\ q,\ \vec{s})_w$ is called admissible if $p \in (0,\ 1]$, $q \in [2,\ \infty) \cap (q_w,\ \infty)$ and $\vec{s} := (s_1,\ s_2)$ with $s_i \in \mathbb{Z}_+$ and $s_i \geqslant \lfloor (\frac{q_w}{p} - 1)\zeta_{i,-}^{-1} \rfloor$, $i = 1,\ 2$. For any $R \in \mathcal{R}$, a function a_R is called a rectangular $(p,\ q,\ \vec{s})_w$-atom if*

(i) *a_R is supported on $R'' := R_1'' \times R_2''$, where $R_i'' := x_{R_i} + B^{(i)}_{v_i(\ell(R_i)-1)+u_i+3\sigma_i}$, $i = 1,\ 2$;*

(ii)

$$\int_{\mathbb{R}^n} a_R(x_1,\ x_2) x_1^\alpha \, dx_1 = 0 \quad \text{for all } |\alpha| \leqslant s_1 \text{ and almost all } x_2 \in \mathbb{R}^m,$$

and

$$\int_{\mathbb{R}^m} a_R(x_1,\ x_2) x_2^\beta \, dx_2 = 0 \quad \text{for all } |\beta| \leqslant s_2 \text{ and almost all } x_1 \in \mathbb{R}^n;$$

(iii) $\|a\|_{L^q_w(\mathbb{R}^n \times \mathbb{R}^m)} \leqslant [w(R)]^{1/q-1/p}$.

We also need to consider the vector-valued space $\mathcal{H} := \{\{f_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}} : f_{k_1,\,k_2}$ is a measurable function on $B^{(1)}_{k_1} \times B^{(2)}_{k_2}$ for any $k_1,\ k_2 \in \mathbb{Z}$ and $|\{f_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}|_{\mathcal{H}} < \infty\}$, where

$$|\{f_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}|_{\mathcal{H}} := \left\{ \sum_{k_1,\,k_2 \in \mathbb{Z}} b_1^{-k_1} b_2^{-k_2} \int_{B^{(1)}_{k_1} \times B^{(2)}_{k_2}} |f_{k_1,\,k_2}(y)|^2 \, dy \right\}^{1/2}.$$

In what follows, for all $x \in \mathbb{R}^n \times \mathbb{R}^m$, we always write

$$|\{f_{k_1,\,k_2}(x)\}_{k_1,\,k_2 \in \mathbb{Z}}|_{\mathcal{H}} := \left\{ \sum_{k_1,\,k_2 \in \mathbb{Z}} b_1^{-k_1} b_2^{-k_2} \int_{B^{(1)}_{k_1} \times B^{(2)}_{k_2}} |f_{k_1,\,k_2}(x-y)|^2 \, dy \right\}^{1/2}.$$

Finally, let $p \in (0,\ \infty)$ and $w \in \mathcal{A}_\infty(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$. Define the vector-valued space $L^p_{w,\,\mathcal{H}}(\mathbb{R}^n \times \mathbb{R}^m)$ as the collection of all sequences $\{f_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}$ of measurable functions on $\mathbb{R}^n \times \mathbb{R}^m$ with the norm

$$\|\{f_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}\|_{L^p_{w,\,\mathcal{H}}(\mathbb{R}^n \times \mathbb{R}^m)} := \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^m} |\{f_{k_1,\,k_2}(x)\}_{k_1,\,k_2 \in \mathbb{Z}}|^p_{\mathcal{H}} \, w(x) \, dx \right\}^{1/p} < \infty.$$

The following conclusion is the vector-valued variant of [4, Corollary 6.5], whose proof is similar to that of [4, Corollary 6.1]; see also [8, Corollary 1.1] for the corresponding result on $H^p(\mathbb{R}^n \times \mathbb{R}^m)$. We omit the details.

Lemma 4.3. *Let $(p,\ q_1,\ \vec{s})_w$ be an admissible triplet as in Definition 4.2, $q_0 \in [q_1,\ \infty)$ and*

$$\{T_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}$$

be an $\mathcal{H}$-valued linear operator bounded from $L^{q_1}_w(\mathbb{R}^n \times \mathbb{R}^m)$ to $L^{q_0}_{w,\,\mathcal{H}}(\mathbb{R}^n \times \mathbb{R}^m)$. Let $q \in [p,\ 2)$ be such that $1/q - 1/p = 1/q_0 - 1/q_1$.

Suppose that there exist positive constants C, ϵ such that for all $\gamma \in \mathbb{Z}_+$ and all rectangular $(p,\ q_1,\ \vec{s})_w$-atoms a_R,

$$\int_{(R_{1,\gamma} \times R_{2,\,\gamma})^\complement} |\{T_{k_1,\,k_2}(a_R)(x)\}_{k_1,\,k_2 \in \mathbb{Z}}|^q_{\mathcal{H}} w(x) \, dx \leqslant C \max\{b_1^{-\gamma\epsilon}, b_2^{-\gamma\epsilon}\},$$

where $R_{i,\gamma} := x_{R_i} + B^{(i)}_{v_i(\ell(R_i)-1)+u_i+5\sigma_i+\gamma}$, $i = 1, 2$. Then $\{T_{k_1,\,k_2}\}_{k_1,\,k_2 \in \mathbb{Z}}$ uniquely extends to a bounded linear operator from $H^p_w(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$ to $L^q_{w,\,\mathcal{H}}(\mathbb{R}^n \times \mathbb{R}^m)$.

We also need the boundedness result for the anisotropic Littlewood-Paley g-function. The proof is similar to that of the anisotropic Lusin-area function; see [4, Theorem 3.2]. We omit the details.

Lemma 4.4. *Let $\varphi \in \mathcal{S}_0(\mathbb{R}^n)$, $p \in (1,\ \infty)$, and $w \in \mathcal{A}_p(\mathbb{R}^n;\ A)$. Then the Littlewood-Paley g-function, which is given by $g_\varphi(f)(x) := \{\sum_{k \in \mathbb{Z}} |f * \varphi_k(x)|^2\}^{1/2}$, is bounded on $L^p_w(\mathbb{R}^n)$.*

Finally, the isotropic and unweighted versions of the following lemma have appeared in several product settings; see, for example, [21, Theorem 4.3] and the proof of [20, Proposition 4]. In particular, Lemma 4.5 can be deduced from the proof of [4, Theorem 5.2] as indicated below.

Lemma 4.5. *Let $w \in \mathcal{A}_\infty(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$ and $p \in (0, 1]$. If $f \in L^2(\mathbb{R}^n \times \mathbb{R}^m) \cap H^p_w(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$, then $f \in L^p_w(\mathbb{R}^n \times \mathbb{R}^m)$. Moreover, there exists a positive constant C_p, independent of f, such that $\|f\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \leqslant C_p\|f\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\, \vec{A})}$.*

Proof. Let $w \in \mathcal{A}_\infty(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$, $p \in (0, 1]$ and $f \in H^p_w(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m)$. By an argument similar to the proof of [21, Theorem 4.3] or [20, Proposition 4], we shall prove that the atomic decomposition of f converges in $L^2(\mathbb{R}^n \times \mathbb{R}^m)$ and thus is pointwise almost everywhere.

Let $\psi := \psi^{(1)} \otimes \psi^{(2)}$ be as in Proposition 2.5. For any $k \in \mathbb{Z}$, let $\Omega_k := \{x \in \mathbb{R}^n \times \mathbb{R}^m : \vec{S}_\psi(f)(x) > 2^k\}$, $\lambda_k := 2^k[w(\Omega_k)]^{1/p}$, and

$$a_k := \lambda_k^{-1} \sum_{P\in m(\widetilde{\Omega}_k)} \sum_{R\in\mathcal{R}_k,\, R^*=P} e_R.$$

Here, our notation is the same as in [4, Lemma 4.6]. Since $f \in L^2(\mathbb{R}^n \times \mathbb{R}^m)$, by [4, Lemma 2.15 and (4.8)], we have that $f = \sum_{k\in\mathbb{Z}} \lambda_k a_k$ holds in $L^2(\mathbb{R}^n \times \mathbb{R}^m)$, and hence also almost everywhere. From this, $\operatorname{supp} a_k \subset \widetilde{\Omega}_k'''$ with $w(\widetilde{\Omega}_k''') \lesssim w(\Omega_k)$ (see [4, (6.5)]), $q \in [2, \infty) \cap (q_w, \infty)$, Hölder's inequality, and the size condition of a_k, it follows that

$$\begin{aligned}
\|f\|^p_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} &\lesssim \sum_{k\in\mathbb{Z}} \lambda_k^p \int_{\widetilde{\Omega}_k'''} |a_k(x)|^p\, w(x)dx \\
&\lesssim \sum_{k\in\mathbb{Z}} \lambda_k^p \|a_k\|^p_{L^q_w(\mathbb{R}^n\times\mathbb{R}^m)} [w(\widetilde{\Omega}_k''')]^{1-p/q} \\
&\lesssim \sum_{k\in\mathbb{Z}} 2^{kp} w(\Omega_k) \lesssim \|\vec{S}_\psi(f)\|^p_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \sim \|f\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\,\vec{A})},
\end{aligned}$$

which completes the proof of Lemma 4.5.

Proof of Theorem 1.6. Let T be a PASIO of order (s_1+1, s_2+1) with kernel K as in Definition 1.3. By the assumption (1.1) which says that $s_i > (q_w/p-1)\log_{|\lambda_{i,1}|} b_i$ for $i = 1, 2$, we can choose $1 < \lambda_{i,-} < |\lambda_{i,1}|$ close to $|\lambda_{i,1}|$, $r \in (q_w, \infty)$ close to q_w such that

$$\eta_i := p[s_i\zeta_{i,-} + 1] - r > 0, \quad i = 1, 2. \tag{4.1}$$

Let $q > \max\{2, r\}$. Then, $(p, q, \vec{s})_w$ is an admissible triplet, where $\vec{s} := (s_1 - 1, s_2 - 1)$.

Let $\psi := \psi^{(1)} \otimes \psi^{(2)}$ and $\phi := \phi^{(1)} \otimes \phi^{(2)}$ be as in Proposition 2.5. By part (iv) of this proposition we have $\psi = \phi * \phi$. Hence, by Theorem 1.5 and Definition 2.7, $T(f)$ is well defined for any $f \in L^2_w(\mathbb{R}^n \times \mathbb{R}^m) \cap H^p_w(\mathbb{R}^n \times \mathbb{R}^m)$ and

$$\begin{aligned}
\|Tf\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m)} &= \|\vec{S}_{\phi*\phi}(Tf)\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \\
&= \|\{\phi_{k_1,k_2} * \phi_{k_1,k_2} * [T(f)]\}_{k_1,k_2\in\mathbb{Z}}\|_{L^p_{w,\mathcal{H}}(\mathbb{R}^n\times\mathbb{R}^m)}.
\end{aligned}$$

To obtain the boundedness of T on $H^p_w(\mathbb{R}^n \times \mathbb{R}^m)$, by Lemma 4.3 and the density of $L^2_w(\mathbb{R}^n \times \mathbb{R}^m)$ $\cap H^p_w(\mathbb{R}^n \times \mathbb{R}^m)$ in $H^p_w(\mathbb{R}^n \times \mathbb{R}^m)$ given by [4, Theorem 5.1(i)], it suffices to prove that for all rectangular $(p, q, \vec{s})_w$-atoms a associated with some $R \in \mathcal{R}$ and all $\gamma \in \mathbb{Z}_+$,

$$\int_{(R_{1,\gamma}\times R_{2,\gamma})^{\complement}} |\{\phi_{k_1,k_2} * \phi_{k_1,k_2} * [T(a)](x)\}_{k_1,k_2\in\mathbb{Z}}|^p_{\mathcal{H}} w(x)\, dx \lesssim \max\{b_1^{-\eta_1\gamma}, b_2^{-\eta_2\gamma}\}, \tag{4.2}$$

where η_i is as in (4.1) and $R_{i,\gamma} := x_{R_i} + B^{(i)}_{v_i(\ell(R_i)-1)+u_i+5\sigma_i+\gamma}$ for $i = 1, 2$.

The left-hand side of (4.2) is less than

$$\begin{aligned}
&\left\{ \int_{R^{\complement}_{1,\gamma}\times R^{\complement}_{2,0}} + \int_{R^{\complement}_{1,\gamma}\times R_{2,0}} + \int_{R_{1,0}\times R^{\complement}_{2,\gamma}} + \int_{R^{\complement}_{1,0}\times R^{\complement}_{2,\gamma}} \right\} \\
&\qquad \times |\{\phi_{k_1,k_2} * \phi_{k_1,k_2} * [T(a)](x)\}_{k_1,k_2\in\mathbb{Z}}|^p_{\mathcal{H}} w(x)\, dx =: \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3 + \mathrm{I}_4.
\end{aligned} \tag{4.3}$$

We only derive the estimate for I_2, since the estimates for the other three items are similar.

Let N_1, N_2 be as in Definition 1.3. Since $\phi^{(i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$ for $i = 1, 2$, then by Lemma 3.1, we obtain that $\phi^{(i)} = \sum_{j_i=0}^{\infty} b_i^{-3j_i} \phi^{(i, j_i)}$, where $\phi^{(i, j_i)} \in \mathcal{S}_0(\mathbb{R}^{n_i})$ is a constant multiple of an $(s_i + N_i + 1)$-normalized bump function associated with $B_{j_i}^{(i)}$. For $j_1, j_2 \in \mathbb{Z}_+$, let $\phi^{\{j_1, j_2\}} := \phi^{(1, j_1)} \otimes \phi^{(2, j_2)}$. Thus, by $\phi = \phi^{(1)} \otimes \phi^{(2)}$, we have

$$\phi * \phi = \sum_{j_1, j_2, \ell_1 \in \mathbb{Z}_+} b_1^{-3(j_1+\ell_1)} b_2^{-3j_2} \phi^{\{j_1, j_2\}} * (\phi^{(1, \ell_1)} \otimes \phi^{(2)}). \tag{4.4}$$

Moreover, by Theorem 1.5 and a density argument, we obtain

$$\phi^{\{j_1, j_2\}} * (\phi^{(1, \ell_1)} \otimes \phi^{(2)}) * [T(a)] = K * [(\phi^{(1, j_1)} *_1 \phi^{(1, \ell_1)}) \otimes \phi^{(2, j_2)}] * (a *_2 \phi^{(2)}), \tag{4.5}$$

where $*_i$ denotes the convolution on $\mathbb{R}^{n_i}$, $i = 1, 2$. In fact, if $a \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$, then the above equality holds. For the rectangular $(p, q, \vec{s})$-atom a, let $\{a_k\}_{k \in \mathbb{N}} \subset \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$ be a sequence of functions approximating to a in $L_w^q(\mathbb{R}^n \times \mathbb{R}^m)$. Noticing that $T(a_k) \to Ta$ in $L_w^q(\mathbb{R}^n \times \mathbb{R}^m)$, we have (4.5).

For $k_1, k_2 \in \mathbb{Z}$ and $j_1, j_2, \ell_1 \in \mathbb{Z}_+$, let $K_{k_1, k_2}^{j_1, j_2, \ell_1} := K * [(\phi^{(1, j_1)} *_1 \phi^{(1, \ell_1)})_{k_1} \otimes \phi_{k_2}^{(2, j_2)}]$. By $w \in \mathcal{A}_r(\mathbb{R}^n \times \mathbb{R}^m; \vec{A})$, [4, Proposition 2.2(i)] and Lemma 4.1(iv), we have $w(R_{1, t_1+\gamma+1} \times R_{2, 0}) \lesssim b_1^{r(\gamma+t_1)} w(R)$. From this, (4.4), (4.5), Minkowski's inequality and Hölder's inequality, it follows that

$$\begin{aligned} \mathrm{I}_2 \lesssim & \sum_{j_1, j_2, \ell_1 \in \mathbb{Z}_+} b_1^{-3p(j_1+\ell_1)} b_2^{-3p\ell_2} \sum_{t_1 \in \mathbb{Z}_+} \bigg\{ \int_{(R_{1, \gamma+t_1+1} \setminus R_{1, \gamma+t_1}) \times R_{2, 0}} \\ & \times |\{K_{k_1, k_2}^{j_1, j_2, \ell_1} * (a *_2 \phi_{k_2}^{(2)})(x)\}_{k_1, k_2 \in \mathbb{Z}}|_{\mathcal{H}}^r w(x)\, dx \bigg\}^{p/r} b_1^{(r-p)(\gamma+t_1)} [w(R)]^{1-p/r}. \end{aligned} \tag{4.6}$$

Let $\widetilde{\ell}_1 := v_1[\ell(R_1) - 1] + u_1 + \gamma + t_1 + 5\sigma_1$. Write

$$\begin{aligned} & |\{K_{k_1, k_2}^{j_1, j_2, \ell_1} * (a *_2 \phi_{k_2}^{(2)})(x)\}_{k_1, k_2 \in \mathbb{Z}}|_{\mathcal{H}}^2 \\ & = \bigg[\sum_{\substack{k_1 < \widetilde{\ell}_1 - j_1 - \ell_1 - 4\sigma_1 \\ k_2 \in \mathbb{Z}}} + \sum_{\substack{k_1 \geqslant \widetilde{\ell}_1 - j_1 - \ell_1 - 4\sigma_1 \\ k_2 \in \mathbb{Z}}} \bigg] \\ & \quad \times b_1^{-k_1} b_2^{-k_2} \int_{B_{k_1}^{(1)}} \int_{B_{k_2}^{(2)}} |K_{k_1, k_2}^{j_1, j_2, \ell_1} * (a *_2 \phi_{k_2}^{(2)})(x - y)|^2\, dy \\ & =: [\mathrm{V}_1(x)]^2 + [\mathrm{V}_2(x)]^2. \end{aligned}$$

We only estimate V_1, since the estimate for V_2 is similar.

For $x \in (R_{1, \gamma+t_1+1} \setminus R_{1, \gamma+t_1}) \times R_{2, 0}$, $y \in B_{k_1}^{(1)} \times B_{k_2}^{(2)}$ and $z \in \mathbb{R}^n \times \mathbb{R}^m$, let

$$\widetilde{K}_{k_1, k_2}^{j_1, j_2, \ell_1}(z_1, z_2) := K_{k_1, k_2}^{j_1, j_2, \ell_1}(x_1 - y_1 - A_1^{\widetilde{\ell}_1} z_1, z_2).$$

For any $\widetilde{y}, \breve{y} \in \mathbb{R}^n \times \mathbb{R}^m$, by Taylor's formula with integral remainder, we have

$$\begin{aligned} \widetilde{K}_{k_1, k_2}^{j_1, j_2, \ell_1}(\widetilde{y}_1, \widetilde{y}_2) = & \sum_{j_1=0}^{s_1-1} \sum_{|\alpha_1|=j_1} (\widetilde{y}_1 - \breve{y}_1)^{\alpha_1} \partial_1^{\alpha_1} \widetilde{K}_{k_1, k_2}^{j_1, j_2, \ell_1}(\breve{y}_1, \widetilde{y}_2) \\ & + \sum_{|\alpha_1|=s_1} \int_0^1 (\widetilde{y}_1 - \breve{y}_1)^{\alpha_1} \partial_1^{\alpha_1} \widetilde{K}_{k_1, k_2}^{j_1, j_2, \ell_1}(\breve{y}_1 + r_1(\widetilde{y}_1 - \breve{y}_1), \widetilde{y}_2) \frac{(1 - r_1)^{s_1-1}}{s_1!}\, dr_1. \end{aligned}$$

Let $\breve{y}_1 := A_1^{-\widetilde{\ell}_1} x_{R_1}$ and $\widetilde{y}_1 := A_1^{-\widetilde{\ell}_1} z_1$. By $\operatorname{supp} a \subset R''$ and the vanishing condition of a up to order $s_1 - 1$, we then have

$$\begin{aligned} & K_{k_1, k_2}^{j_1, j_2, \ell_1} * (a *_2 \phi_{k_2}^{(2)})(x - y) \\ & \quad = \int_{\mathbb{R}^n \times \mathbb{R}^m} \widetilde{K}_{k_1, k_2}^{j_1, j_2, \ell_1}(A_1^{-\widetilde{\ell}_1} z_1, x_2 - y_2 - z_2)(a *_2 \phi_{k_2}^{(2)})(z)\, dz \end{aligned}$$

$$
\begin{aligned}
&= \sum_{|\alpha_1|=s_1} \int_0^1 \int_{R_1''\times\mathbb{R}^m} (a *_2 \phi_{k_2}^{(2)})(z)(A_1^{-\widetilde{\ell}_1}(z_1 - x_{R_1}))^{\alpha_1} \frac{(1+r_1)^{s_1-1}}{s_1!} \\
&\quad \times \partial_1^{\alpha_1} \widetilde{K}_{k_1,\,k_2}^{j_1,\,j_2,\,\ell_1}(A_1^{-\widetilde{\ell}_1}[x_{R_1} + (1-r_1)(x_{R_1} - z_1)],\, x_2 - y_2 - z_2)\, dz\, dr_1. \qquad (4.7)
\end{aligned}
$$

Moreover, for $x_1 \in R_{1,\,\gamma+t_1+1} \setminus R_{1,\,\gamma+t_1}$, $z_1 \in R_1''$, $r_1 \in (0,\,1)$ and $\widetilde{\ell}_1 > k_1 + j_1 + \ell_1 + 4\sigma_1$, by (2.2), (2.3) and Lemma 4.1(iv), we have $\rho_1(x_1 - x_{R_1}) = b_1^{\widetilde{\ell}_1}$ and $\rho_1(z_1 - x_{R_1}) \leqslant b_1^{-2\sigma_1 - t_1 - \gamma - 1}\rho_1(x_1 - x_{R_1})$, which together with $k_1 < \widetilde{\ell}_1 - j_1 - \ell_1 - 4\sigma_1$ further means that

$$\rho_1(x_1 - x_{R_1} - (1-r_1)(x_{R_1} - z_1)) \sim b_1^{\widetilde{\ell}_1},$$

and that for $y_1 \in B_{k_1}^{(1)}$,

$$\rho_1(x_1 - x_{R_1} - (1-r_1)(x_{R_1} - z_1) - y_1) \leqslant b_1^{k_1}.$$

From this and Lemma 3.4, it follows that

$$
\begin{aligned}
&\left|\partial_1^{\alpha_1} \widetilde{K}_{k_1,\,k_2}^{j_1,\,j_2,\,\ell_1}(A_1^{-\widetilde{\ell}_1}[x_{R_1} + (1-r_1)(x_{R_1} - z_1)],\, x_2 - y_2 - z_2)\right| \\
&\quad = \left|\partial_1^{\alpha_1}[K_{k_1,\,k_2}^{j_1,\,j_2,\,\ell_1}(A_1^{\widetilde{\ell}_1}\cdot,\, x_2 - y_2 - z_2)](A_1^{-\widetilde{\ell}_1}[x_1 - x_{R_1} - (1-r_1)(x_{R_1} - z_1) - y_1])\right| \\
&\quad \lesssim b_1^{(j_1+\ell_1)(1+\epsilon_1)+k_1\epsilon_1-\widetilde{\ell}_1(1+\epsilon_1)} \frac{b_2^{k_2\epsilon_2}}{\left[b_2^{k_2} + b_2^{-j_2}\rho_2(x_2 - z_2)\right]^{1+\epsilon_2}},
\end{aligned}
$$

where we used the fact that $b_2^{k_2} + b_2^{-j_2}\rho_2(x_2 - y_2 - z_2) \sim b_2^{k_2} + b_2^{-j_2}\rho_2(x_2 - z_2)$.

Furthermore, for $z_1 \in R_1''$, by [4, (2.6)], we have $|A_1^{-\widetilde{\ell}_1}(z_1 - x_{R_1})| \lesssim b_1^{-(\gamma+t_1)\zeta_{1,\,-}}$. Thus, for $x \in (R_{1,\,\gamma+t_1+1} \setminus R_{1,\,\gamma+t_1}) \times R_{2,\,0}$, by the above two estimates, (4.7), $\widetilde{\ell}_1 = v_1[\ell(R_1) - 1] + u_1 + \gamma + t_1 + 5\sigma_1$, a similar proof to that of (3.9), and Minkowski's inequality, we obtain

$$
\begin{aligned}
\mathrm{V}_1(x) &\lesssim \Bigg\{ b_1^{2(j_1+\ell_1)(1+\epsilon_1)-2\widetilde{\ell}_1(1+\epsilon_1)} \sum_{\substack{k_1 < \widetilde{\ell}_1 - j_1 - \ell_1 - 4\sigma_1 \\ k_2 \in \mathbb{Z}}} b_1^{2k_1\epsilon_1} b_2^{j_2(1+\epsilon_2)} \\
&\qquad \times \left[\int_{R_1''} b_1^{-(\gamma+t_1)s_1\zeta_{1,\,-}} \mathcal{M}^{(2)}(a(z_1,\,\cdot) *_2 \phi_{k_2}^{(2)})(x_2)\, dz_1\right]^2 \Bigg\}^{1/2} \\
&\lesssim b_1^{(j_1+\ell_1)-(t_1+\gamma)s_1\zeta_{1,\,-}} b_2^{j_2(1+\epsilon_2)} \\
&\qquad \times \frac{1}{b_2^{v_1\ell(R_1)+t_1+\gamma}} \int_{R_{1,\,\gamma+t_1+1}} \left\{\sum_{k_2\in\mathbb{Z}} |\mathcal{M}^{(2)}(a(z_1,\,\cdot) *_2 \phi_{k_2}^{(2)})(x_2)|^2\right\}^{1/2} dz_1 \\
&\lesssim b_1^{j_1+\ell_1-(t_1+\gamma)s_1\zeta_{1,\,-}} b_2^{j_2(1+\epsilon_2)} \\
&\qquad \times \mathcal{M}^{(1)}\left(\left\{\sum_{k_2\in\mathbb{Z}} [\mathcal{M}^{(2)}(a *_2 \phi_{k_2}^{(2)})(x_2)]^2\right\}^{1/2}\right)(x_1),
\end{aligned}
$$

where, and in what follows, $\mathcal{M}^{(i)}$ denotes the Hardy-Littlewood maximal function on $\mathbb{R}^{n_i}$, $i = 1,\, 2$.

Then, by the above estimate of $\mathrm{V}_1(x)$, the $L^r_{w(\cdot,\,x_2)}(\mathbb{R}^n)$-boundedness of $\mathcal{M}^{(1)}$ for all $x_2 \in \mathbb{R}^m$, the weighted vector-valued inequality for the Hardy-Littlewood maximal operator $\mathcal{M}^{(2)}$ with $w(x_1,\,\cdot) \in \mathcal{A}_r(\mathbb{R}^m;\, A_2)$ for all $x_1 \in \mathbb{R}^n$ (see [2, Theorem 2.5]), Lemma 4.4 with $g_{\phi^{(2)}}$, $\operatorname{supp} a \subset R''$, $r > q > 1$, Hölder's inequality, and the size condition of a, we have

$$
\begin{aligned}
&\left\{\int_{(R_{1,\,\gamma+t_1+1}\setminus R_{1,\,\gamma+t_1})\times R_{2,\,0}} [\mathrm{V}_1(x)]^r\, w(x)\, dx\right\}^{1/r} \\
&\quad \lesssim b_1^{j_1+\ell_1-(t_1+\gamma)s_1\zeta_{1,\,-}} b_2^{j_2(1+\epsilon_2)} \|g_{\phi^{(2)}}(a)\|_{L^r_w(\mathbb{R}^n\times\mathbb{R}^m)} \\
&\quad \lesssim b_1^{j_1+\ell_1-(t_1+\gamma)s_1\zeta_{1,\,-}} b_2^{j_2(1+\epsilon_2)} \|a\|_{L^q_w(\mathbb{R}^n\times\mathbb{R}^m)} [w(R'')]^{1/r-1/q} \\
&\quad \lesssim b_1^{j_1+\ell_1-(t_1+\gamma)s_1\zeta_{1,\,-}} b_2^{j_2(1+\epsilon_2)} [w(R)]^{1/r-1/p}.
\end{aligned}
$$

From this and $\eta_1 = p[(s_1+1)\zeta_{1,-}+1] - r > 0$, it follows that

$$\begin{aligned}
&\sum_{t_1\in\mathbb{Z}_+}\left\{\int_{(R_{1,\gamma+t_1+1}\setminus R_{1,\gamma+t_1})\times R_{2,0}}[\mathrm{V}_1(x)]^r w(x)\,dx\right\}^{p/r} b_1^{(r-p)(\gamma+t_1)}[w(R)]^{1-p/r}\\
&\quad\lesssim \sum_{t_1\in\mathbb{Z}_+} b_1^{p(j_1+\ell_1)} b_2^{pj_2(1+\epsilon_2)}[w(R)]^{p/r-1} b_1^{-p(t_1+\gamma)s_1\zeta_{1,-}} b_1^{(t_1+\gamma)(r-p)}[w(R)]^{1-p/r}\\
&\quad\lesssim b_1^{p(j_1+\ell_1)-\gamma\eta_1} b_2^{pj_2(1+\epsilon_2)},
\end{aligned}$$

which together with (4.6) yields that $\mathrm{I}_2 \lesssim b_1^{-\gamma\eta_1}$.

By an estimate similar to that of I_2, we also have $\mathrm{I}_1+\mathrm{I}_3+\mathrm{I}_4 \lesssim \max\{b_1^{-\gamma\eta_1}, b_2^{-\gamma\eta_2}\}$, where η_1 and η_2 are as in (4.1). By this and (4.3), we obtain (4.2) and hence the boundedness of T on $H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$.

Finally, let us prove that T is bounded from $H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ to $L^p_w(\mathbb{R}^n\times\mathbb{R}^m)$ with $p\in(0,1]$ satisfying (1.1) by borrowing some ideas from the proof of [21, Theorem 1.11]. Assume that $f\in L^2(\mathbb{R}^n\times\mathbb{R}^m)\cap H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$. By Theorem 1.5, Lemma 4.5 and the boundedness of T on $H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$, we obtain that

$$Tf \in L^p_w(\mathbb{R}^n\times\mathbb{R}^m)\cap H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})\cap L^2(\mathbb{R}^n\times\mathbb{R}^m)$$

and

$$\|Tf\|_{L^p_w(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim \|Tf\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})} \lesssim \|f\|_{H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})}.$$

This together with the density of $L^2(\mathbb{R}^n\times\mathbb{R}^m)\cap H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ in $H^p_w(\mathbb{R}^n\times\mathbb{R}^m;\vec{A})$ given by [4, Theorem 5.1(i)] implies that T extends to a linear bounded operator from $H^p_w(\mathbb{R}^n\times\mathbb{R}^m)$ to $L^p_w(\mathbb{R}^n\times\mathbb{R}^m)$. This finishes the proof of Theorem 1.6.

Acknowledgements This work was partially supported by Start-up Funding Doctor of Xinjiang University (Grant No. BS090109), National Science Foundation of US (Grant No. DMS 0653881) and National Natural Science Foundation of China (Grant No. 10871025).

References

1 Bownik M. Anisotropic Hardy spaces and wavelets. Mem Amer Math Soc, 2003, 164: 1–122

2 Bownik M, Ho K P. Atomic and molecular decompositions of anisotropic Triebel-Lizorkin spaces. Trans Amer Math Soc, 2006, 358: 1469–1510

3 Bownik M, Li B, Yang D, et al. Weighted anisotropic Hardy spaces and their applications in boundedness of sublinear operators. Indiana Univ Math J, 2008, 57: 3065–3100

4 Bownik M, Li B, Yang D, et al. Weighted anisotropic product Hardy spaces and boundedness of sublinear operators. Math Nachr, 2010, 283: 392–442

5 Calderón A P, Torchinsky A. Parabolic maximal functions associated with a distribution. Adv Math, 1975, 16: 1–64

6 Calderón A P, Torchinsky A. Parabolic maximal functions associated with a distribution. II. Adv Math, 1977, 24: 101–171

7 Chang D C, Sadosky C. Functions of bounded mean oscillation. Taiwanese J Math, 2006, 10: 573–601

8 Chang D C, Yang D, Zhou Y. Boundedness of sublinear operators on product Hardy spaces and its application. J Math Soc Japan, 2010, 62: 321–353

9 Christ M. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. Colloq Math, 1990, 60/61: 601–628

10 Coifman R R, Weiss G. Extensions of Hardy spaces and their use in analysis. Bull Amer Math Soc, 1977, 83: 569–645

11 Fefferman R. Harmonic analysis on product spaces. Ann of Math (2), 1987, 126: 109–130

12 Fefferman R. A^p weights and singular integrals. Amer J Math, 1988, 110: 975–987

13 Fefferman C, Stein E M. H^p spaces of several variables. Acta Math, 1972, 129: 137–193

14 Fefferman R, Stein E M. Singular integrals on product spaces. Adv Math, 1982, 45: 117–143

15 Folland G B, Stein E M. Hardy Spaces on Homogeneous Group, Mathematical Notes. Princeton, NJ: Princeton University Press, 1982

16 García-Cuerva J, Rubio de Francia J L. Weighted Norm Inequalities and Related Topics. Amsterdam: North-Holland Publishing Co, 1985

17 Grafakos L. Classical Fourier Analysis. New York: Springer Press, 2008

18 Grafakos L. Modern Fourier Analysis. New York: Springer Press, 2008

19 Gundy R F, Stein E M. H^p theory for the poly-disc. Proc Nat Acad Sci USA, 1979, 76: 1026–1029

20 Han Y, Lee M Y, Lin C C, et al. Calderón-Zygmund operators on product Hardy spaces. J Funct Anal, 2010, 258: 2834–2861

21 Han Y, Lu G. Discrete Littlewood-Paley-Stein theory and multi-parameter Hardy spaces associated with flag singular integrals. arXiv: 0801.1701 (available at http://arxiv.org/abs/0801.1701)

22 Han Y, Yang D. H^p boundedness of Calderón-Zygmund operators on product spaces. Math Z, 2005, 249: 869–881

23 Han Y, Yang D. Boundedness of Calderón-Zygmund operators in product Hardy spaces. Appl Math J Chinese Univ Ser B, 2009, 24: 321–335

24 Haroske D D, Tamási E. Wavelet frames for distributions in anisotropic Besov spaces. Georgian Math J, 2005, 12: 637–658

25 Müller D, Ricci F, Stein E M. Marcinkiewicz multipliers and multi-parameter structure on Heisenberg(-type) groups. I. Invent Math, 1995, 119: 199–233

26 Nagel A, Stein E M. On the product theory of singular integrals. Rev Mat Iberoamericana, 2004, 20: 531–561

27 Nagel A, Stein E M. The $\overline{\partial}_b$-complex on decoupled boundaries in $\mathbb{C}^n$. Ann of Math (2), 2006, 164: 649–713

28 Sato S. Weighted inequalities on product domains. Studia Math, 1989, 92: 59–72

29 Schmeisser H J, Triebel H. Topics in Fourier Analysis and Function Spaces. Leipzig: Akademische Verlagsgesellschaft Geest & Portig K G, 1987

30 Stein E M. Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton, NJ: Princeton University Press, 1993

31 Strömberg J O, Torchinsky A. Weighted Hardy Spaces. Berlin: Springer-Verlag, 1989

32 Triebel H. Fractals and Spectra. Related to Fourier Analysis and Function Spaces. Basel: Birkhäuser Verlag, 1997

33 Vybíral J. Function spaces with dominating mixed smoothness. Dissertationes Math (Rozprawy Mat), 2006, 436: 1–73