Representation of Fourier Integral Operators using Shearlets

Kanghui Guo and Demetrio Labate

December 18, 2006; revised June 16, 2007; revised August 6, 2007

Abstract

Traditional methods of time-frequency and multiscale analysis, such as wavelets and Gabor frames, have been successfully employed for representing most classes of pseudodifferential operators. However these methods are not equally effective in dealing with Fourier Integral Operators in general. In this paper, we show that the shearlets, recently introduced by the authors and their collaborators, provide very efficient representations for a large class of Fourier Integral Operators. The shearlets are an affine-like system of well-localized waveforms at various scales, locations and orientations, which are particularly efficient in representing anisotropic functions. Using this approach, we prove that the matrix representation of a Fourier Integral Operator with respect to a Parseval frame of sherlets is sparse and well-organized. This fact recovers a similar result recently obtained by Candès and Demanet using curvelets, which illustrates the benefits of directional multiscale representations (such as curvelets and shearlets) in the study of those functions and operators where traditional multiscale methods are unable to provide the appropriate geometric analysis in the phase space.

1 Introduction

Fourier Integral Operators were originally introduced by Lax [17], with the construction of parametrics in the Cauchy problem for strongly hyperbolic equations. Since then they have become an important tool in a variety of problems arising in partial differential equations. For example, consider the hyperbolic equation

$$\frac{\partial}{\partial t}\, u = i\, \lambda(t, x, D_x)\; u, \tag{1.1}$$

where $\lambda(t, x, D_x)$ is a smooth one-parameter family of pseudodifferential operators with symbol $\lambda(t, x, \xi) \sim \lambda_1(t, x, \xi) + \lambda_0(t, x, \xi)$, where $\lambda_j(t, x, \xi)$ is homogeneous of degree j in ξ and $\lambda_1(t, x, \xi)$ is real valued (which makes the equation hyperbolic). The solution operator $S(t, s)$ to (1.1) taking $u(s)$ to $u(t)$ is essentially (modulo a smoothing operator) given by an integral of the form

$$Tu(t) = \int e^{2\pi i \Phi(t,x,\xi)}\, \sigma(t, x, \xi)\, \hat{f}(\xi)\, d\xi,$$

where the *phase* $\Phi(t, x, \xi)$ and the *amplitude* $\sigma(t, x, \xi)$ are C^∞ functions. In particular, if $Tu(t) = S(t, s)\, u(s)$ is the solution to (1.1), then, for small t, $\Phi(t, x, \xi)$ is close to $\xi\, x$ and Φ solves the eikonal equation

$$\frac{\partial}{\partial t}\, \Phi = \lambda_1(t, x, \nabla_x \Phi), \quad \Phi(0, x, \xi) = \xi\, x.$$

It follows (see, for example, [23, Ch.8], [21]) that the *canonical transformation*

$$(\nabla_\xi \Phi(t, x, \xi), \xi) \;\to\; (x, \nabla_x \Phi(t, x, \xi))$$

[0] *Math Subject Classifications.* 42C15, 42C40.
[0] *Keywords and Phrases.* Curvelets, Fourier Integral Operators, Pseudo Differential Operators, Shearlets, Sparsity, Wavelets.
[0] *Acknowledgements.* The second author was supported in part by a National Science Foundation grant DMS 0604561.

defines a bijective map $C(t)$ for all t close to $t = 0$. Then one can show that $C(t)$ corresponds to the Hamiltonian flow (or bicharacteristic equation):

$$\begin{aligned} \dot{\xi}(t) &= \nabla_x \lambda_1(t, x(t), \xi(t)), \quad \xi(0) = \xi_0, \\ \dot{x}(t) &= -\nabla_\xi \lambda_1(t, x(t), \xi(t)), \quad x(0) = x_0. \end{aligned}$$

Traditional methods of time-frequency and multiscale analysis, such as wavelets and Gabor frames, have proven to be very effective in providing efficient representations for a large class of operators, including pseudodifferential operators and Calderón–Zygmund operators (see, for example, [18, 10]). Unfortunately, they fail to provide equally efficient representations for most Fourier Integral Operators. This paper is motivated by some remarkable recent results by H. Smith [19, 20] and E. Candès and L. Demanet [2] who have shown how to obtain sparse representations of Fourier Integral Operators by combining the methods of multiscale analysis with a notion of anisotropy and directionality. More specifically, in [2], a Fourier Integral Operator T is represented using a Parseval frame of *curvelets*. This is a multiscale system of functions $\{\phi_\mu : \mu \in \mathcal{M}\}$, at various scales, locations and directions, having anisotropic compact support in the frequency domain [3]. Then, denoting by $T(\mu, \mu')$ the matrix elements of T with respect to the curvelets, it is proven that, for each $N > 0$,

$$|T(\mu, \mu')| \leq C_N \, \omega(\mu, h(\mu'))^{-N}, \tag{1.2}$$

where ω is a certain distance and h is an index mapping. Let us recall that the ability to provide sparse representations of an operator is very significant for the theoretical analysis of the operator as well as its numerical implementation. In fact, sparse representations are useful to deduce sharp estimates [20] and can be exploited to design low complexity algorithms [1].

In this paper, we show that the shearlets, introduced by the authors and their collaborators [13, 15, 11] provide an alternative approach to the construction of sparse representations of Fourier Integral Operators. Shearlets are Parseval frames of well localized functions at various scales, locations and directions. Similarly to the curvelets, they provide optimally sparse representations for two-dimensional functions that are smooth away from discontinuities along curves [12] and have the ability to exactly characterize the wavefront set of distributions [16]. In addition, the shearlets have some distinctive features:

- They form an affine-like system. That is, they are generated from the action of translation and dilation operators on a single function.
- They are defined on a Cartesian grid.
- Thanks to their affine-like structure, they exhibit a group structure.

These properties do not hold for curvelets, and provide a number of advantages for some theoretical as well as numerical applications of such representations (see further discussions in [6, 7, 15]).

In this paper we, show that, by using the shearlets to represent a Fourier integral operator, one obtains an estimate on the matrix representation of the same type as (1.2). Thus, shearlets provide sparse representations of such operators. As we will describe later in the paper, the main outline of our proof follows the one in [2]. However, to prove our sparsity result, we had to introduce new ideas to deal with the distinctive mathematical structure of the shearlets. In particular, the notion of almost orthogonality of shearlet molecules, in Sec. 4.2, has to be adapted to the Cartesian geometry, and to the combination of both horizontal and vertical shearlets. For the analysis of the operator T2, in Sec. 4.3, since we do not have radial symmetry, we cannot use the Smith approach involving Fourier multipliers. Our argument introduces a new atomic decomposition. The methods developed in our proof are of independent interest, and show the potential of representation methods based on shearlets in the study of operators and functions spaces.

Shearlets are a special case of composite wavelets, introduced by the authors and their collaborators in [13, 14, 15]. This theory unifies and extends the theory of wavelet systems, and provides a general framework for constructing multiscle directional systems using the affine framework. Some of the benefits of this approach are the ability to extend this framework to any dimensions, the existence of a multiresolution analysis (MRA), and the ability to exploit a group structure. In particular, thanks to this group structure, there are function spaces naturally associated to composite wavelets, which are useful in variational methods for image processing. As an additional benefit, the shearlet decomposition of FIO indicates that the whole machinery of composite wavelets can be brought into play. For example, higher dimensional extensions can be easily obtained, and the MRA structure can be exploited.

The paper is organized as follows. In Section 2 we recall the construction of the shearlets. In Section 3 we defined the class of Fourier Integral Operators that will be discussed in this paper and present our main theorem. The various steps needed to prove the main theorem are presented in Section 4.

1.1 Notation and definitions

We adopt the convention that $x \in \mathbb{R}^n$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^n$ is a row vector, i.e., $\xi = (\xi_1, \dots, \xi_n)$. A vector x multiplying a matrix $M \in GL_n(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying M on the left is a row vector. Thus, $Mx \in \mathbb{R}^n$ and $\xi M \in \widehat{\mathbb{R}}^n$. The Fourier transform is defined as

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi x}\, dx,$$

where $\xi \in \widehat{\mathbb{R}}^n$, and the inverse Fourier transform is

$$\check{f}(x) = \int_{\widehat{\mathbb{R}}^n} f(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

For any $E \subset \widehat{\mathbb{R}}^n$, we denote by $L^2(E)^\vee$ the space $\{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset E\}$.

Recall that a countable collection $\{\psi_i\}_{i \in I}$ in a Hilbert space $\mathcal{H}$ is a *Parseval frame* (sometimes called a *normalized tight frame*) for $\mathcal{H}$ if

$$\sum_{i \in I} |\langle f, \psi_i \rangle|^2 = \|f\|^2, \qquad \text{for all } f \in \mathcal{H}.$$

This is equivalent to the reproducing formula $f = \sum_i \langle f, \psi_i \rangle\, \psi_i$, for all $f \in \mathcal{H}$, where the series converges in the norm of $\mathcal{H}$. This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. We refer the reader to [4, 5] for more details about frames.

2 Shearlets

The *shearlets* are elements of an affine collection of functions in $L^2(\mathbb{R}^2)$ of the form

$$\mathcal{A}_{AB}(\psi) = \{\psi_{j,\ell,k}(x) = |\det A|^{j/2}\, \psi(B^\ell A^j x - k) : (j, \ell, k) \in \mathcal{M}\}, \tag{2.3}$$

where $\mathcal{M} = \{j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}$, $\psi \in L^2(\mathbb{R}^2)$ and A, $B \in GL_n(\mathbb{R})$. It will be more convenient to define the function ψ and examine the shearlets in the frequency domain. After computing the Fourier transform, the functons $\psi_{j,\ell,k}$, given by (2.3), have the form

$$\hat{\psi}_{j,\ell,k}(\xi) = |\det A|^{-j/2}\, \hat{\psi}(\xi A^{-j} B^{-\ell})\, e^{-2\pi i \xi A^{-j} B^{-\ell} k}.$$

For any $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$, $\xi_1 \neq 0$, let ψ be given by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where $\hat\psi_1, \hat\psi_2 \in C^\infty(\widehat{\mathbb{R}})$, $\operatorname{supp}\hat\psi_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and $\operatorname{supp}\hat\psi_2 \subset [-1, 1]$. We assume that

$$\sum_{j\geq 0} |\hat\psi_1(2^{-2j}\omega)|^2 = 1 \qquad \text{for } |\omega| \geq \frac{1}{8}, \tag{2.4}$$

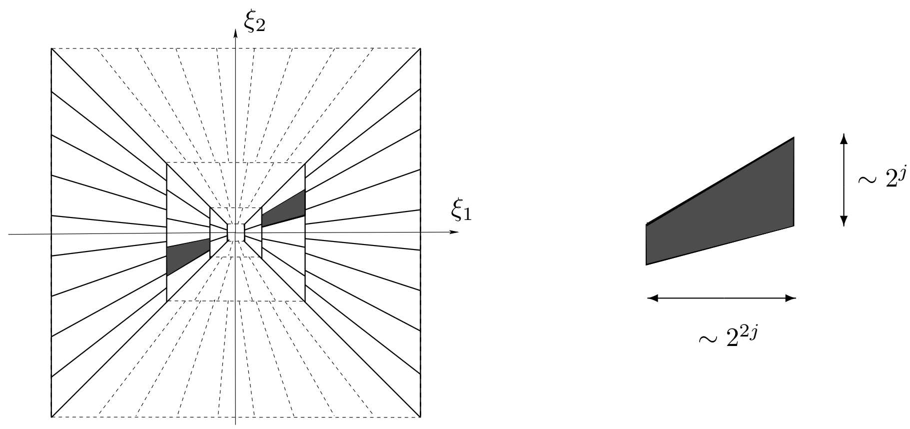
and

$$|\hat\psi_2(\omega-1)|^2 + |\hat\psi_2(\omega)|^2 + |\hat\psi_2(\omega+1)|^2 = 1 \qquad \text{for } |\omega| \leq 1.$$

There are several examples of functions ψ_1, ψ_2 satisfying these properties (see [12, 15]). It follows from the last equation that, for any $j \geq 0$,

$$\sum_{\ell=-2^j}^{2^j} |\hat\psi_2(2^j\,\omega + \ell)|^2 = 1 \qquad \text{for } |\omega| \leq 1. \tag{2.5}$$

It also follows from our assumptions that $\hat\psi \in C_0^\infty(\widehat{\mathbb{R}}^2)$, with $\operatorname{supp}\hat\psi \subset [-\frac{1}{2}, \frac{1}{2}]^2$.



(a) (b)

Figure 1: *(a) The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by the shearlets. (b) Frequency support of the shearlet $\psi_{j,\ell,k}$, for $\xi_1 > 0$. The other half of the support, for $\xi_1 < 0$, is symmetrical.*

Let A and B be the matrices given by

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Observe that $(\xi_1, \xi_2)\, A^{-j} B^{-\ell} = (2^{-2j}\xi_1, -\ell 2^{-2j}\xi_1 + 2^{-j}\xi_2)$. Thus, using (2.4) and (2.5) we have that:

$$\begin{aligned} \sum_{j\geq 0}\sum_{\ell=-2^j}^{2^j} |\hat\psi(\xi\, A^{-j}B^{-\ell})|^2 &= \sum_{j\geq 0}\sum_{\ell=-2^j}^{2^j} |\hat\psi_1(2^{-2j}\,\xi_1)|^2\, |\hat\psi_2(2^j\,\frac{\xi_2}{\xi_1} - \ell)|^2 \\ &= \sum_{j\geq 0} |\hat\psi_1(2^{-2j}\,\xi_1)|^2 \sum_{\ell=-2^j}^{2^j} |\hat\psi_2(2^j\,\frac{\xi_2}{\xi_1} + \ell)|^2 = 1, \end{aligned}$$

for $(\xi_1, \xi_2) \in \mathcal{D}_C$, where $\mathcal{D}_C = \{(\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \geq \frac{1}{8}, |\frac{\xi_2}{\xi_1}| \leq 1\}$. This equation describes the tiling of the region $\mathcal{D}_C$ in the frequency plane illustrated in Figure 1(a) (solid line). Since $\hat\psi$ is supported inside $[-\frac{1}{2}, \frac{1}{2}]^2$, this implies that the collection of *horizontal shearlets*:

$$\mathcal{S}(\psi) = \{\psi_{j,\ell,k}(x) = 2^{\frac{3j}{2}}\, \psi(B^\ell A^j x - k) : \, j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}, \tag{2.6}$$

is a Parseval frame for the subspace of L^2 defined by

$$L^2(\mathcal{D}_C)^\vee = \{f \in L^2(\mathbb{R}^2) : \, \operatorname{supp} \hat{f} \subset \mathcal{D}_C\}.$$

In order to obtain a Parseval frame for the whole space $L^2(\mathbb{R}^2)$, we can construct another Parseval frame of shearlets for the functions whose frequency support is contained vertical cone $\mathcal{D}_{\widetilde{C}} = \{(\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : \, |\xi_2| \geq \frac{1}{8}, |\frac{\xi_1}{\xi_2}| \leq 1\}$. Namely, let

$$A_{(v)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad B_{(v)} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

and $\psi^{(v)}$ be given by

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2 \left(\frac{\xi_1}{\xi_2}\right).$$

Then a similar computation shows that the collection of *vertical shearlets*:

$$\mathcal{S}(\psi^{(v)}) = \{\psi^{(v)}_{j,\ell,k}(x) = 2^{\frac{3j}{2}}\, \psi^{(v)}(B^\ell_{(v)} A^j_{(v)} x - k) : \, j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}, \tag{2.7}$$

is a Parseval frame for the subspace of $L^2(\mathbb{R}^2)$ given by $L^2(\widetilde{C})^\vee$. Finally, we can easily construct an orthonormal basis for $L^2([-\frac{1}{8}, \frac{1}{8}]^2)^\vee$ (for example, by using Fourier series). As a result, any function in $L^2(\mathbb{R}^2)$ can be expressed as the sum $f = P_C f + P_{\widetilde{C}} f + P_0 f$, where each component corresponds to the orthogonal projection of f into one of the 3 subspaces of $L^2(\mathbb{R}^2)$ described above. The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by this system is illustrated in Figure 1(a). Additional details about this construction can be found in [12, 11].

The conditions on the support of $\hat{\psi}_1$ and $\hat{\psi}_2$ imply that, for each j, ℓ, the functions $\hat{\psi}_{j,\ell,k}$ have frequency support contained in the sets:

$$W_{j,\ell} = \{(\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \, |\tfrac{\xi_2}{\xi_1} - \ell\, 2^{-j}| \leq 2^{-j}\}.$$

As shown in Figure 1, for each j, ℓ, the set $W_{j,\ell}$ is a pair of trapezoids centered about $\pm\xi_{j,\ell}$, where $\xi_{j,\ell} = 2^{-2}(2^{2j}, \ell 2^j)$. Each trapezoid is approximately contained in a box of size $2^{2j} \times 2^j$ in the frequency domain (see Figure 1(b)). Thus the frequency support of the shearlets becomes increasingly thin as j increases. In addition, since the function ψ is in C_0^∞, then the shearlets are well localized. In the spatial domain, they are centered at $k_{j,\ell} = A^{-j} B^{-\ell} k$, and oriented along the line $\frac{x_1}{x_2} = -\ell 2^{-j}$. Similar observations hold for the vertical shearlets $\hat{\psi}^{(v)}_{j,\ell,k}$. These properties are essential in the ability of shearlets to provide very efficient representations for functions containing distributed discontinuities. In fact, shearlets provide optimally sparse representations for a large class of functions with discontinuities along curves [12]. These properties will also play a crucial role in the arguments used to prove the main results of this paper.

In the following, we will use the notation $\{\psi_\mu : \mu \in \mathcal{M}\}$ and $\{\psi^{(v)}_\mu : \mu \in \mathcal{M}\}$ to denote the collection of shearlets (2.6) and (2.7), respectively, where μ stands for the multi-index (j, ℓ, k). Each shearlet ψ_μ has frequency support in the set $W_{j,\ell}$ and is associated to a location $(k_{j,\ell}, \pm\xi_{j,\ell})$ in the phase space. A direct calculation shows that the set $W_{j,\ell}$ has area $\frac{63}{128}\, 2^{3j}$. Similar observations hold for the vertical shearlets $\{\hat{\psi}^{(v)}_\mu : \mu \in \mathcal{M}\}$.

3 Shearlet Representation of Fourier Integral Operators

In this paper, we consider *Fourier Integral Operators* T of the form

$$Tf(x) = \int e^{2\pi i \Phi(x,\xi)}\, \sigma(x,\xi)\, \hat{f}(\xi)\, d\xi, \tag{3.8}$$

where the *phase function* $\Phi(x,\xi)$ and the *amplitude function* $\sigma(x,\xi)$ satisfy the following assumptions:

- $\Phi(x,\xi)$ is a real-valued function, C^∞ in (x,ξ), for $\xi \neq 0$, on the support of σ, and homogeneous of degree 1 in ξ; that is, $\Phi(x,\lambda\xi) = \lambda\,\Phi(x,\xi)$, for all $\lambda > 0$. In addition, we assume that

$$|\det \Phi_{x\xi}(x,\xi)| > c > 0, \tag{3.9}$$

  uniformly in x and ξ, where $\Phi_{x\xi} = \nabla_x\ \nabla_\xi \Phi$;

- σ is a standard symbol of order 0, that is, σ is in C^∞ and

$$|\partial_\xi^\alpha\, \partial_x^\beta \sigma(x,\xi)| \le C_{\alpha\beta}\,(1+|\xi|)^{-|\alpha|}; \tag{3.10}$$

  in addition, we assume that σ has compact support in the x variable.

Observe that, for each t fixed, the solution operator to the hyperbolic problem described in the Introduction is of the form (3.8). The above assumptions about Φ and σ are consistent with the study of Fourier Integral Operator associated with the solutions of hyperbolic problems (see [22, Ch.9]).

As described in the introduction, there is a notion of Hamiltonian flow associated with hyperbolic equations. The same notion also exists for Fourier Integral Operators and is encoded in the phase function Φ. Indeed the canonical transformation associated with the phase $\Phi(x,\xi)$ is the mapping $(x,\xi) \to (y,\eta)$ of the phase space:

$$y = \nabla_\xi \Phi(x,\xi), \quad \eta = \nabla_x \Phi(x,\xi).$$

As mentioned in the introduction, this formulation is equivalent to that involving trajectories along the Hamiltonian flow. This canonical transformation induces a bijective mapping on the indices of the shearlet denoted by $h(\mu)$, for $\mu \in \mathcal{M}$.

This is the main theorem of this paper:

Theorem 3.1. *Let T be an Fourier Integral Operator satisfying the assumptions given above, acting on functions on $\mathbb{R}^2$. For $\mu,\mu' \in \mathcal{M}$, let $T(\mu,\mu') = \langle T\psi_\mu, \psi_{\mu'}\rangle$, where ψ_μ and $\psi_{\mu'}$ are elements of the Parseval frame of shearlets $\{\psi_\mu : \mu \in \mathcal{M}\} \cup \{\psi_\mu^{(v)} : \mu \in \mathcal{M}\}$. Then, for each $N > 0$, there is a constant $C_N > 0$ such that*

$$|T(\mu,\mu')| \le C_N\ \omega(\mu, h_{\mu'}(\mu'))^{-N}.$$

In the statement of the theorem, the function ω is the dyadic parabolic pseudo-distance associated with the indices $\mu,\mu' \in \mathcal{M}$, whose definition will be given in Section 4.2. As mentioned above, for each $\mu' \in \mathcal{M}$, the function $h_{\mu'}$ is a bijective mapping on $\mathcal{M}$, induced by the canonical transformation associated with the phase Φ of T (see further discussion in Section 4.3). As we observed above, the Parseval frame of shearlets is made up of two collections: vertical and horizontal shearlets, given by (2.6) and (2.7), respectively. Theorem 3.1 holds for any combination of vertical and horizontal shearlets. Since the structure of the two systems is very similar, in the following we will analyze in detail the representation of the operator T with respect to the Parseval frame of horizontal shearlets only. We will point out the differences of the representation with respect to the vertical shearlets where this is needed (in particular, in Section 4.2).

As described in the Introduction, a result similar to Theorem 3.1 is obtained by Candès and Demanet in [2]. As in [2], using Schur's Lemma, it follows from Theorem 3.1 that, for every $0 < p \le \infty$, $T(\mu,\mu')$ is bounded from ℓ^p to ℓ^p.

We shall now start to examine the matrix representation of a Fourier Integral Operator T with respect to a Parseval frame of shearlets. For $j \geq 0$ and $|\ell| \leq 2^j$, let

$$B_{j,\ell} = \begin{pmatrix} 1 & -\ell\, 2^{-j} \\ 0 & 1, \end{pmatrix}$$

and define the sets

$$W_j = W_{j,0} = W_{j,\ell}\, B_{j,\ell} = \{(\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}],\ |\tfrac{\xi_2}{\xi_1}| \leq 2^{-j}\}$$

and

$$E = W_{j,\ell}\, B_{j,\ell}\, A^{-j} = W_j\, A^{-j} = \{(\xi_1, \xi_2) : \xi_1 \in [-2^{-1}, -2^{-4}] \cup [2^{-4}, 2^{-1}],\ |\tfrac{\xi_2}{\xi_1}| \leq 1\}.$$

That is, the shear matrix $B_{j,\ell}$ is mapping $W_{j,\ell}$ into another pair of trapezoids oriented along the ξ_1 axis, and $B_{j,\ell}\, A^{-j}$ is mapping $W_{j,\ell}$ into a pair of trapezoids inside the unit square $[-1/2, 1, 2]^2$. Also observe that

$$B_{j,\ell}\, A^{-j} = A^{-j}\, B^{-\ell}.$$

For μ a fixed index triple in $\mathcal{M}$, let ψ_μ be a shearlet with location (x_μ, ξ_μ) in the phase space. Then, using the notation we have introduced and the change of variables $\xi = \eta\, B_{j,\ell}^{-1}$ and $\omega = \eta A^{-j}$, we have

$$\begin{aligned} T\,\psi_\mu(x) &= \int e^{2\pi i \Phi(x,\xi)}\, \sigma(x,\xi)\, \hat{\psi}_\mu(\xi)\, d\xi \\ &= 2^{-3j/2} \int_{W_{j,\ell}} e^{2\pi i \left(\Phi(x,\xi) - \xi A^{-j} B^{-\ell} k\right)}\, \sigma(x,\xi)\, \hat{\psi}(\xi A^{-j} B^{-\ell})\, d\xi \\ &= 2^{-3j/2} \int_{W_j} e^{2\pi i \left(\Phi(x,\eta\, B_{j,\ell}^{-1}) - \eta\, B_{j,\ell}^{-1} A^{-j} B^{-\ell} k\right)}\, \sigma(x, \eta\, B_{j,\ell}^{-1})\, \hat{\psi}(\eta\, B_{j,\ell}^{-1} A^{-j} B^{-\ell})\, d\eta \\ &= 2^{-3j/2} \int_{W_j} e^{2\pi i \left(\Phi(x,\eta\, B_{j,\ell}^{-1}) - \eta\, A^{-j} k\right)}\, \sigma(x, \eta\, B_{j,\ell}^{-1})\, \hat{\psi}(\eta\, A^{-j})\, d\eta \\ &= 2^{3j/2} \int_{E} e^{2\pi i \left(\Phi(x,\omega\, B^{\ell} A^{j}) - \omega k\right)}\, \sigma(x, \omega\, B^{\ell} A^{j})\, \hat{\psi}(\omega)\, d\omega \end{aligned} \tag{3.11}$$

In order to proceed with the analysis of T, it will be convenient to locally linearize the phase $\Phi(x,\xi)$ to separate the nonlinearities of ξ from those of x. This is a standard approach in the study of of Fourier Integral Operators and can be found, for example, in [22, Ch.9].

We define

$$\delta_{j,\ell}(x,\eta) = \begin{cases} \Phi(x, \eta\, B_{j,\ell}^{-1}) - \eta\, B_{j,\ell}^{-1} \cdot \nabla_\xi \Phi\left(x, (1,0)\, B_{j,\ell}^{-1}\right), & \text{for } \eta \in W_j^+; \\ \Phi(x, \eta\, B_{j,\ell}^{-1}) - \eta\, B_{j,\ell}^{-1} \cdot \nabla_\xi \Phi\left(x, (-1,0)\, B_{j,\ell}^{-1}\right), & \text{for } \eta \in W_j^-, \end{cases} \tag{3.12}$$

where

$$W_j^+ = W_j \bigcap \{(\xi_1,\xi_2) :\ \xi \geq 0\}, \quad W_j^- = W_j \bigcap \{(\xi_1,\xi_2) :\ \xi < 0\}. \tag{3.13}$$

For μ fixed, this allows us to decompose T as

$$T = T_\mu^{(2)}\, T_\mu^{(1)},$$

where

$$\begin{aligned} T_\mu^{(1)} f(x) &= \int_{W_j} e^{2\pi i \eta\, B_{j,\ell}^{-1} x}\, \beta_\mu(x,\eta)\, \hat{f}(\eta\, B_{j,\ell}^{-1})\, d\eta, \\ T_\mu^{(2)} f(x) &= f(\phi_\mu(x)), \end{aligned}$$

with $\beta_\mu(x,\eta) = e^{2\pi i \delta_{j,\ell}(\phi_\mu^{-1}(x),\eta)}\,\sigma(\phi_\mu^{-1}(x), \eta\, B_{j,\ell}^{-1})$, and $\phi_\mu(x) = \nabla_\xi \Phi\left(x, (1,0)\, B_{j,\ell}^{-1}\right)$.

Observe that the operator $T_\mu^{(1)}$ obtained from this decomposition has linear phase but is not a 'standard' pseudodifferential operator. That is, $\beta_\mu(x,\eta)$ is not a standard class symbol in general. To illustrate this point, consider the following example. Let $\Phi(\xi) = |\xi|$, for $\xi = (\xi_1, \xi_2) \neq 0$, and $\xi_\mu = 2^j\, e_\mu$, where $e_\mu = (\cos\theta_\mu, \sin\theta_\mu)$. Then

$$\nabla_\xi \Phi(\xi_\mu) = \frac{\xi_\mu}{|\xi_\mu|} = e_\mu$$

and

$$\delta_\mu(\xi) = \Phi(\xi) - \nabla_\xi \Phi(\xi_\mu)\, \xi = \Phi(\xi) - e_\mu\, \xi.$$

For $\theta_\mu = 0$, then $e_\mu = (1,0)$ and

$$\delta_\mu(\xi) = |\xi| - \xi_1 = \sqrt{\xi_1^2 + \xi_2^2} - \xi_1,$$

which is unbounded. Notice however that $\delta_\mu(\xi)$ is a bounded function if we impose that the components of $\xi = (\xi_1, \xi_2)$ satisfy a parabolic scaling condition $\xi_1 = c\, \xi_2^2$, for some $c \in \mathbb{R}$. This is one reason why the shearlets and the curvelets are effective in dealing with the operator T (see Lemma 4.3). Also notice that the derivatives of $\delta_\mu(\xi)$ are homogeneous of degree 0 in ξ. Hence they exhibit no decay in ξ and, as a result, $\beta_\mu(x,\xi)$ does not satisfy (3.10) unless $\delta_\mu(x,\xi) = 0$.

The detailed analysis of the operators $T_\mu^{(1)}$ and $T_\mu^{(1)}$ and the proof of Theorem 3.1 will be given in the next section.

4 Proof of Main Theorem

In order to prove Theorem 3.1, we will now proceed with the analysis of the operator $T_\mu^{(1)}$ and $T_\mu^{(2)}$.

4.1 Analysis of the operator $T_\mu^{(1)}$

In this section, we show that the operator $T_\mu^{(1)}$ maps a shearlet ψ_μ into a shearlet-like function m_μ which, for analogy with similar notions in wavelet analysis, will be referred to as a *shearlet molecule*.

Definition 4.1. *For $\mu = (j,\ell,k) \in \mathcal{M}$, the function $m_\mu(x) = 2^{3j/2}\, a_\mu(B^\ell A^j x - k)$ is an horizontal shearlet molecule with regularity R if the α_μ satisfies the following properties:*

(i) for each $\gamma = (\gamma_1, \gamma_2) \in \mathbb{N} \times \mathbb{N}$ and each $N \geq 0$ there is a constant $C_N > 0$ such that

$$|\partial_x^\gamma a_\mu(x)| \leq C_N\, (1 + |x|)^{-N}; \tag{4.14}$$

(ii) for each $M \leq R$ and each $N \geq 0$ there is a constant $C_{N,M} > 0$ such that

$$|\hat{a}_\mu(\xi)| \leq C_{N,M}\, (1 + |\xi|)^{-N}\, (2^{-2j} + |\xi_1|)^M. \tag{4.15}$$

For $\mu = (j,\ell,k) \in \mathcal{M}$, the function $m_\mu^{(v)}(x) = 2^{3j/2}\, a_\mu(B_{(v)}^\ell A_{(v)}^j x - k)$ is a vertical shearlet molecule with regularity R if the $\alpha_\mu^{(v)}$ satisfies (4.14) and for each $M \leq R$ and each $N \geq 0$ there is a constant $C_{N,M} > 0$ such that

$$\left|\hat{a}_\mu^{(v)}(\xi)\right| \leq C_{N,M}\, (1 + |\xi|)^{-N}\, (2^{-2j} + |\xi_2|)^M.$$

The constants C_N and $C_{N,M}$ are independent of μ.

The second factor in the inequality (4.15) is associated with the almost vanishing moments. By this property, the frequency support of a shearlet molecule m_μ is mostly concentrated around $|\xi| \approx 2^{2j}$. Observe that a shearlet ψ_μ is also a shearlet molecule, but a shearlet molecule has no compact support in the frequency domain, in general. Coarse scale molecules are defined as elements of the form $\{a_\mu(x-k) : k \in \mathbb{Z}^2\}$, where a_μ satisfies (4.14).

Our definition of shearlet molecule is inspired by the curvelet molecules introduced by Candès and Demanet [2]. Both definitions adapt the notion of vaguelettes of Coifman and Meyer [18].

Let us examine a few implications of Definition 4.1. If $m_\mu(x)$ is an horizontal shearlet molecule with regularity R, then it follows from (4.14) that

$$|(2\pi i\xi)^\gamma \, \hat{a}_\mu(\xi)| \le \|\partial^\gamma a_\mu\|_{L^1} \le C_\gamma,$$

and, thus, for all $N \ge 0$ there is a constant C_N such that

$$|\hat{a}_\mu(\xi)| \le C_N \, (1+|\xi|)^{-N}.$$

It follows that for all $N \ge 0$ there is a constant C_N such that

$$|\hat{m}_\mu(\xi)| \le C_N \, 2^{-3j/2} \, (1+|\xi A^{-j} B^{-\ell}|)^{-N}. \tag{4.16}$$

On the other hand, from (4.15) it follows that for each $M \le R$ and each $N \ge 0$ there is a constant $C_{N,M} > 0$ such that

$$|\hat{m}_\mu(\xi)| = |\hat{a}_\mu(\xi \, A^{-j} \, B^{-\ell})| \le C_{N,M} \, 2^{-3j/2} \, \left\{2^{-2j} \, (1+|\xi_1|)\right\}^M \, \left(1+|\xi \, A^{-j} \, B^{-\ell}|\right)^{-N} \tag{4.17}$$

Thus, combining (4.16) and (4.17), it follows that for each $M \le R$ and each $N \ge 0$ there is a constant $C_{N,M} > 0$ such that

$$|\hat{m}_\mu(\xi)| \le C_{N,M} \, 2^{-3j/2} \, \min\left\{1, 2^{-2j} \, (1+|\xi_1|)\right\}^M \, \left(1+|\xi \, A^{-j} \, B^{-\ell}|\right)^{-N}. \tag{4.18}$$

Similarly if $m_\mu^{(v)}(x)$ is a vertical shearlet molecule with regularity R, then for all $N \ge 0$ and all $M \le R$ there is a constant $C_{N,M}$ independent of μ such that:

$$|\hat{m}_\mu^{(v)}(\xi)| \le C_{N,M} \, 2^{-3j/2} \, \min\left\{1, 2^{-2j} \, (1+|\xi_2|)\right\}^M \, \left(1+|\xi \, A_{(v)}^{-j} \, B_{(v)}^{-\ell}|\right)^{-N}. \tag{4.19}$$

Let $m_\mu(x) = m_{j,\ell}(x - A^{-j} B^{-\ell} k)$, where we use the notation $m_{j,\ell}(x) = m_{j,\ell,0}(x)$. Then $m_{j,\ell}(x) = 2^{3j/2} \, a_\mu(B^\ell A^j x)$ and

$$\hat{m}_{j,\ell}(\xi) = 2^{-3j/2} \, \hat{a}_\mu(\xi A^{-j} B^{-\ell}).$$

Thus, by direct computation, observing that $(\xi_1,\xi_2) \, A^{-j} \, B^{-\ell} = (2^{-2j}\xi_1, 2^{-j}(\xi_2 - \ell 2^{-j}\xi_1))$, we have that

$$\begin{aligned} \partial_{\xi_1} \hat{m}_{j,\ell}(\xi) &= 2^{-\frac{3}{2}j} \, \left(2^{-2j} \, P_1(\xi) - \ell 2^{-2j} \, P_2(\xi)\right) \\ \partial_{\xi_2} \hat{m}_{j,\ell}(\xi) &= 2^{-\frac{3}{2}j} \, 2^{-j} \, P_2(\xi), \end{aligned}$$

where $P_n(\xi) = \partial_{\eta_n} \hat{a}(\eta)|_{\eta=(\xi \, A^{-j} \, B^{-\ell})}$, for $n = 1, 2$. Similarly for higher order derivatives. Observe that, by an argument similar to the one given above, it follows that also the functions $P_n(\xi)$ satisfy (4.18). Thus, for each $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N} \times \mathbb{N}$, we have that for all $N \ge 0$ there is a constant C_N such that:

$$\left|\partial_\xi^\alpha \, \hat{m}_{j,\ell}(\xi)\right| \le C_N \, 2^{-\frac{3}{2}j} \, 2^{-j(2\alpha_1+\alpha_2)} \, (1+|\ell|)^{\alpha_1} \, \min\left\{1, 2^{-2j} \, (1+|\xi_1|)\right\}^M \, \left(1+|\xi \, A^{-j} \, B^{-\ell}|\right)^{-N}. \tag{4.20}$$

Next, using the partial derivatives we can compute the directional derivative $D_\theta \, m_\mu(\xi)$ in the direction θ as follows

$$\begin{aligned} D_\theta \, \hat{m}_{j,\ell}(\xi) &= \cos\theta \, \partial_{\xi_1} \hat{m}_{j,\ell}(\xi) + \sin\theta \, \partial_{\xi_2} \hat{m}_{j,\ell}(\xi) \\ &= 2^{-\frac{3}{2}j} \, 2^{-2j} \, P_1(\xi) \cos\theta + (2^{-j} \sin\theta - \ell 2^{-2j} \cos\theta) \, P_2(\xi) \\ &= 2^{-\frac{3}{2}j} \, \cos\theta \, \left(2^{-2j} \, P_1(\xi) + 2^{-j} \, (\tan\theta - \ell 2^{-j}) \, P_2(\xi)\right). \end{aligned}$$

Thus we have that for all $N > 0$ there is a constant C_N such that:

$$|D_\theta \, \hat{m}_{j,\ell}(\xi)| \le C_N \, 2^{-\frac{3}{2}j} \, 2^{-2j} (1 + 2^j |\tan\theta - \ell 2^{-2j}|) \, \min\left\{1, 2^{-2j}\,(1+|\xi_1|)\right\}^M \, \left(1 + |\xi \, A^{-j} \, B^{-\ell}|\right)^{-N}. \quad (4.21)$$

Similar estimates can be derived for the vertical molecules $m_\mu^{(v)}(x)$.

We have the following result.

Theorem 4.2. *Let $\{\psi_\mu : \mu \in \mathcal{M}\}$ be a Parseval frame of shearlets. For each $\mu \in \mathcal{M}$ the operator $T_\mu^{(1)}$ maps ψ_μ into a shearlet molecule $m_\mu = T_\mu^{(1)} \psi_\mu$ with arbitrary regularity R, uniformly in μ. That is, the constant in Definition 4.1 is independent of μ.*

The same result holds for the vertical shearlets $\{\psi_\mu^{(v)} : \mu \in \mathcal{M}\}$. Since the argument is essentially the same, in the following we will only examine the case of horizontal shearlets.

To prove Theorem 4.2 we need the following lemmata. We will use the notation introduced in Section 3. In particular, $\delta_{j,\ell}$ is given by (3.12) and $W_j^{\pm}$ is given by (3.13).

Lemma 4.3. *For $x \in \operatorname{supp}\sigma$, $\eta \in W_j$ and $\alpha = (\alpha_1, \alpha_2)$, we have*

$$\left|\partial_x^\beta \, \partial_\eta^\alpha \, \delta_{j,\ell}(x,\eta)\right| \le C_{\alpha,\beta} \, 2^{-(2\alpha_1+\alpha_2)j},$$

where $C_{\alpha,\beta}$ is independent of x, η, j and ℓ.

Proof. Without loss of generality, we will consider only the case $\eta \in W_j^+$. The case $\eta \in W_j^-$ is similar. Since $\Phi(x,\eta)$ is homogeneous of degree one in η, it follows that

$$\Phi(x,\eta) = \eta \cdot \nabla_\eta \Phi(x,\eta)$$

for all x and η. Also $\nabla_\eta^\alpha \Phi(x,\eta) = O(|\eta|^{1-|\alpha|})$ and $\eta \cdot \nabla_\eta^2 \Phi(x,\eta) = 0$ and . By the definition of $\delta_{j,\ell}(x,\eta)$, we have that $\delta_{j,\ell}(x,\eta)$ is homogeneous of degree one in η, $\delta_{j,\ell}(x,\eta_1,0) = 0$ and $\partial_{\eta_2}\delta_{j,\ell}(x,\eta_1,0) = 0$ for all $\eta \ge 0$. These equations imply that, for each n, $\partial_{\eta_1}^n \delta_{j,\ell}(x,\eta_1,0) = 0$ and $\partial_{\eta_2}\partial_{\eta_1}^n \delta_{j,\ell}(x,\eta_1,0) = 0$. Thus, the Taylor series expansion of $\delta_{j,\ell}(x,\eta)$ about $\eta_2 = 0$ is

$$\delta_{j,\ell}(x,\eta) = \partial_{\eta_2}^2 \delta_{j,\ell}(x,\eta_1,0)\, \frac{\eta_2^2}{2} + h.o.t. \quad (4.22)$$

Recall that on the set W_j^+, we have that:

$$\frac{1}{16}\, 2^{2j} \le \eta_1 \le \frac{1}{2}\, 2^{2j} \quad \text{and } \eta_2 \le \frac{1}{2}\, 2^j.$$

Thus, using (4.22) and the homogeneity assumptions on Φ, we obtain that [1]:

$$\begin{aligned}
\partial_{\eta_1}^{\alpha_1}\delta_{j,\ell}(x,\eta) &= O(|\eta_2|^2\,|\eta|^{-1-\alpha_1}) = O(2^{-2\alpha_1 j}) \text{ on } W_j^+; \\
\partial_{\eta_2}\partial_{\eta_1}^{\alpha_1}\delta_{j,\ell}(x,\eta) &= O(|\eta_2|\,|\eta|^{-1-\alpha_1}) = O(2^{-j-2\alpha_1 j}) \text{ on } W_j^+; \\
\partial_{\eta_2}^{\alpha_2}\partial_{\eta_1}^{\alpha_1}\delta_{j,\ell}(x,\eta) &= O(|\eta|^{1-\alpha_1-\alpha_2}) = O(2^{-2\alpha_1 j-\alpha_2 j}) \text{ on } W_j^+, \text{ for } \alpha_2 \ge 2.
\end{aligned}$$

In the last estimate, we have used the observation that, for $\alpha_2 \ge 2$,

$$|\eta|^{1-\alpha_1-\alpha_2} \le |\eta_1|^{1-\alpha_1-\alpha_2} \le C\, 2^{-2\alpha_1 j} 2^{2(1-\alpha_2)j} \le C\, 2^{-2\alpha_1 j - \alpha_2 j}.$$

Therefore:

$$\left|\partial_\eta^\alpha e^{2\pi i \delta_{j,\ell}(\phi_\mu(x)^{-1},\eta)}\right| \le C_\alpha\, 2^{-2\alpha_1 j-\alpha_2 j}, \text{ on } W_j^+. \quad \square$$

[1] We only need these estimates for $x \in \operatorname{supp}\sigma(\phi_\mu^{-1}(x),\xi)$. Since σ has compact support in x, the estimates are uniform in x.

Lemma 4.4. *For $x \in \operatorname{supp} \sigma$ and $\omega \in E$ we have*

$$\left|\partial_x^\beta\, \partial_\omega^\alpha\, \left(\beta_\mu(x, \omega A^j)\, \hat{\psi}(\omega)\right)\right| \le C_{\alpha,\beta},$$

where $C_{\alpha,\beta}$ is independent of x and $\mu = (j, \ell, k)$.

Proof. It follows from Lemma 4.3 that

$$\sup_{\omega \in E,\, x \in \phi_\mu^{-1}(\operatorname{supp} \sigma)} \left|\partial_x^\beta\, \partial_\omega^\alpha\, \delta_{j,\ell}(\phi_\mu^{-1}(x), \omega A^j)\right| \le C_{\alpha,\beta}, \tag{4.23}$$

where $C_{\alpha,\beta}$ is independent of j and ℓ. Notice that $\omega \neq 0$ on the set E. Also recall that σ is a symbol of order 0. Thus $|\partial_x^\beta \partial_\xi^\alpha \sigma(\phi_\mu^{-1}(x), \xi)| \le C_\alpha\, (1 + |\xi|)^{-|\alpha|}$. This implies that

$$\sup_{\omega \in E,\, x \in \phi_\mu^{-1}(\operatorname{supp} \sigma)} |\partial_x^\beta \partial_\omega^\alpha \sigma(\phi_\mu^{-1}(x), \omega A^j B_{j,\ell}^{-1})| \le C'_{\alpha,\beta}, \tag{4.24}$$

where $C'_{\alpha,\beta}$ is independent of j and ℓ. It follows from (4.23) and (4.24) that

$$\sup_{\omega \in E,\, x \in \phi_\mu^{-1}(\operatorname{supp} \sigma)} \left|\partial_x^\beta\, \partial_\omega^\alpha\, \left(\beta_\mu(x, \omega A^j)\, \hat{\psi}(\omega)\right)\right| \le K_{\alpha,\beta},$$

where $K_{\alpha,\beta}$ is independent of j and ℓ. $\square$

Lemma 4.5. *Suppose that*

$$f(x) = \int_{\mathbb{R}^n} e^{2\pi i \xi (Ax - k)}\, F(x, \xi)\, d\xi,$$

where $A \in GL_n(\mathbb{R})$, $k \in \mathbb{R}^n$, F is C_0^∞ and, for all $\alpha = (\alpha_1, \alpha_2)$ satisfies $\left|\partial_\xi^\alpha\, F(x, \xi)\right| \le C_\alpha$, where C_α is independent of ξ and x. Then, for each $N \in \mathbb{N}$, there is a constant C_N such that, for any $x \in \mathbb{R}^n$, we have

$$|f(x)| \le C_N\, (1 + |Ax - k)|^2)^{-N}. \tag{4.25}$$

If, in addition, $\left|\partial_x^\beta\, F(x, \xi)\right| \le C_\beta$, where C_β is independent of ξ and x, then

$$|\partial_x^\beta\, f(x)| \le |\beta|\, C_N \sup_{\xi \in \operatorname{supp} F} \left|(\xi A)_1^{\beta_1} (\xi A)_2^{\beta_2}\right|\, (1 + |Ax - k)|^2)^{-N}. \tag{4.26}$$

In particular, $C_N = N\, m(R) \left(\|\hat{\psi}\|_\infty + \|\triangle_\xi^N \hat{\psi}\|_\infty\right)$, where $R = \operatorname{supp}_\xi F$, $\triangle_\xi = \sum_{i=1}^n \frac{\partial^2}{\partial \xi_i^2}$ is the frequency domain Laplacian operator and $m(R)$ is the Lebesgue measure of R.

Proof. First observe that

$$|f(x)| \le m(R)\, \|F\|_\infty. \tag{4.27}$$

An integration by parts shows that

$$\int_R e^{2\pi i \xi (Ax - k)}\, \triangle_\xi F(x, \xi)\, d\xi = -(2\pi)^2\, |Ax - k|^2\, F(x)$$

and thus, for every $x \in \operatorname{supp} F$,

$$(2\pi\, |Ax - k|)^{2N}\, |f(x)| \le m(R)\, \|\triangle^N \hat{F}\|_\infty. \tag{4.28}$$

Using (4.27) and (4.28), we have

$$\left(1 + (2\pi\, |Ax - k|)^{2N}\right) |f(x)| \le m(R) \left(\|F\|_\infty + \|\triangle^N F\|_\infty\right). \tag{4.29}$$

Observe that, for each $N \in \mathbb{N}$,

$$(1+|x|^2)^N \le \left(1+(2\pi)^2\,|x|^2\right)^N \le N\left(1+(2\pi\,|x|)^{2N}\right).$$

Using this last inequality and (4.29), we have that for each $x \in \mathbb{R}^n$

$$|f(x)| \le N\,m(R)\,(1+|Ax-k|^2)^{-N}\left(\|F\|_\infty + \|\triangle^N F\|_\infty\right).$$

This proves inequality (4.25). In order to prove (4.26), observe that

$$\partial_x^\beta f(x) = \sum_{\alpha+\gamma=\beta} c_{\alpha,\gamma} \int_{\mathbb{R}^n} (\xi\,A)^\alpha\, e^{2\pi i \xi(Ax+k)}\, \partial_x^\gamma F(x,\xi)\,d\xi.$$

Inequality (4.26) now follows using the same argument as above. □

We can now prove Theorem 4.2.

Proof of Theorem 4.2.
From (3.11), using the change of variables $\omega = \eta\,A^{-j}$, we observe that:

$$\begin{aligned} m_\mu(x) = T_\mu^{(1)}\psi_\mu(x) &= 2^{-3j/2} \int_{W_j} e^{2\pi i\left(\eta\,B_{j,\ell}^{-1}x - \eta\,A^{-j}k\right)}\,\beta_\mu(x,\eta)\,\hat\psi(\eta\,A^{-j})\,d\eta \\ &= 2^{3j/2} \int_E e^{2\pi i\left(\omega A^j\,B_{j,\ell}^{-1}x - \omega k\right)}\,\beta_\mu(x,\omega\,A^j)\,\hat\psi(\omega)\,d\omega \\ &= 2^{3j/2} \int_E e^{2\pi i\omega\left(B^\ell A^j x - k\right)}\,\beta_\mu(x,\omega\,A^j)\,\hat\psi(\omega)\,d\omega. \end{aligned} \tag{4.30}$$

For $\mu \in \mathcal{M}$, let a_μ be defined by

$$m_\mu(x) = 2^{3j/2}\,a_\mu(B^\ell A^j x - k).$$

It follows from (4.30) that

$$a_\mu(y) = 2^{-3j/2}\,m_\mu(A^{-j}\,B^{-\ell}(y+k)) = \int_E e^{2\pi i\omega y}\,\beta_\mu(A^{-j}\,B^{-\ell}(y+k),\omega\,A^j)\,\hat\psi(\omega)\,d\omega. \tag{4.31}$$

We need to show that a_μ satisfies (4.14) and (4.15).

By Lemmata 4.4 and 4.5 applied to (4.31) it follows that, for all $N > 0$, there is a constant C_N such that

$$|a_\mu(y)| \le C_N\,(1+|y|)^{-N}. \tag{4.32}$$

For any $\gamma = (\gamma_1,\gamma_2) \in \mathbb{N}\times\mathbb{N}$, we have that

$$\partial_y^\gamma a_\mu(y) = \int_E \partial_y^\gamma\left(e^{2\pi i\omega y}\,\beta_\mu(A^{-j}\,B^{-\ell}(y+k),\omega\,A^j)\right)\,\hat\psi(\omega)\,d\omega.$$

By applying to this expression Lemmata 4.4 and 4.5 we obtain the same estimates as $a_\mu(y)$, given by (4.32). This gives (4.14).

To deduce the second estimate, let us compute the Fourier transform of $a_\mu(y)$, given by (4.31). We have:

$$\hat a_\mu(\xi) = \int_{\mathbb{R}^2}\int_E e^{2\pi i\omega y}\,e^{-2\pi i\xi y}\,\beta_\mu\left(A^{-j}\,B^{-\ell}\,(y+k)\,,\omega\,A^j\right)\,\hat\psi(\omega)\,d\omega\,dy. \tag{4.33}$$

The integral (4.33) is well defined since $\beta_\mu(y,\omega)$ has compact support in y. Next set $D_1 = \frac{1}{2\pi i}\partial_{y_1}$ and observe that

$$D_1\,e^{2\pi i\omega y} = \omega_1\,e^{2\pi i\omega y}.$$

Thus, using an integration by parts, we have that

$$\begin{aligned}\hat{a}_\mu(\xi) &= \int_{\mathbb{R}^2}\int_E D_1^{-M}\left(e^{2\pi i\omega y}\right)\, D_1^M\left(e^{-2\pi i\xi y}\beta_\mu\left(A^{-j}\,B^{-\ell}\,(y+k)\,,\omega\,A^j\right)\right)\,\hat{\psi}(\omega)\,d\omega\,dy\\ &= \int_{\mathbb{R}^2}\int_E \omega_1^{-N}\,e^{2\pi i\omega y}\,D_1^M\left(e^{-2\pi i\xi y}\beta_\mu(A^{-j}\,B^{-\ell}(y+k),\omega\,A^j)\right)\,\hat{\psi}(\omega)\,d\omega\,dy.\end{aligned}$$

This expression can be written as a sum of the form

$$\hat{a}_\mu(\xi) = \sum_{l=0}^{M} c_l\,\xi_1^l\,2^{-2j(M-l)}\int_{\mathbb{R}^2} e^{-2\pi i y\xi}F_l(y)\,dy = \sum_{l=0}^{M} c_l\,\xi_1^l\,2^{-2j(M-l)}\,\hat{F}_l(\xi), \tag{4.34}$$

where

$$F_l(y) = \int_E e^{2\pi i\omega y}\,2^{2j(M-l)}\,\partial_{y_1}^{M-l}\left(\beta_\mu(A^{-j}\,B^{-\ell}(y+k),\omega\,A^j)\right)\,\omega_1^{-M}\,\hat{\psi}(\omega)\,d\omega.$$

Observe that, in the expression above, $\omega_1^{-1}\le 16$ on E and $|2^{2j(M-l)}\,\partial_{y_1}^{M-l}\left(\beta_\mu(A^{-j}\,B^{-\ell}(y+k),\omega\,A^j)\right)|\le C_M$ uniformly. Thus, by applying Lemma 4.5, we have that for all $M\ge 0$ there is a $C_{\mathcal{M}}$ such that

$$|F_l(y)|\le C_M\,\left(1+|y|^2\right)^{-M},$$

and, similarly, that for all $M\ge 0$ and all α

$$|\partial_y^\alpha F_l(y)|\le C_{\alpha,M}\,\left(1+|y|^2\right)^{-M}.$$

It follows that, for all $\xi\in\operatorname{supp}F_l$,

$$|(2\pi\xi)^\alpha\,F_l(\xi)|\le\|\partial^\alpha F_l\|_{L^1}\le C_\alpha,$$

and, thus, for all $M>0$ there is a constant C_M such that

$$|\hat{F}_l(\xi)|\le C_N\,(1+|\xi|)^{-N}. \tag{4.35}$$

The estimate (4.15) now follows from (4.34), the binomial theorem, and (4.35). □

The following observation shows that the molecules m_μ, $\mu\in\mathcal{M}$ are well localized.

Proposition 4.6. *Let $m_\mu = T_\mu^{(1)}\psi_\mu$. For all $N\in\mathbb{N}$ there is a constant C_N, independent of j,ℓ,k,x such that*

$$|m_\mu(x)|\le C_N\,2^{3j/2}\,(1+|B^\ell\,A^j x-k)|^2)^{-N}. \tag{4.36}$$

Further, for all $\beta=(\beta_1,\beta_2)$ and all $N\in\mathbb{N}$ there is a constant K_N, independent of j,ℓ,k,x such that

$$|\partial_x^\beta m_\mu(x)|\le C_N\,2^{3j/2}\,2^{2j(\beta_1+\beta_2)}\,(1+|B^\ell\,A^j x-k)|^2)^{-N}. \tag{4.37}$$

Proof.

To prove inequality (4.36) we apply Lemma 4.5 to the function $m_\mu(x)$, given by (4.30). Observe that the assumptions of Lemma 4.5 are satisfied by Lemma 4.4. Next observe that

$$(\omega_1,\omega_2)B^\ell A^j = (2^{2j}\omega_1, 2^j(\ell\omega_1+\omega_2)).$$

Thus, inequality (4.37) also follows from Lemmata 4.4 and 4.5 by observing that, for $\omega\in E$:

$$\left|(\omega B^\ell A^j)_1^{\beta_1}(\omega B^\ell A^j)_2^{\beta_2}\right|\le C\,2^{2j\beta_1}\,2^{j\beta_2}|\ell|^{\beta_2}\le C\,2^{2j(\beta_1+\beta_2)}.\quad\square$$

4.2 Almost orthogonality

While shearlets do not form an orthogonal family, they are associated to a notion of almost orthogonality. In fact, the inner product of two shearlets ψ_μ and $\psi_{\mu'}$ exhibits an almost exponential decay as a function of an appropriate distance defined on the indices μ and μ' in $\mathcal{M}$.

Given a pair of indices $\mu = (j, \ell, k)$ and $\mu' = (j', \ell', k')$ in $\mathcal{M}$, the *dyadic parabolic pseudo-distance* $\omega(\mu, \mu')$ is defined by:

$$\omega(\mu, \mu') = 2^{|j-j'|} \left(1 + 2^{\min(j,j')} \, d(\mu, \mu')\right),$$

where

$$d(\mu, \mu') = |\ell 2^{-j} - \ell' 2^{-j'}|^2 + |k_{j,\ell} - k'_{j',\ell'}|^2 + |\langle e_\mu, k_{j,\ell} - k'_{j',\ell'} \rangle|,$$

$e_\mu = (\cos\theta_\mu, \sin\theta_\mu)$ and $\theta_\mu = \arctan(\ell 2^{-j})$. Observe that the term $|\langle e_\mu, k_{j,\ell} - k'_{j',\ell'} \rangle|$ induces a non-Euclidean notion of distance. This definition is motivated by a similar pseudo-distance introduced by Smith [19] which was later modified and used in the work of Candès and Demanet [2].

Proposition 4.7. *For $\mu, \mu', \mu'', \mu_0 \in \mathcal{M}$, the dyadic parabolic pseudo-distance ω satisfies the following properties.*

(1) Symmetry. $\omega(\mu, \mu') \sim \omega(\mu', \mu)$.

(2) Triangle Inequality. There is a constant $C > 0$ such that $d(\mu, \mu') \le C\, d(\mu, \mu'') + d(\mu'', \mu')$.

(3) Composition. For every $N > 0$, there is a constant $C_N > 0$ such that

$$\sum_{\mu''} \omega(\mu, \mu'')^{-N} \, \omega(\mu'', \mu')^{-N} \le C_N \, \omega(\mu, \mu')^{-N+1}.$$

(4) Invariance under Hamiltonian flows: $\omega(\mu, \mu') \sim \omega(\mu'(t), \mu(t))$, where $\mu(t)$ is the shearlet index μ evolved along the Hamiltonian system.

Proof. The proof follows essentially from the arguments in [2, Prop.2.2]. For brevity, in the following, we will only indicate how to adapt those arguments to our definition.

(1.) In our definition of d, the term $|\ell 2^{-j} - \ell' 2^{-j'}|^2 + |k_{j,\ell} - k'_{j',\ell'}|^2$ is obviously symmetric. Thus we only have to show that

$$|\langle e_\mu, k_{j,\ell} - k'_{j',\ell'} \rangle| \sim |\langle e'_\mu, k_{j,\ell} - k'_{j',\ell'} \rangle|. \tag{4.38}$$

Since this non-Euclidean term is defined formally as in [2], the proof of (4.38) follows verbatim as in the proof of [2, Prop.2.2].

(2.) A direct calculation shows that:

$$\begin{aligned} |k_{j,\ell} - k'_{j',\ell'}|^2 &= |k_{j,\ell} - k''_{j'',\ell''} + k''_{j'',\ell''} - k'_{j',\ell'}|^2 \\ &\le 2\left(|k_{j,\ell} - k''_{j'',\ell''}|^2 + |k''_{j'',\ell''} - k'_{j',\ell'}|^2\right). \end{aligned}$$

Thus, the triangle inequality certainly holds on the symmetric part of d. As above, for the non-Euclidean term the proof follows verbatim as in [2, Prop.2.2].

(3.) The proof is again very similar to [2, Prop.2.2]. We will only indicate the minor adjustments that are needed. In particular, using our definition of d, equation (7.4) in [2] becomes

$$\sum_{\ell=0}^{2^j-1} \sum_{k \in \mathbb{Z}^2} (1 + 2^q \, d(\mu, \mu'))^{-N} \le C \sum_{\ell=0}^{2^j-1} \sum_{k \in \mathbb{Z}^2} \left(1 + 2^q \left(|2^{-j}\ell|^2 + |2^{-j} k_2|^2 + |2^{-2j} k_1|\right)\right)^{-N}.$$

Now, arguing as in [2, Prop.2.2], we have that

$$\sum_{\ell=0}^{2^j-1} \sum_{k \in \mathbb{Z}^2} (1 + 2^q \, d(\mu, \mu'))^{-N} \le C \, 2^{2(2j-q)_+}, \tag{4.39}$$

where the subscript + denotes the positive part. The remaining part of the argument follows as in [2, Prop.2.2]. For example (we refer there for the definition of I_{μ_1}), using (4.39), the estimate for the case $0 \le j_2 \le j_1$ is

$$\begin{aligned}\sum_{\mu_1} I_{\mu_1} &\le C\,(1+2^{j_0}\,d_{02})^{-N} \sum_{j_1\ge j_2} 2^{-(2j_1-j_0-j_2)N}\,2^{4j_1-2j_2}\\ &\le C\,2^{-(j_2-j_0)N}\,(1+2^{j_0}\,d_{02})^{-N} = C\,\omega(\mu_0,\mu_1)^{-N}.\end{aligned}$$

(4.) This property follows from [20, p.804]. Indeed, the symmetric part of d is clearly the invariance along the Hamiltonian flow. We only have to consider the non-Euclidean part of d. However, this term is defined as in the curvelet case and the proof for this part is given in [20, p.804]. □

The following Propositions 4.8 and 4.9 show that the shearlet molecules, defined by Definition 4.1, form an almost orthogonal family with respect to the dyadic parabolic pseudo-distance ω. The almost orthogonality will play a major role in the proof of Theorem 3.1.

Proposition 4.8. *Let m_μ and $m_{\mu'}$ be two shearlet molecules with regularity R. Let $j, j' \ge 0$. For every $N \le f(R)$, there is a constant $C_N > 0$ such that*

$$|\langle m_\mu, m_{\mu'}\rangle| \le C_N\,\omega(\mu,\mu')^{-N}.$$

The number $f(R)$ increases with R and goes to infinity as R goes to infinity.

Proof. The following argument adapts some ideas from Lemma 2.3 in [2].

Consider the integral:

$$I(\mu,\mu') = \int\!\!\int_{\mathbb{R}^2} \hat{p}_\mu(\xi_1,\xi_2)\,\overline{\hat{p}_{\mu'}(\xi_1,\xi_2)}\,d\xi_1\,d\xi_2,$$

where the functions p_μ, p'_μ satisfy the estimate (4.18).

For $j' \ge j$, we split the integral as

$$I(\mu,\mu') = I_1(\mu,\mu') + I_2(\mu,\mu') + I_3(\mu,\mu') + I_4(\mu,\mu'),$$

where I_1 corresponds to the region of integration $\{(\xi_1,\xi_2) : |\xi_1| \ge 2^{2j'}\}$, I_2 to the region of integration $\{(\xi_1,\xi_2) : 2^{2j} \le |\xi_1| < 2^{2j'}\}$, I_3 to the region of integration $\{(\xi_1,\xi_2) : 1 \le |\xi_1| < 2^{2j}\}$ and I_4 to the region of integration $\{(\xi_1,\xi_2) : 0 \le |\xi_1| < 1\}$.

We can assume $\xi_1 \ge 0$ (the case $\xi_1 < 0$ is similar). Using the change of variables $\xi_2 = u\,\xi_1$, we have

$$\begin{aligned} I_1(\mu,\mu') &\le C_N \int_{2^{2j'}}^{\infty}\int_{-\infty}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}}{\left(1+2^{-2j}\,\xi_1+2^{-j}\,|\xi_2-\ell\,2^{-j}\,\xi_1|\right)^N\,\left(1+2^{-2j'}\,\xi_1+2^{-j'}\,|\xi_2-\ell'\,2^{-j'}\,\xi_1|\right)^N}\,d\xi_2\,d\xi_1\\ &= C_N \int_{2^{2j'}}^{\infty}\int_{-\infty}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{\left(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|\right)^N\,\left(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|u-\ell'\,2^{-j'}|\right)^N}\,du\,d\xi_1\\ &= C_N \int_{2^{2j'}}^{\infty}\int_{-\infty}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{\left(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u|\right)^N\,\left(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|u+\Delta_{\mu,\mu'}|\right)^N}\,du\,d\xi_1, \end{aligned}$$

where $\Delta_{\mu,\mu'} = \ell'\,2^{-j'} - \ell\,2^{-j}$. By introducing the variables $\alpha = \frac{2^{-j}\xi_1}{1+2^{-2j}\xi_1}$ and $\alpha' = \frac{2^{-j'}\xi_1}{1+2^{-2j'}\xi_1}$, we can factorize the fraction inside the integral sign to obtain

$$I_1(\mu,\mu') \le C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\,\xi_1)^N\,(1+2^{-2j'}\,\xi_1)^N} \int_{-\infty}^{\infty} \frac{1}{\left(1+\alpha|u|\right)^N\,\left(1+\alpha'\,|u+\Delta_{\mu,\mu'}|\right)^N}\,du\,d\xi_1.$$

A classical estimate (see [9, Appendix K]) gives that, for $\alpha \ge \alpha'$, and $N > 1$,

$$\int_{-\infty}^{\infty} \frac{1}{\left(1+\alpha\,|u|\right)^N\,\left(1+\alpha'\,|u+\Delta_{\mu,\mu'}|\right)^N}\,du \le C_N\,\frac{1}{\alpha\,\left(1+\alpha'\,|\Delta_{\mu,\mu'}|\right)^N}. \qquad (4.40)$$

From the definition α', for $\xi_1 \geq 2^{2j'}$, we have that $\frac{1}{2}\,2^{j'} \leq \alpha' \leq 2^{j'}$. Thus, for $\xi_1 \geq 2^{2j'}$, provided $N > 1$, from the last inequality we obtain

$$\int_{-\infty}^{\infty} \frac{1}{(1+\alpha\,|u|)^N\,(1+\alpha'\,|u+\Delta_{\mu,\mu'}|)^N}\,du \leq C_N\,\frac{(1+2^{-2j}\,\xi_1)}{2^{-j}\,\xi_1}\,\left(1+2^{j'}\,|\Delta_{\mu,\mu'}|\right)^{-N}. \tag{4.41}$$

If $\alpha' \geq \alpha$, arguing in a similar way, we obtain that, for $N > 1$ and $\xi_1 \geq 2^{2j'}$:

$$\int_{-\infty}^{\infty} \frac{1}{(1+\alpha\,|u|)^N\,(1+\alpha'\,|u+\Delta_{\mu,\mu'}|)^N}\,du \leq C_N\,\frac{(1+2^{-2j'}\,\xi_1)}{2^{-j'}\,\xi_1}\,\left(1+2^{j}\,|\Delta_{\mu,\mu'}|\right)^{-N}.$$

Also observe that, for each $N, N' \in \mathbb{N}$, we have that

$$\begin{aligned}
\int_{2^{2j'}}^{\infty} \frac{1}{(1+2^{-2j'}\,\xi_1)^{N'}\,(1+2^{-2j}\,\xi_1)^N}\,d\xi_1 &\leq \left(1+2^{2(j'-j)}\right)^{-N}\int_{2^{2j'}}^{\infty}\frac{1}{(1+2^{-2j'}\,\xi_1)^{N'}}\,d\xi_1\\
&= \left(1+2^{2(j'-j)}\right)^{-N}\,2^{2j'}\int_{1}^{\infty}\frac{1}{(1+\xi_1)^{N'}}\,d\xi_1.
\end{aligned}$$

For $N' > 1$, it follows that

$$\int_{2^{2j'}}^{\infty} \frac{1}{(1+2^{-2j'}\,\xi_1)^{N'}\,(1+2^{-2j}\,\xi_1)^N}\,d\xi_1 \leq C_{N'}\,2^{2j'}\,\left(1+2^{2(j'-j)}\right)^{-N}. \tag{4.42}$$

Thus, using (4.41) and (4.42), for $\alpha \geq \alpha' > 1$, we have that

$$\begin{aligned}
I_1(\mu,\mu') &\leq C_N\int_{2^{2j'}}^{\infty}\frac{2^{-\frac{3}{2}(j+j')}}{(1+2^{-2j}\,\xi_1)^{N-1}\,(1+2^{-2j'}\,\xi_1)^N}\,2^j\,\left(1+2^{j'}\,|\Delta_{\mu,\mu'}|\right)^{-N}\,d\xi_1\\
&\leq C_N\,2^{-\frac{3}{2}(j'+j)}\,2^{j+2j'}\,\left(1+2^{2(j'-j)}\right)^{-N}\,\left(1+2^{j'}\,|\Delta_{\mu,\mu'}|\right)^{-N}\\
&\leq C_N\,2^{\frac{1}{2}(j'-j)}\,2^{-2N(j'-j)}\,\left(1+2^{j'}\,|\Delta_{\mu,\mu'}|\right)^{-N}\\
&\leq C_N\,2^{\frac{1}{2}(j'-j)}\,2^{-2N(j'-j)}\,\left(1+2^{j}\,|\Delta_{\mu,\mu'}|\right)^{-N}.
\end{aligned} \tag{4.43}$$

The same estimate holds for $\alpha' \geq \alpha$.

To estimate the integral I_2 we will use the additional factor $2^{-2j'N}\,(1+|\xi_1|)^N$ from (4.18) (we choose $N = M$). We still assume that $j' \geq j > 1$. Thus, for $\alpha \geq \alpha'$ (the case $\alpha < \alpha'$ is handled in a similar way) we have:

$$I_2(\mu,\mu') \leq C_N\int_{2^{2j}}^{2^{2j'}}\frac{2^{-\frac{3}{2}(j+j')}\,\xi_1\,2^{-2j'N}\,(1+\xi_1)^N}{(1+2^{-2j'}\,\xi_1)^{N'}\,(1+2^{-2j}\,\xi_1)^N}\int_{-\infty}^{\infty}\frac{1}{(1+\alpha|u|)^N\,\left(1+\alpha'\,|u+\Delta_{\mu,\mu'}|\right)^N}\,du\,d\xi_1.$$

Thus, using (4.40) we have

$$\begin{aligned}
I_2(\mu,\mu') &\leq C_N\int_{2^{2j}}^{2^{2j'}}\frac{2^{-\frac{3}{2}(j+j')}\,\xi_1\,2^{-2j'N}\,(1+\xi_1)^N\,\left(1+2^{-2j}\,\xi_1\right)\left(1+2^{-2j'}\,\xi_1\right)^N}{(1+2^{-2j'}\,\xi_1)^N\,(1+2^{-2j}\,\xi_1)^N\,2^{-j}\,\xi_1\,\left(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|\right)^N}\,d\xi_1\\
&= C_N\,2^{-\frac{1}{2}(j+3j')}\,2^{-2j'N}\int_{2^{2j}}^{2^{2j'}}\frac{(1+\xi_1)^N}{(1+2^{-2j}\,\xi_1)^{N-1}\,\left(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|\right)^N}\,d\xi_1\\
&\leq C_N\,2^{-\frac{1}{2}(j+3j')}\,2^{-2j'N}\,\frac{1}{(1+2^{2j-j'}\,|\Delta_{\mu,\mu'}|)^N}\int_{2^{2j}}^{2^{2j'}}\frac{\xi_1^N}{(1+2^{-2j}\,\xi_1)^N}\,d\xi_1\\
&= C_N\,2^{-\frac{1}{2}(j+3j')}\,2^{-2j'N}\,2^{2j}\,\frac{1}{(1+2^{2j-j'}\,|\Delta_{\mu,\mu'}|)^N}\int_{1}^{2^{2(j'-j)}}\frac{2^{2jN}\,u^N}{(1+u)^N}\,d\xi_1.
\end{aligned}$$

Thus, observing that

$$\int_1^{2^{2(j'-j)}} \frac{u^N}{(1+u)^N}\, d\xi_1 \le 2^{2(j'-j)},$$

we conclude that

$$\begin{aligned} I_2(\mu,\mu') &\le C_N\, 2^{\frac12(j'-j)}\, 2^{-2(j'-j)N} \left(1+2^{2j-j'}\,|\Delta_{\mu,\mu'}|\right)^{-N} \\ &\le C_N\, 2^{\frac12(j'-j)}\, 2^{-(j'-j)N} \left(1+2^{j}\,|\Delta_{\mu,\mu'}|\right)^{-N}. \end{aligned} \tag{4.44}$$

To estimate the integral I_3 we will use the additional factor $2^{-2(j'+j)M}\,(1+|\xi_1|)^M$ (notice that we choose $M \neq N$). We still assume that $j' \ge j > 1$. Thus, for $\alpha \ge \alpha'$ (the case $\alpha < \alpha'$ is handled in a similar way) we have:

$$I_3(j,j',\ell,\ell') \le C_{N,M} \int_1^{2^{2j}} \frac{2^{-\frac32(j+j')}\,\xi_1\, 2^{-2(j+j')M}\,(1+\xi_1)^{2M}}{(1+2^{-2j'}\,\xi_1)^{N'}\,(1+2^{-2j}\,\xi_1)^{N}} \int_{-\infty}^{\infty} \frac{1}{(1+\alpha|u|)^N\,\,(1+\alpha'\,|u+\Delta_{\mu,\mu'}|)^N}\, du\, d\xi_1.$$

Using again (4.40) we have

$$\begin{aligned} I_3(\mu,\mu') &\le C_{N,M} \int_1^{2^{2j}} \frac{2^{-\frac32(j+j')}\,\xi_1\, 2^{-2(j+j')M}\,(1+\xi_1)^{2M}\,\left(1+2^{-2j}\,\xi_1\right)\left(1+2^{-2j'}\,\xi_1\right)^N}{(1+2^{-2j'}\,\xi_1)^{N}\,(1+2^{-2j}\,\xi_1)^{N}\; 2^{-j}\,\xi_1\,\,(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &= C_{N,M}\, 2^{-\frac12(j+3j')}\, 2^{-2(j+j')M} \int_1^{2^{2j}} \frac{(1+\xi_1)^{2M}}{(1+2^{-2j}\,\xi_1)^{N-1}\,\,(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &\le C_{N,M}\, 2^{-\frac12(j+3j')}\, 2^{-2(j+j')M} \int_1^{2^{2j}} \frac{\xi_1^{2M}}{(1+2^{-2j}\,\xi_1)^{N}\,\,(1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &\le C_{N,M}\, 2^{-\frac12(j+3j')} \int_1^{2^{2j}} \frac{2^{-2j'M}\,\xi_1^{M}}{(1+2^{-2j}\,\xi_1)^{N}\,\,(1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1. \end{aligned}$$

Let $M = 2N$. Then

$$\begin{aligned} I_3(\mu,\mu') &\le C_N 2^{-\frac12(j+3j')} \int_1^{2^{2j}} \frac{2^{-2j'N}\,\xi_1^{N}}{(1+2^{-2j}\,\xi_1)^{N}\,\,2^{2j'N}\,\xi_1^{-N}\,\,(1+2^{-j'}\,\xi_1\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &\le C_N 2^{-\frac12(j+3j')} \int_1^{2^{2j}} \frac{2^{-2j'N}\,\xi_1^{N}}{(1+2^{-2j}\,\xi_1)^{N}\,\,(1+2^{j'}\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &\le C_N 2^{-\frac12(j+3j')} \int_1^{2^{2j}} \frac{2^{-2j'N}\,\xi_1^{N}}{(1+2^{j}\,|\Delta_{\mu,\mu'}|)^N}\, d\xi_1 \\ &\le C_N 2^{-\frac32(j'-j)}\, 2^{-2(j'-j)N} \left(1+2^{j}\,|\Delta_{\mu,\mu'}|\right)^{-N}. \end{aligned} \tag{4.45}$$

To estimate the integral I_4 we choose $M = N$. Then, still assuming that $j' \ge j > 1$, we have

$$\begin{aligned} I_4(\mu,\mu') &\le C_N \int_0^1 \int_{-\infty}^{\infty} \frac{2^{-\frac32(j+j')}\,\xi_1\, 2^{-2(j+j')N}\,(1+\xi_1)^{2N}}{(1+2^{-j}\,\xi_1\,|u|)^N}\, du\, d\xi_1 \\ &\le C_N\, 2^{-\frac32(j+j')}\, 2^{-2(j+j')N} \int_0^1 \int_{-\infty}^{\infty} 2^j\, \frac{(1+\xi_1)^{2N}}{(1+|v|)^N}\, dv\, d\xi_1 \\ &\le C_N\, 2^{-\frac12(j+3j')}\, 2^{-2(j+j')N}. \end{aligned} \tag{4.46}$$

Combining the estimates for (4.43), (4.44), (4.45) and (4.46) we conclude that, if $j' \geq j$, for all $N \in \mathbb{N}$ there is a constant C_N such that:

$$I(\mu, \mu') \leq C_N \, 2^{-(j'-j)N} \, \left(1 + 2^j \, |\Delta_{\mu,\mu'}|\right)^{-N}. \tag{4.47}$$

We will now deduce another estimate for the decay associated with the spatial location. We still assume that $j' \geq j$. Recall that, for $\mu = (j, \ell, k)$, $m_\mu(x) = m_{j,\ell}(x - A^{-j}B^{-\ell}k)$ and $\hat{m}_\mu(\xi) = \hat{m}_{j,\ell}(\xi)\, e^{-2\pi i \xi A^{-j} B^{-\ell} k}$. Let $\triangle = \partial^2_{\xi_1} + \partial^2_{\xi_2}$ be the frequency Laplacian operator. By (4.20), we have that for all $N, M > 0$ there is a constant $C_{N,M}$ such that:

$$\left|\triangle \left(\hat{m}_{j,\ell}(\xi)\, \overline{\hat{m}_{j',\ell'}(\xi)}\right)\right| \leq C_{N,M} \, \frac{2^{-2j}\, 2^{-\frac{3}{2}(j+j')}\, \min\left\{1, 2^{-2j}(1+|\xi_1|)\right\}^M \, \min\left\{1, 2^{-2j'}(1+|\xi_1|)\right\}^M}{(1+2^{-2j}\,|\xi_1|+2^{-j}\,|\xi_2 - \ell\, 2^{-j}\, \xi_1|)^N \left(1+2^{-2j'}\,|\xi_1|+2^{-j'}\,|\xi_2 - \ell'\, 2^{-j'}\, \xi_1|\right)^N}. \tag{4.48}$$

Let D_μ be the directional derivative in the direction of μ, that is, $\tan\theta = 2^{-j}\ell$. Then, using (4.21) we have that

$$\left|D^2_\mu \, \hat{m}_{j,\ell}(\xi)\right| \leq C_{N,M} \, 2^{-\frac{3}{2}j}\, 2^{-4j} \, \min\left\{1, 2^{-2j}(1+|\xi_1|)\right\}^M \, \left(1 + |\xi \, A^{-j}\, B^{-\ell}|\right)^{-N}$$

and

$$\left|D^2_\mu \, \hat{m}_{j',\ell'}(\xi)\right| \leq C_{N,M} \, 2^{-\frac{3}{2}j'}\, 2^{-4j'} (1 + 2^{2j'}|\Delta_{\mu,\mu'}|^2) \, \min\left\{1, 2^{-2j'}(1+|\xi_1|)\right\}^M \, \left(1 + |\xi \, A^{-j}\, B^{-\ell}|\right)^{-N}.$$

It follows that

$$\begin{aligned} &\left|\frac{2^{4j}}{(1+2^{2j}|\Delta_{\mu,\mu'}|^2)} D^2_\theta \left(\hat{m}_{j,\ell}(\xi)\, \overline{\hat{m}_{j',\ell'}(\xi)}\right)\right| \\ &\leq C_{N,M} \, \frac{2^{-\frac{3}{2}(j+j')}\, \min\left\{1, 2^{-2j}(1+|\xi_1|)\right\}^M \, \min\left\{1, 2^{-2j'}(1+|\xi_1|)\right\}^M}{(1+2^{-2j}\,|\xi_1|+2^{-j}\,|\xi_2 - \ell\, 2^{-j}\, \xi_1|)^N \, (1+2^{-2j'}\,|\xi_1|+2^{-j'}\,|\xi_2 - \ell'\, 2^{-j'}\, \xi_1|)^N}. \end{aligned} \tag{4.49}$$

Set

$$L = I - 2^{2j}\, \frac{\triangle}{(2\pi)^2} - \frac{2^{4j}}{(2\pi)^2\,(1+2^{2j}|\Delta_{\mu,\mu'}|^2)}\, D^2_\mu.$$

Then, using (4.48) and (4.49), we have that, for all $N, M > 0$ and for each $n \leq N$ there is a $C_{N,M}$ such that

$$\left|L^n \left(\hat{m}_{j,\ell}(\xi)\, \overline{\hat{m}_{j',\ell'}(\xi)}\right)\right| \leq C_{N,M} \, \frac{2^{-\frac{3}{2}(j+j')}\, \min\left\{1, 2^{-2j}(1+|\xi_1|)\right\}^M \, \min\left\{1, 2^{-2j'}(1+|\xi_1|)\right\}^M}{(1+2^{-2j}\,|\xi_1|+2^{-j}\,|\xi_2 - \ell\, 2^{-j}\, \xi_1|)^N \left(1+2^{-2j'}\,|\xi_1|+2^{-j'}\,|\xi_2 - \ell'\, 2^{-j'}\, \xi_1|\right)^N}. \tag{4.50}$$

We also have that for each $n \in \mathbb{N}$

$$L^{-n}\left(e^{-2\pi i \xi (k_{j,\ell} - k'_{j,\ell})}\right) = \left(1 + 2^{2j}\,|k_{j,\ell} - k'_{j,\ell}|^2 + \frac{2^{4j}}{(1+2^{2j}|\Delta_{\mu,\mu'}|^2)} |\langle e_\mu, k_{j,\ell} - k'_{j,\ell}\rangle|^2\right)^{-n} e^{-2\pi i \xi (k_{j,\ell} - k'_{j,\ell})}. \tag{4.51}$$

where $k_{j,\ell} = A^{-j}B^{-\ell}k$, $k'_{j,\ell} = A^{-j'}B^{-\ell'}k'$. Repeated integrations by parts give

$$\begin{aligned} \langle m_\mu, m_{\mu'}\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{m}_{j,\ell}(\xi)\, \overline{\hat{m}_{j',\ell'}(\xi)}\, e^{-2\pi i \xi (k_{j,\ell} - k'_{j,\ell})}\, d\xi_1\, d\xi_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L^n\left(\hat{m}_{j,\ell}(\xi)\, \overline{\hat{m}_{j',\ell'}(\xi)}\right) \, L^{-n}\left(e^{-2\pi i \xi (k_{j,\ell} - k'_{j,\ell})}\right)\, d\xi_1\, d\xi_2. \end{aligned}$$

Therefore, by (4.50) and (4.51) we have that for all $N, M > 0$ there is a constant $C_{N,M}$ such that:

$$\begin{aligned} |\langle m_\mu, m_{\mu'}\rangle| \;\leq\; & C_{N,M} \left(1 + 2^{2j}\,|k_{j,\ell} - k'_{j,\ell}|^2 + \frac{2^{4j}}{(1+2^{2j}|\Delta_{\mu,\mu'}|^2)} |\langle e_\mu, k_{j,\ell} - k'_{j,\ell}\rangle|^2\right)^{-N} \\ & \times \int\int \frac{2^{-\frac{3}{2}(j+j')}\, \min\left\{1, 2^{-2j}(1+|\xi_1|)\right\}^M \, \min\left\{1, 2^{-2j'}(1+|\xi_1|)\right\}^M}{(1+2^{-2j}\,|\xi_1|+2^{-j}\,|\xi_2 - \ell\, 2^{-j}\, \xi_1|)^N \left(1+2^{-2j'}\,|\xi_1|+2^{-j'}\,|\xi_2 - \ell'\, 2^{-j'}\, \xi_1|\right)^N}\, d\xi_1\, d\xi_2, \end{aligned}$$

where the integral is of the same form as $I(\mu,\mu')$. Thus, from the last expression, using (4.47), we have that for every $N>0$ there is a constant C_N such that:

$$\begin{aligned}
|\langle m_\mu, m_{\mu'}\rangle| &\le C_N\, 2^{-(j'-j)N}\,\left(1+2^{2j}\,|\Delta_{\mu,\mu'}|^2\right)^{-N}\left(1+2^{2j}|k_{j,\ell}-k'_{j,\ell}|^2+\frac{2^{4j}}{(1+2^{2j}|\Delta_{\mu,\mu'}|^2)}|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle|^2\right)^{-N}\\
&\le C_N\, 2^{-(j'-j)N}\left(1+2^{2j}\left(|\Delta_{\mu,\mu'}|^2+|k_{j,\ell}-k'_{j,\ell}|^2\right)+\frac{2^{4j}}{(1+2^{2j}|\Delta_{\mu,\mu'}|^2)}|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle|^2\right)^{-N}\\
&\le C_N\, 2^{-(j'-j)N}\,\left(1+2^{2j}\left(|\Delta_{\mu,\mu'}|^2+|k_{j,\ell}-k'_{j,\ell}|^2+|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle|\right)\right)^{-N}.
\end{aligned}$$

In the last step we have used the observation that

$$\begin{aligned}
2^{2j}\,|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle| &= \sqrt{1+2^{2j}|\Delta_{\mu,\mu'}|^2}\;\frac{2^{2j}\,|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle|}{\sqrt{1+2^{2j}|\Delta_{\mu,\mu'}|^2}}\\
&\le C\left((1+2^{2j}\,|\Delta_{\mu,\mu'}|^2)+\frac{2^{4j}\,|\langle e_\mu, k_{j,\ell}-k'_{j,\ell}\rangle|^2}{1+2^{2j}|\Delta_{\mu,\mu'}|^2}\right)
\end{aligned}$$

The case $j'\le j$ is symmetric to $j\le j'$ and follows easily from the estimates above. □

The almost orthogonality of shearlets molecules extends to the situation where there are both vertical and horizontal molecules. The case where both shearlet molecules are of vertical type follows easily by adapting Proposition 4.8. The case where one molecule is horizontal and the other one is vertical is proved by the following proposition.

Proposition 4.9. *Let m_μ and $m^{(v)}_{\mu'}$ be a horizontal and a shearlet molecules with regularity R, respectively. Let $j,j'\ge 0$. For every $N\le f(R)$, there is a constant $C_N>0$ such that*

$$|\langle m_\mu, m^{(v)}_{\mu'}\rangle|\le C_N\,\omega(\mu,\mu')^{-N}.$$

The number $f(R)$ increases with R and goes to infinity as R goes to infinity.

Proof. Consider the integral:

$$I(\mu,\mu')=\int\!\!\int_{\mathbb{R}^2}\hat p_\mu(\xi_1,\xi_2)\,\overline{\hat p^{(v)}_{\mu'}(\xi_1,\xi_2)}\,d\xi_1\,d\xi_2,$$

where the functions $\hat p_\mu$ and $\hat p^{(v)}_{\mu'}$ satisfy the estimates (4.18) and (4.19), respectively. It will be sufficient to estimate the integral in the region $\Gamma=\{|\xi_2|\le|\xi_1|\}$. The estimate in the complementary region follows by interchanging the roles of ξ_1 and ξ_2.

As in the proof of Proposition 4.8, for $j'\ge j$, we split the integral as

$$I(\mu,\mu')=I_1(\mu,\mu')+I_2(\mu,\mu')+I_3(\mu,\mu')+I_4(\mu,\mu').$$

Our estimates will make use of the following inequality, valid for $|a|,|b|\le 1$:

$$|1-a\,b|\ge|a-b|. \tag{4.52}$$

Observe that, unlike Proposition 4.8 the case $j'<j$ is not symmetric to the case $j\ge j'$ and will be discussed later.

For $\xi_1\ge 0$ (the case $\xi_1<0$ is similar), using the change of variables $\xi_2=u\,\xi_1$ we have:

$$\begin{aligned}
I_1(\mu,\mu') &\le C_N\int_{2^{2j'}}^{\infty}\int_{-\infty}^{\infty}\frac{2^{-\frac{3}{2}(j+j')}}{\left(1+2^{-2j}\,\xi_1+2^{-j}\,|\xi_2-\ell\,2^{-j}\,\xi_1|\right)^N\,\left(1+2^{-2j'}\,|\xi_2|+2^{-j'}\,|\xi_1-\ell'\,2^{-j'}\,\xi_2|\right)^N}\,d\xi_2\,d\xi_1\\
&= C_N\int_{2^{2j'}}^{\infty}\int_{-\infty}^{\infty}\frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{\left(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|\right)^N\,\left(1+2^{-2j'}\,\xi_1|u|+2^{-2j'}\ell'\,\xi_1\,|u-\frac{2^{j'}}{\ell'}|\right)^N}\,du\,d\xi_1\\
&\le C_N\int_{2^{2j'}}^{\infty}\frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\xi_1)^N}\int_{-\infty}^{\infty}\frac{1}{\left(1+\alpha|u-\ell\,2^{-j}|\right)^N\,\left(1+\alpha'|u-\ell'\,2^{-j'}|\right)^N}\,du\,d\xi_1,
\end{aligned}$$

where $\alpha = \frac{2^{-j}\xi_1}{1+2^{-2j}\xi_1}$ and $\alpha' = 2^{-2j'}\,\ell'\,\xi_1$.

If $\alpha \geq \alpha'$, using (4.40) and (4.52) we have

$$\begin{aligned}
I_1(\mu,\mu') &\leq C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,2^{j}}{(1+2^{-2j}\xi_1)^{N-1}\left(1+2^{-2j'}\,\ell'\,\xi_1|\frac{2^{j'}}{\ell'}-\ell\,2^{-j}|\right)^N}\,d\xi_1 \\
&= C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{1}{2}(j+3j')}}{(1+2^{-2j}\xi_1)^{N-1}\,(1+2^{-j'}\,\xi_1|1-\ell'\,2^{-j'}\ell\,2^{-j}|)^N}\,d\xi_1 \\
&\leq C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{1}{2}(j+3j')}}{(1+2^{-2j}\xi_1)^{N-1}\,(1+2^{-j'}\,\xi_1|\ell'\,2^{-j'}-\ell\,2^{-j}|)^N}\,d\xi_1 \\
&\leq C_N\,(1+2^{j'}\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^{-N}\int_{2^{2j'}}^{\infty} \frac{2^{-\frac{1}{2}(j+3j')}}{(1+2^{-2j}\xi_1)^{N-1}}\,d\xi_1 \\
&\leq C_N\,2^{-\frac{3}{2}(j'-j)}\,2^{-2(j'-j)(N-2)}\,(1+2^{j'}\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^{-N}.
\end{aligned}$$

If $\alpha < \alpha'$, using (4.40) and (4.52) we have

$$\begin{aligned}
I_1(\mu,\mu') &\leq C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\xi_1)^{N}\,\alpha'\left(1+\alpha\,|\frac{2^{j'}}{\ell'}-\ell\,2^{-j}|\right)^N}\,d\xi_1 \\
&\leq C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\xi_1)^{N}\,\alpha\,(1+\alpha\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^N}\,d\xi_1 \\
&\leq C_N \int_{2^{2j'}}^{\infty} \frac{2^{-\frac{1}{2}(j+3j')}}{(1+2^{-2j}\xi_1)^{N-1}\,(1+2^{j}\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^N}\,d\xi_1 \\
&\leq C_N\,2^{-\frac{3}{2}(j'-j)}\,2^{-2(j'-j)(N-2)}\,(1+2^{j'}\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^{-N}.
\end{aligned}$$

For I_2, when $\xi_1 \geq 0$, using the change of variables $\xi_2 = u\,\xi_1$ we have:

$$\begin{aligned}
I_2(\mu,\mu') &\leq C_N \int_{2^{2j}}^{2^{2j'}} \int_{|\xi_2|\leq\xi_1} \frac{2^{-\frac{3}{2}(j+j')}\,2^{-2j'N}\,(1+|\xi_2|)^N}{(1+2^{-2j}\,\xi_1+2^{-j}\,|\xi_2-\ell\,2^{-j}\,\xi_1|)^N\,(1+2^{-2j'}\,|\xi_2|+2^{-j'}\,|\xi_1-\ell'\,2^{-j'}\,\xi_2|)^N}\,d\xi_2\,d\xi_1 \\
&\leq C_N \int_{2^{2j}}^{2^{2j'}} \int_{-1}^{1} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1\,2^{-2j'N}\,(1+\xi_1|u|)^N}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-j'}\,\xi_1|1-\ell'\,2^{-j'}\,u|)^N}\,du\,d\xi_1 \\
&\leq C_N \int_{2^{2j}}^{2^{2j'}} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1\,2^{-2j'N}\,(1+\xi_1|u|)^N}{(1+2^{-2j}\xi_1)^N} \int_{-\infty}^{\infty} \frac{1}{(1+\alpha|u-\ell\,2^{-j}|)^N\,(1+\alpha'|u-\ell'\,2^{-j'}|)^N}\,du\,d\xi_1, \\
&\leq C_N \int_{2^{2j}}^{2^{2j'}} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1^{N+1}\,2^{-2j'N}}{(1+2^{-2j}\xi_1)^N} \int_{-\infty}^{\infty} \frac{1}{(1+\alpha|u-\ell\,2^{-j}|)^N\,(1+\alpha'|u-\ell'\,2^{-j'}|)^N}\,du\,d\xi_1.
\end{aligned}$$

Now we proceed as for I_1. When $\alpha \geq \alpha'$, using (4.40) and (4.52), and then proceeding as in Proposition 4.7

we have:

$$\begin{aligned}
I_2(\mu,\mu') &\le C_N \int_{2^{2j}}^{2^{2j'}} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1^N\,2^{-2j'N}\,2^j}{(1+2^{-2j}\xi_1)^{N-1}\left(1+2^{-2j'}\,\ell'\,\xi_1|\frac{2^{j'}}{\ell'}-\ell\,2^{-j}|\right)^N}\,d\xi_1\\
&\le C_N \int_{2^{2j}}^{2^{2j'}} \frac{2^{-\frac{1}{2}(j+3j')}\,\xi_1^N\,2^{-2j'N}}{(1+2^{-2j}\xi_1)^{N-1}\,(1+2^{-j'}\,\xi_1|\ell' 2^{-j'}-\ell\,2^{-j}|)^N}\,d\xi_1\\
&\le C_N\,2^{-\frac{1}{2}(j+3j')}\,(1+2^{2j-j'}\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^{-N}\int_{2^{2j}}^{2^{2j'}}\frac{\xi_1^N\,2^{-2j'N}}{(1+2^{-2j}\xi_1)^{N-1}}\,d\xi_1\\
&\le C_N\,2^{-\frac{1}{2}(j-j')}\,2^{-(j'-j)N}\,(1+2^j\,|\ell'\,2^{-j'}-\ell\,2^{-j}|)^{-N}.
\end{aligned}$$

The case where $\alpha<\alpha'$ is similar.

The integrals I_3 and I_4 are treated as in Proposition 4.7.

We next consider the case $j'<j$.

For $\xi_1\ge 0$ (the case $\xi_1<0$ is similar), using the change of variables $\xi_2=u\,\xi_1$ and (4.52) we have:

$$\begin{aligned}
I_1(\mu,\mu') &\le C_N \int_{2^{2j}}^{\infty}\int_{-\infty}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1|u|+2^{-j'}\,\xi_1\,|1-\ell' 2^{-j'}u|)^N}\,du\,d\xi_1\\
&\le C_N \int_{2^{2j}}^{\infty}\int_{|u|\ge 1/2} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|u-\ell' 2^{-j'}|)^N}\,du\,d\xi_1\\
&\quad + C_N \int_{2^{2j}}^{\infty}\int_{|u|<1/2} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1|u|+2^{-j'}\,\xi_1\,|1-\ell' 2^{-j'}u|)^N}\,du\,d\xi_1
\end{aligned}$$

Observe that the first of the two integrals in the last sum can be estimated similarly to the integral I_1 in Proposition 4.7. Since $|u|<1/2$ (and, thus, $|1-\ell' 2^{-j'}u|>1/2$), the second integral is controlled by

$$\begin{aligned}
C_N \int_{2^{2j}}^{\infty} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1}{(1+2^{-2j}\,\xi_1)^N\,(1+2^{-j'}\,\xi_1)^N}\,d\xi_1 &\le C_N\,2^{\frac{1}{2}(5j-3j')}\,2^{-(2j-j')N}\int_{2^{2j}}^{\infty}\frac{y}{(1+y)^N}\,dy\\
&\le C_N\,2^{\frac{1}{2}(5j-3j')}\,2^{-(j-j')N}\,2^{-jN}.
\end{aligned}$$

For I_2, when $\xi_1\ge 0$, using the change of variables $\xi_2=u\,\xi_1$ we have:

$$\begin{aligned}
I_2(\mu,\mu') &\le C_N \int_{2^{2j'}}^{2^{2j}}\int_{-1}^{1} \frac{2^{-\frac{3}{2}(j+j')}\,\xi_1\,2^{-2jN}\,(1+\xi_1|u|)^N}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1|u|+2^{-j'}\,\xi_1|1-\ell'\,2^{-j'}\,u|)^N}\,du\,d\xi_1\\
&\le C_N \int_{2^{2j'}}^{2^{2j}}\int_{1\ge|u|\ge 1/2} \frac{2^{-\frac{3}{2}(j+j')}\,2^{-2jN}\,\xi_1^{N+1}}{(1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1+2^{-j'}\,\xi_1\,|u-\ell' 2^{-j'}|)^N}\,du\,d\xi_1\\
&\quad + C_N \int_{2^{2j'}}^{2^{2j}}\int_{|u|<1/2} \frac{2^{-\frac{3}{2}(j+j')}\,2^{-2jN}\,\xi_1^{N+1}}{(1+2^{-2j}\,\xi_1+2^{-j}\,\xi_1|u-\ell\,2^{-j}|)^N\,(1+2^{-2j'}\,\xi_1|u|+2^{-j'}\,\xi_1\,|1-\ell' 2^{-j'}u|)^N}\,du\,d\xi_1
\end{aligned}$$

Observe that the first of the two integrals in the last sum can now be estimated similarly to the integral I_2, case $j'\ge j$. Since $|u|<1/2$ (and, thus, $|1-\ell' 2^{-j'}u|>1/2$), the second integral is controlled by

$$\begin{aligned}
C_N \int_{2^{2j'}}^{2^{2j}} \frac{2^{-\frac{3}{2}(j+j')}\,2^{-2jN}\,\xi_1^{N+1}}{(1+2^{-j'}\,\xi_1)^N}\,d\xi_1 &= C_N\,2^{-\frac{1}{2}(3j-j')}\,2^{-(2j-j')N}\int_{2^{2j'}}^{2^{2j}}\frac{y^{N+1}}{(1+y)^N}\,dy\\
&\le C_N\,2^{\frac{1}{2}(3j+j')}\,2^{-(j-j')N}\,2^{-jN}.
\end{aligned}$$

The integrals I_3 and I_4 are treated in a similar way.

The remaining part of the proof, involving the estimate for the spatial decay, is very similar to the one in Proposition 4.8. □

4.3 Analysis of the operator $T_\mu^{(2)}$

To analyze the operator $T_\mu^{(1)}$ in Section 4.1 we have taken advantage of the compact frequency support of the shearlets $\{\psi_\mu : \mu \in \mathcal{M}\}$. In contrast, to analyze the operator $T_\mu^{(2)}$ it will be convenient to introduce a family of shearlet-like functions with compact support in the space domain. Using this analyzing family, it will be possible to introduce an atomic decomposition

$$f(x) = \sum_\mu \nu_\mu \, \rho_\mu(x),$$

for functions $f \in L^2(R^2)$, where the *shearlet atoms* ρ_μ have compact support and satisfy certain regularity and vanishing moments conditions. Their precise definition will be given after Proposition 4.18. Recall that the notion of atomic decomposition is standard in harmonic analysis (see, for example, [8]).

We have the following result.

Theorem 4.10. *Let $\{\rho_{\mu'} : \mu' \in \mathcal{M}\}$ be a family of shearlet atoms with regularity R. For each $\mu' \in \mathcal{M}$, the operator $T_\mu^{(2)}$ maps $\rho_{\mu'}$ into a shearlet atom $m_{h_\mu(\mu')}$ with the same regularity R, uniformly over $\mu' \in \mathcal{M}$.*

As in Theorem 3.1, for each $\mu \in \mathcal{M}$, the function h_μ is a bijective mapping on $\mathcal{M}$. The precise definition of h will be given later in this section.

We will construct a family of shearlet-like functions with compact support of the form

$$\psi_{ast}(x) = |\det A_a|^{-1/2} \, \psi(A_a^{-1} \, B_s^{-1}(x - t),$$

where

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad B_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix},$$

and a, s, t are continuous parameters satisfying: $0 < a \le 1$, $|s| \le 2$, $t \in \mathbb{R}$. Notice that, in the Fourier domain, $\hat{\psi}_{ast}(\xi) = a^{3/4} \, e^{-2\pi i \xi t} \, \psi(\xi B_s A_a)$. The following construction adapts several ideas from H. Smith [19].

Definition 4.11. *We say that a Schwartz function ϕ has k vanishing moments in the x_1 direction if ϕ can be expressed as*

$$\phi(x) = \partial_{x_1}^k \tilde{\phi}(x)$$

for some other Schwartz function $\tilde{\phi}$.

Observe that if ϕ has a certain number of vanishing moments in the x_1 direction, then $\hat{\phi}(\xi) = (2\pi i \xi_1)^k \widehat{\tilde{\phi}}(\xi)$, and, thus, $\hat{\phi}(0, \xi_2) = 0$. This shows that $\hat{\phi}(\xi)$ is concentrated along the ξ_1 axis way from the origin. As a consequence, $\hat{\phi}(\xi B_s A_a)$ is concentrated in elongated regions, symmetric with respect to the origin, along the direction $\xi_2 = s\xi_1$. These regions become increasingly elongated as $a \to 0$.

Let $\mu(x)$ be a Schwartz function with k vanishing moments in the x_1 direction. That is

$$\mu(x) = \partial_{x_1}^k \phi(x),$$

where ϕ is another Schwartz function. We assume that $\hat{\phi}(\pm 1, 0) \neq 0$.

We use the notation

$$\mu_a(x) = a^{-3/2} \mu(A_a^{-1} x) = a^{-3/2} \mu(a^{-1} x_1, a^{-1/2} x_2)$$

and

$$\mu_{as}(x) = a^{-3/2} \mu(A_a^{-1} B_s^{-1} x) = \mu_a(x_1 - s x_2, x_2) = a^{-3/2} \mu(a^{-1}(x_1 - s x_2), a^{-1/2} x_2)$$

Since $\hat{\mu}(\xi) = (2\pi i \xi_1)^k \hat{\phi}(\xi)$, it follows that

$$\hat{\mu}_a(\xi) = \hat{\mu}(\xi A_a) = \hat{\mu}(a\xi_1, a^{1/2} \xi_2) = (2\pi i a \xi_1)^k \, \hat{\phi}(a\xi_1, a^{1/2} \xi_2),$$

and

$$\hat{\mu}_{as}(\xi) = \hat{\mu}(\xi B_s A_a) = \hat{\mu}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)) = (2\pi i a\xi_1)^k\, \hat{\phi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)).$$

Let

$$p_\beta(\xi) = \int_{-2}^{2}\int_0^1 a^{-\beta}\, |\hat{\mu}_{as}(\xi)|^2 \frac{da}{a} ds,$$

and Γ be the region in the frequency plane:

$$\Gamma = \left\{(\xi_1,\xi_2) \in \mathbb{R}^2 :\, |\xi_1| \geq 1,\, \left|\frac{\xi_2}{\xi_1}\right| \leq 1\right\}.$$

Lemma 4.12. *For each $\xi \in \Gamma$ and $k > \beta/2$, we have that*

$$\int_{a\leq 1} a^{-\beta}\, |\hat{\mu}_a(\xi)|^2 \, \frac{da}{a} \leq C\, |\xi|^\beta \, \left(1 + \frac{|\xi_2|^2}{|\xi|}\right)^{\beta-2k}$$

Proof. Since ϕ is a Schwartz function, for any $N > 0$ we have:

$$\left|\hat{\phi}(a\,\xi_1, \sqrt{a}\,\xi_2)\right| \leq C_N\, \left(1 + a\,|\xi_1| + \sqrt{a}\,|\xi_2|\right)^{-N}.$$

This implies that

$$|\hat{\mu}_a(\xi)|^2 \leq C_N\, |a\,\xi_1|^{2k}\, (1 + a\,|\xi|)^{-N}\, \left(1 + a\,|\xi_2|^2\right)^{-N/2}.$$

Using the last inequality and the change of variables $u = a\,|\xi|$, it follows that

$$\begin{aligned}
\int_{a\leq 1} a^{-\beta}\, |\hat{\mu}_a(\xi)|^2 \, \frac{da}{a} &\leq C_N\, |\xi|^\beta \int_{a\leq 1} (a\,\xi)^{2k-\beta}\, (1 + a\,|\xi|)^{-N} \left(1 + a\,|\xi| \frac{|\xi_2|^2}{|\xi|}\right)^{-N/2} \frac{da}{a} \\
&= C_N\, |\xi|^\beta \int_{u\leq|\xi|} u^{2k-\beta}\, (1+u)^{-N} \left(1 + u\, \frac{|\xi_2|^2}{|\xi|}\right)^{-N/2} \frac{du}{u}. \qquad (4.53)
\end{aligned}$$

Using the change of variables $v = u\, \frac{|\xi_2|^2}{|\xi|}$, for $\xi_2 \neq 0$, we have

$$\begin{aligned}
\int_{a\leq 1} a^{-\beta}\, |\hat{\mu}_a(\xi)|^2 \, \frac{da}{a} &\leq C_N\, |\xi|^\beta \int_{v\leq|\xi_2|^2} \left(v\, \frac{|\xi|}{|\xi_2|^2}\right)^{2k-\beta} \left(1 + v\, \frac{|\xi|}{|\xi_2|^2}\right)^{-N} (1+v)^{-N/2}\, \frac{dv}{v} \\
&\leq C_N\, |\xi|^\beta \left(\frac{|\xi|}{|\xi_2|^2}\right)^{2k-\beta} \int_{v\leq|\xi_2|^2} (v)^{2k-\beta-1}\, (1+v)^{-N/2}\, dv \\
&\leq C\, |\xi|^\beta \left(\frac{|\xi|}{|\xi_2|^2}\right)^{2k-\beta}, \qquad (4.54)
\end{aligned}$$

where, in the last inequality, we have used $k > \beta/2$ and N large. We still have to consider the case $\xi_2 = 0$. However, it is easy to see from (4.53) that, for all $\xi \in \Gamma$,

$$\int_{a\leq 1} a^{-\beta}\, |\hat{\mu}_a(\xi)|^2 \, \frac{da}{a} \leq C\, |\xi|^\beta. \qquad (4.55)$$

Thus, the Lemma follows from (4.54) and (4.55). □

Lemma 4.13. *For each $\xi \in \Gamma$ and $k > \frac{1}{2}(\beta + 1/2)$, we have that*

$$\int_{|s|\leq 2}\int_{a\leq 1} a^{-\beta}\, |\hat{\mu}_a(\xi)|^2 \, \frac{da}{a} \, ds \leq C\, |\xi|^{\beta-1/2},$$

where $C > 0$ is independent of ξ.

Proof. Using Lemma 4.12, we have that

$$\int_{a\le 1} a^{-\beta}\,|\hat\mu_{as}(\xi)|^2\,\frac{da}{a} = \int_{a\le 1} a^{-\beta}\,|\hat\mu_{a}(\xi B_s)|^2\,\frac{da}{a} \le C\,|\xi B_s|^\beta \left(1+|\xi B_s|\,\left(\frac{|(\xi B_s)_2|}{|\xi B_s|}\right)^2\right)^{\beta-2k} \tag{4.56}$$

Since $|s|\le 2$, then $|\xi B|\simeq|\xi|$. Also, for $\xi\in\Gamma$ and $|s|\le 2$, we have

$$\left(\frac{|(\xi B_s)_2|}{|\xi B_s|}\right)^2 = \frac{|\xi_2 - s\xi_1|^2}{\xi_1^2+(\xi_2-s\xi_1)^2} = \frac{|\frac{\xi_2}{\xi_1}-s|^2}{1+(\frac{\xi_2}{\xi_1}-s)^2} \simeq (\tfrac{\xi_2}{\xi_1}-s)^2.$$

Thus, from (4.56) we have that

$$\int_{a\le 1} a^{-\beta}\,|\hat\mu_{as}(\xi)|^2\,\frac{da}{a} \le C\,|\xi|^\beta\,\left(1+|\xi|\,\left(\tfrac{\xi_2}{\xi_1}-s\right)^2\right)^{\beta-2k} \tag{4.57}$$

Using (4.57) and the change of variables $t=|\xi|^{1/2}\,(\frac{\xi_2}{\xi_1}-s)$, it follows that for $\xi\in\Gamma$ and $k>\frac12(\beta+1/2)$:

$$\begin{aligned}\int_{|s|\le 2}\int_{a\le 1} a^{-\beta}\,|\hat\mu_{as}(\xi)|^2\,\frac{da}{a}\,ds &\le C\,|\xi|^\beta \int_{|s|\le 2}\left(1+|\xi|\,\left(\tfrac{\xi_2}{\xi_1}-s\right)^2\right)^{\beta-2k} ds\\ &\le C\,|\xi|^\beta\,|\xi|^{-1/2}\int_{\mathbb{R}}\left(1+t^2\right)^{\beta-2k}\,ds\\ &\le C\,|\xi|^{\beta-1/2}. \qquad \square\end{aligned}$$

Lemma 4.14. *Assume that $\hat\phi(\pm1,0)\neq 0$. Then, for each $\xi\in\Gamma$*

$$p_\beta(\xi)\ge C\,|\xi|^{\beta-1/2},$$

where $C>0$ is independent of ξ.

Proof. We will only consider the case $\xi_1>0$, under the assumption that $\hat\phi(1,0)\neq 0$. The situation where $\xi_1<0$ is treated in a similar way, using the assumption that $\hat\phi(-1,0)\neq0$.

Using the change of variables $u=a\,\xi_1$, we have

$$\begin{aligned}p_\beta(\xi) &= \int_{|s|\le 2}\int_{a\le 1} a^{-\beta}\,(a\,\xi_1)^{2k}\,|\hat\phi(a\,\xi_1,\sqrt{a}\,(\xi_2-s\,\xi_1))|^2\,\frac{da}{a}\,ds\\ &= \int_{|s|\le 2}\int_{u\le \xi_1} \xi_1^\beta\,u^{2k-\beta}\,|\hat\phi(u,\sqrt{u\xi_1}\,(\tfrac{\xi_2}{\xi_1}-s))|^2\,\frac{du}{u}\,ds.\end{aligned}$$

Since $\hat\phi(1,0)\neq0$, we can find a $\delta>0$ such that $\hat\phi(\omega_1,\omega_2)\ge C>0$ for $|\omega_1-1|<\delta$, $|\omega_2|<\delta$. Since $\xi\in\Gamma$, then $\xi_1>1$ and therefore, by (4.57), we have

$$p(\xi)\ge C(\delta)\,\xi_1^\beta\int_{E_\delta}\int_{1-\delta}^1 u^{2k-\beta}\,\frac{du}{u}\,ds,$$

where $E_\delta=\{s:\,|s|\le 2,\,|s-\frac{\xi_2}{\xi_1}|\le C(\delta)\,\xi_1^{-1/2}\}$ with $m(E_\delta)\ge C(\delta)\,\xi_1^{-1/2}$. Thus, it follows that $p(\xi)\ge C\,\xi_1^{\beta-1/2}$ for $\xi\in\Gamma$. $\square$

We can show that there choices of ϕ such that, for $\xi\in\Gamma$ and $k>\frac12(\beta+1/2)$,

$$C_1\,|\xi|^{\beta-1/2}\le p_\beta(\xi)\le C_2\,|\xi|^{\beta-1/2},$$

for $C_1, C_2 > 0$. In fact, let

$$\psi(t) = \begin{cases} \frac{1}{1-\frac{1}{16}t^2}, & \text{if } |t| < 1/4 \\ 0, & \text{if } |t| \geq 1/4, \end{cases}$$

and define $\phi(x_1, x_2) = \psi(x_1)\,\psi(x_2)$. Then it is easy to verify that

(i) $\phi \in \mathcal{S}(\mathbb{R}^2)$;

(ii) $\operatorname{supp}\phi \subset [-\frac{1}{4}, \frac{1}{4}]^2$;

(iii) $\hat{\phi}(\pm 1, 0) \neq 0$.

To verify (iii) observe that

$$\begin{aligned} \hat{\phi}(1,0) = \hat{\psi}(1)\,\hat{\psi}(0) &= \int_{-1/4}^{1/4} e^{-2\pi i t}\,\psi(t)\,dt \int_{-1/4}^{1/4} \psi(t)\,dt \\ &= \int_{-1/4}^{1/4} \cos(2\pi t)\,\psi(t)\,dt \int_{-1/4}^{1/4} \psi(t)\,dt \neq 0. \end{aligned}$$

Lemma 4.15. *For each $\xi \in \Gamma$ and $k > \frac{1}{2}(\beta + 1/2)$, we have*

$$\left|\partial_\xi^\alpha\, p_\beta(\xi)\right| \leq C\,|\xi|^{\beta - \frac{1}{2} - \frac{|\alpha|}{2}},$$

where $C > 0$ is independent of ξ.

Proof. By direct calculation:

$$\partial_{\xi_2}\, p_\beta(\xi) = \int_{|s|\leq 2}\int_{a\leq 1} a^{-\beta+1/2}\,(a\,\xi_1)^{2k}\,\partial_{\eta_2}\left(|\hat{\phi}(a\,\xi_1, \sqrt{a}\,(\xi_2 - s\,\xi_1))|^2\right)\frac{da}{a}\,ds.$$

This show that $\partial_{\xi_2}\, p_\beta(\xi)$ behaves essentially as $p_{\beta-1/2}(\xi)$ and, by Lemma 4.13,

$$|\partial_{\xi_2}\, p_\beta(\xi)| \leq C\,|\xi|^{\beta-1}, \text{ for } \xi \in \Gamma,$$

provided $k > \frac{1}{2}(\beta + 1/2)$. By repeating the partial integration we have that

$$\left|\partial_{\xi_2}^N\, p_\beta(\xi)\right| \leq C\,|\xi|^{\beta-(N+1)/2}, \text{ for } \xi \in \Gamma.$$

In order to estimate the partial derivative with respect to ξ_1, it will be useful to write the function $p_\beta(\xi)$ in the following form. For $\xi_1 \geq 0$ (the case $\xi_1 < 0$ is similar), using the change of variables $a\xi_1 = u$, we have

$$\begin{aligned} p_\beta(\xi) &= \int_{|s|\leq 2}\int_0^1 a^{-\beta}\,(a\,\xi_1)^{2k}\,|\hat{\phi}(a\,\xi_1, \sqrt{a}\,(\xi_2 - s\,\xi_1))|^2\,\frac{da}{a}\,ds \\ &= \xi_1^\beta \int_{|s|\leq 2}\int_0^{\xi_1} (u)^{2k-\beta}\,|\hat{\phi}(u, u^{\frac{1}{2}}\,\xi_1^{\frac{1}{2}}(\frac{\xi_2}{\xi_1} - s\,\xi_1))|^2\,\frac{du}{u}\,ds. \end{aligned}$$

From the last expression a direct computation gives

$$\partial_{\xi_1}\, p_\beta(\xi) = I_1(\xi) + I_2(\xi) + I_3(\xi),$$

where

$$\begin{aligned} I_1(\xi) &= \beta\,\xi_1^{\beta-1}\int_{|s|\leq 2}\int_0^{\xi_1}(u)^{2k-\beta}\,|\hat{\phi}(u, u^{\frac{1}{2}}\,\xi_1^{\frac{1}{2}}(\frac{\xi_2}{\xi_1} - s\,\xi_1))|^2\,\frac{du}{u}\,ds; \\ I_2(\xi) &= \xi_1^\beta\int_{|s|\leq 2}(\xi_1)^{2k-\beta}\,|\hat{\phi}(\xi_1, \xi_1(\frac{\xi_2}{\xi_1} - s\,\xi_1))|^2\,\frac{du}{u}\,ds; \\ I_3(\xi) &= \xi_1^\beta\int_{|s|\leq 2}\int_0^{\xi_1}(u)^{2k-\beta}\,\partial_{\eta_2}\left(|\hat{\phi}(u, u^{\frac{1}{2}}\,\xi_1^{\frac{1}{2}}(\frac{\xi_2}{\xi_1} - s\,\xi_1))|^2\right)\partial_{\xi_1}\left(u^{\frac{1}{2}}\,\xi_1^{\frac{1}{2}}(\frac{\xi_2}{\xi_1} - s\,\xi_1))\right)\frac{du}{u}\,ds. \end{aligned}$$

Since $I_1(\xi) = \beta\, \xi_1^{-1}\, p_\beta(\xi)$, then, by Lemma 4.13, $|I_1(\xi)| \le C\, |\xi|^{\beta - \frac{1}{2} - 1}$ for $\xi \in \Gamma$. Also, since $\hat{\phi} \in \mathcal{S}(\mathbb{R}^2)$, we certainly have that I_2 satisfies the same type of decay. Thus, it only remains to control I_3. After computing the partial derivative with respect to ξ_1 inside the integral in I_3, we can write

$$I_3(\xi) = I_{3,1}(\xi) + I_{3,2}(\xi),$$

where

$$\begin{aligned} I_{3,1}(\xi) &= \xi_1^\beta \int_{|s|\le 2} \int_0^{\xi_1} (u)^{2k-\beta}\, \partial_{\eta_2} \left(|\hat{\phi}(u, u^{\frac{1}{2}}\, \xi_1^{\frac{1}{2}} (\frac{\xi_2}{\xi_1} - s\, \xi_1))|^2 \right) \frac{1}{2}\, u^{\frac{1}{2}}\, \xi_1^{-\frac{1}{2}} \left(\frac{\xi_2}{\xi_1} - s\, \xi_1 \right) \frac{du}{u}\, ds; \\ I_{3,2}(\xi) &= \xi_1^\beta \int_{|s|\le 2} \int_0^{\xi_1} (u)^{2k-\beta}\, \partial_{\eta_2} \left(|\hat{\phi}(u, u^{\frac{1}{2}}\, \xi_1^{\frac{1}{2}} (\frac{\xi_2}{\xi_1} - s\, \xi_1))|^2 \right) \left(\frac{-u^{\frac{1}{2}}\, \xi_1^{\frac{1}{2}} \xi_2}{\xi_1^2}) \right) \frac{du}{u}\, ds. \end{aligned}$$

Using the change of variables $u = a\xi_1$, we have that

$$I_{3,1}(\xi) = \frac{1}{2}\, \xi_1^{-1} \int_{|s|\le 2} \int_0^1 a^{-\beta}\, (a\, \xi_1)^{2k}\, \partial_{\eta_2} \left(|\hat{\phi}(a\, \xi_1, \sqrt{a}\, (\xi_2 - s\, \xi_1))|^2 \right) \sqrt{a}\, (\xi_2 - s\, \xi_1))\, \frac{da}{a}\, ds,$$

which shows that $I_{3,1}(\xi)$ behaves like $\frac{1}{2}\, \xi_1^{-1}\, p_\beta(\xi)$ and thus $|I_{3,1}(\xi)| \le C\, |\xi|^{\beta - \frac{1}{2} - 1}$ for $\xi \in \Gamma$. Similarly, using the change of variables $u = a\xi_1$, we have that

$$I_{3,2}(\xi) = -\frac{\xi_2}{\xi_1^{3/2}} \int_{|s|\le 2} \int_0^1 a^{-\beta}\, (a\, \xi_1)^{2k+1/2}\, \partial_{\eta_2} \left(|\hat{\phi}(a\, \xi_1, \sqrt{a}\, (\xi_2 - s\, \xi_1))|^2 \right) \frac{da}{a}\, ds,$$

which behaves like $-\frac{\xi_2}{\xi_1^{3/2}}\, p_\beta(\xi)$ (with k replaced by $k+1/4$). Thus $|I_{3,1}(\xi)| \le C\, |\xi|^{\beta - \frac{1}{2} - \frac{1}{2}}$ for $\xi \in \Gamma$. Therefore, by combining the estimated for $I_1(\xi)$, $I_2(\xi)$ and $I_3(\xi)$, we have that, for $\xi \in \Gamma$,

$$|\partial_{\xi_1}\, p_\beta(\xi)| \le C\, |\xi|^{\beta - \frac{1}{2} - \frac{1}{2}}.$$

The proof is completed by repeating the partial integration. □

We can now prove the following result, which is similar to [19, Lemma 2.11].

Proposition 4.16. *Let ψ be a Schwartz obeying $\hat{\psi}(\pm 1, 0) \ne 0$ and having at least one vanishing moment in the x_1 direction. Then there is a function $q(\xi)$ such that the following formula holds*

$$q(\xi) \int_{|s|\le 2} \int_{a\le 1} a^{3/2} \left| \hat{\psi}(\xi B_s A_a) \right|^2 \frac{da}{a^3}\, ds = 1, \qquad \textit{for } \xi \in \Gamma.$$

$q(\xi)$ is a smooth function satisfying $|\partial^\alpha q(\xi)| \le C\, |\xi|^{-\frac{|\alpha|}{2}}$ on Γ.

Proof. Observe that

$$I(\xi) = \int_{|s|\le 2} \int_{a\le 1} a^{3/2} \left| \hat{\psi}(\xi B_s A_a) \right|^2 \frac{da}{a^3}\, ds = \int_{|s|\le 2} \int_{a\le 1} a^{-1/2} \left| \hat{\psi}(\xi B_s A_a) \right|^2 \frac{da}{a}\, ds = p_{1/2}(\xi).$$

Thus, by Lemmata 4.13 and 4.14, there are constants $C_1, C_2 > 0$ such that

$$C_1 \le I(\xi) \le C_2, \text{ for } \xi \in \Gamma,$$

provided $k > 1/2$. Thus, for $\xi \in \Gamma$, the function

$$q(\xi) = \frac{1}{I(\xi)}$$

is well defined and smooth. Observe that

$$\partial_{\xi_i} q(\xi) = -\frac{I_{\xi_i}(\xi)}{I(\xi)^2}, \quad \partial^2_{\xi_i} q(\xi) = \frac{2I(\xi)\, I^2_{\xi_i}(\xi) - I_{\xi_i\xi_i}(\xi)\, I^2(\xi)}{I(\xi)^4}, \quad i = 1, 2$$

and so on for higher order derivatives. The estimate on $|\partial^\alpha q(\xi)|$ then follows by applying Lemma 4.15 with $\beta = 1/2$. $\square$

There many choices of ψ satisfying the assumptions of Proposition 4.16. In the following, we will choose ψ of the form

$$\psi(x_1, x_2) = \psi_1(x_1)\, \psi_2(x_2),$$

where ψ_1, and ψ_2 are C^∞ functions with compact support, satisfying $\operatorname{supp}\psi_1, \operatorname{supp}\psi_2 \subset [0,1]$. In addition, we assume that ψ_1 has vanishing moments up to order R, that is,

$$\int_{\mathbb{R}} \psi_1(x)\, x^k\, dx = 0, \quad k = 0, 1, \ldots, R.$$

We thus obtain the following reproducing formula.

Proposition 4.17. *Suppose that $\hat{f}$ vanishes outside Γ. Then we have the reproducing formula*

$$f(x) = \int_{\mathbb{R}^2} \int_{|s|\le 2} \int_{a \le 1} \langle q(D)f, \psi_{ast}\rangle\, \psi_{ast}(x)\, \frac{da}{a^3}\, ds\, dt, \tag{4.58}$$

where $(q(D)f)^\wedge(\xi) = q(\xi)\, \hat{f}(\xi)$ and $\psi_{ast}(x) = a^{-3/4}\, \psi(A_a^{-1} B_s^{-1}(x - t))$.

Proof. Using Proposition 4.16 we have:

$$\begin{aligned}
&\int_{\mathbb{R}^2} \int_{|s|\le 2} \int_{a \le 1} \langle q(D)f, \psi_{ast}\rangle\, \psi_{ast}(x)\, \frac{da}{a^3}\, ds\, dt \;= \\
&= \int_{\mathbb{R}^2} \int_{|s|\le 2} \int_{a \le 1} \left(\int_{\mathbb{R}^2} q(\xi)\, \hat{f}(\xi)\, \overline{\hat{\psi}(\xi B_a A_a)}\, e^{2\pi i \xi t}\, d\xi \right) \psi(A_a^{-1} B_s^{-1}(x-t))\, \frac{da}{a^3}\, ds\, dt \\
&= \int_{\mathbb{R}^2} \hat{f}(\xi) \int_{|s|\le 2} \int_{a \le 1} \left(\int_{\mathbb{R}^2} e^{2\pi i \xi t}\, \psi(A_a^{-1} B_s^{-1}(x-t))\, dt \right) q(\xi)\, \overline{\hat{\psi}(\xi B_a A_a)}\, \frac{da}{a^3}\, ds\, d\xi \\
&= \int_{\mathbb{R}^2} \hat{f}(\xi)\, e^{2\pi i \xi x} \int_{|s|\le 2} \int_{a \le 1} a^{3/2}\, q(\xi) \left| \hat{\psi}(\xi B_s A_a) \right|^2 \frac{da}{a^3}\, ds\, d\xi \\
&= \int_{\mathbb{R}^2} \hat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi = f(x). \qquad \square
\end{aligned}$$

The reproducing formula (4.58) can be written as an *atomic decomposition* where the integral is broken into several components associated with distinct regions. For $\mu = (j, \ell, k)$, let

$$Q_\mu = \{(a, s, t) :\, 2^{-2(j+1)} \le a < 2^{-2j},\, \ell 2^{-j} \le s < (\ell+1)2^{-j},\, A^{-j} B^{-\ell} t \in [k_1, k_1+1) \times [k_2, k_2+1)\}.$$

Observe that the sets Q_μ are disjoint and $\bigcup_{j\ge 0} \bigcup_{\ell=-2^{j+1}}^{2^{j+1}-1} \bigcup_{(k_a,k_2)\in\mathbb{Z}^2} Q_\mu = \{(a,s,t) : a \le 1, |s| \le 2, t \in \mathbb{R}^2\}$. Then, by breaking the integral (4.58) into components arising from different cells Q_μ we have:

$$f(x) = \sum_{j\ge 0} \sum_{\ell=-2^{j+1}}^{2^{j+1}-1} \sum_{(k_a,k_2)\in\mathbb{Z}^2} \nu_\mu\, \rho_\mu(x), \tag{4.59}$$

where

$$\rho_\mu(x) = \frac{1}{\nu_\mu} \iiint_{Q_\mu} \langle q(D)f, \psi_{ast}\rangle \, \psi_{ast}(x) \, \frac{da}{a^3} \, ds \, dt, \quad \nu_\mu = \left(\iiint_{Q_\mu} |\langle q(D)f, \psi_{ast}\rangle|^2 \, \frac{da}{a^3} \, ds \, dt \right)^{1/2}.$$

Also observe that

$$\iiint_{Q_\mu} \frac{da}{a^3} \, ds \, dt = \frac{1}{2} \, (2^4 - 1).$$

Define the functions α_μ by:

$$\alpha_\mu(x) = 2^{-\frac{3}{2}j} \, \rho_\mu(A^{-j} B^{-\ell}(x+k)) = 2^{-\frac{3}{2}j} \, \rho_\mu(B_{j,\ell} A^{-j}(x+k)).$$

They satisfy the following properties:

Proposition 4.18. *For all μ, the function α_μ satisfy the following properties:*

(i) *Compact support:* $\operatorname{supp} \alpha_\mu \subset C\,[-1,1]^2$*, where C is independent of μ and f.*

(ii) *Regularity: for each $\beta = (\beta_1, \beta_2)$, there is a constant C_β independent of μ and f such that*

$$\left| \partial_x^\beta \alpha_\mu(x) \right| \le C_\beta.$$

(iii) *Vanishing moments along the x_1 direction: for all $n = 0, 1, \ldots, R$,*

$$\int_{\mathbb{R}} \alpha_\mu(x_1, x_2) \, x_1^n \, dx_1 = 0.$$

We will refer to the elements α_μ satisfying (i)–(iii) as *atoms* with regularity R. The corresponding functions ρ_μ, given by

$$\rho_\mu(x) = 2^{\frac{3}{2}j} \, \alpha_\mu(B^\ell A^j x - k)$$

will be referred to as *shearlet atoms* with regularity R. It is an easy exercise to verify that a shearlet atom is also a shearlet molecule, according to Definition 4.1.

Proof of Proposition 4.18.
(i) By direct computation we have

$$\begin{aligned} \alpha_\mu(x) &= \frac{1}{\nu_\mu} \iiint_{Q_\mu} \langle q(D)f, \psi_{ast}\rangle \, 2^{-\frac{3}{2}j} \, a^{-\frac{3}{4}} \, \psi(A_a^{-1} B_s^{-1}(B_{j,\ell} A^{-j}(x+k) - t)) \, \frac{da}{a^3} \, ds \, dt \\ &= \frac{1}{\nu_\mu} \iiint_{Q_\mu} \langle q(D)f, \psi_{ast}\rangle \, |\det M|^{-1/2} \, \psi(M^{-1}(x - (\tau - k)) \, \frac{da}{a^3} \, ds \, dt \end{aligned}$$

where $\tau = B_{j,\ell} A^{-j} t = A^{-j} B^{-\ell} t$ and

$$M^{-1} = A_a^{-1} B_s^{-1} B_{j,\ell} A^{-j} = A_a^{-1} B_{(s+\ell 2^j)}^{-1} A^{-j} = \begin{pmatrix} a^{-1} 2^{-2j} & -a^{-1} 2^{-j} \, (s + \ell 2^{-j}) \\ 0 & a^{-1/2} 2^{-j} \end{pmatrix}.$$

Observe that, for $t \in Q_\mu$, then $\tau \in [k_1, k_1+1) \times [k_2, k_2+1)$ and, thus, $\operatorname{supp} \psi(M^{-1}(x - (\tau - k)) \subset \operatorname{supp} \psi(M^{-1}x) + [-1,1]^2$. In addition, since ψ is compactly supported with support on the set $[0,1]^2$, over the region Q_μ we have that $\operatorname{supp} \psi(M^{-1}x) \subset M\,[-1,1]^2 \subset [c_1, c_2) \times [d_1, d_2)$, where c_1, c_2, d_1, d_2 are constants independent of j and ℓ. This shows that $\alpha_m u(x)$ is compactly supported.

(ii) By the Cauchy–Schwarz inequality

$$\begin{aligned}
|\alpha_\mu(x)| &\le \frac{1}{\nu_\mu}\iiint_{Q_\mu} |\langle q(D)f,\psi_{ast}\rangle|\,|\det M|^{-1/2}\,\|\psi\|_\infty\,\frac{da}{a^3}\,ds\,dt\\
&\le \|\psi\|_\infty\,\frac{1}{\nu_\mu}\left(\iiint_{Q_\mu} |\langle q(D)f,\psi_{ast}\rangle|^2\,\frac{da}{a^3}\,ds\,dt\right)^{1/2}\left(\iiint_{Q_\mu} |\det M|^{-1}\,\frac{da}{a^3}\,ds\,dt\right)^{1/2}\\
&\le \frac{1}{2}\,(2^4-1)\,\|\psi\|_\infty,
\end{aligned}$$

where we have used the fact that $|\det M|^{-1} \le 1$ for $(a,s,t) \in Q_\mu$. Estimates for the derivatives of α_μ are obtained in a similar way.

(iii) By the assumptions on ψ, we have that

$$\int_{\mathbb{R}} \psi(x_1,x_2)\,x_1^n\,dx_1 = 0, \quad n = 0,1,\dots,R,$$

and, more generally, by changing the variable x_1 into $y = c\,x_1 + d\,x_2$ and expanding the polynomial $(y - d\,x_2)^k$, we have:

$$\int_{\mathbb{R}} \psi(c\,x_1 + d\,x_2, x_2)\,x_1^n\,dx_1 = 0, \quad n = 0,1,\dots,R.$$

In particular, since M is an upper triangular matrix, the last equality implies that

$$\int_{\mathbb{R}} \psi(M^{-1}x)\,x_1^n\,dx_1 = 0, \quad n = 0,1,\dots,R. \tag{4.60}$$

Using the Cauchy–Schwarz inequality we have that:

$$\begin{aligned}
&\left|\int_{\mathbb{R}} \alpha_\mu(x)\,x_1^n\,dx_1\right|\\
\le\ & \frac{1}{\nu_\mu}\iiint_{Q_\mu} |\langle q(D)f,\psi_{ast}\rangle|\,|\det M|^{-1/2}\int_{\mathbb{R}} \psi(M^{-1}(x-(\tau-k))\,x_1^n\,dx_1\,\frac{da}{a^3}\,ds\,dt\\
\le\ & \frac{1}{\nu_\mu}\left(\iiint_{Q_\mu} |\langle q(D)f,\psi_{ast}\rangle|^2\,\frac{da}{a^3}\,ds\,dt\right)^{1/2}\left(\iiint_{Q_\mu} |\det M|^{-1}\left|\int_{\mathbb{R}} \psi(M^{-1}(x-(\tau-k))\,x_1^n\,dx_1\right|^2\frac{da}{a^3}\,ds\,dt\right)^{1/2},
\end{aligned}$$

and this expression is equal to zero for $n = 0,1,\dots,R$ by (4.60). □

Proof of Theorem 4.10

Using the notation introduced above, we have that

$$\rho_{\mu'}(x) = 2^{\frac{3}{2}j'}\,\alpha_{\mu'}(B^{\ell'}A^{j'}(x - k_{\mu'})),$$

where $k_{\mu'} = A^{-j'}B^{-\ell'}k'$ and $a_{\mu'}(x)$ satisfies Proposition 4.18. By expanding the function ϕ_μ about the point $y_{\mu,\mu'} = \phi_\mu^{-1}(k_{\mu'})$, we can write $\phi_\mu(x)$ as

$$\phi_\mu(x) = k_{\mu'} + L_{\mu,\mu'}\,(x - y_{\mu,\mu'}) + g(x - y_{\mu,\mu'}),$$

where $L_{\mu,\mu'} = \nabla\phi_\mu(y_{\mu,\mu'})$ and g is a C^∞ function (with $g(x) = x^2 g'(x)$, and g' is a C^∞ function). Thus, with this notation:

$$T_\mu^{(2)}\,\rho_{\mu'}(x) = \rho_{\mu'}(\phi_\mu(x)) = 2^{\frac{3}{2}j'}\,\alpha_{\mu'}\left(B^{\ell'}A^{j'}\,(L_{\mu,\mu'}\,(x - y_{\mu,\mu'}) + g(x - y_{\mu,\mu'}))\right),$$

which shows that the function $T_\mu^{(2)}\,\rho_{\mu'}(x)$ is centered about the point $y_{\mu,\mu'}$.

Without loss of generality we can assume that $y_{\mu,\mu'} = 0$. For simplicity let us assume for now that $L_{\mu,\mu'} = I$, so that

$$\rho_{\mu'}(\phi(x)) = 2^{\frac{3}{2}j'}\,\alpha_{\mu'}\left(B^{\ell'}A^{j'}\,(x+g(x))\right) = 2^{\frac{3}{2}j'}\,\beta_{\mu}(B^{\ell'}A^{j'}x),$$

where $\beta_\mu(x) = \alpha_{\mu'}(x + B^{\ell}A^{j'}g(A^{-j'}B^{-\ell'}x))$. The general case $L_{\mu,\mu'} \neq I$ will be examined later.

By Proposition 4.18, $\operatorname{supp}\alpha_{\mu'} \subset C\,[-1,1]^2$, for some constant $C > 0$. Therefore, for each j', ℓ', the support of $\beta_{\mu'}$ is contained on a box of side length C. In particular, the support conditions imply that over the support of $\rho_{\mu'} \circ \phi_\mu$:

$$|x| \le C\,2^{-j'} \quad |g(x)| \le C\,2^{-2j'}, \quad |\partial^\gamma g(x)| \le C_\gamma\,2^{-j'}, \text{ for } |\gamma| = 1, \quad |\partial^\gamma g(x)| \le C_\gamma, \text{ for } |\gamma| > 1.$$

This implies that, over the support of $\rho_{\mu'} \circ \phi_\mu$ (in particular, for $|x| \le C$), for each γ:

$$\|B^{\ell'}A^{j'}\partial^\gamma g(A^{-j'}B^{-\ell'}x)\| \le C_\gamma.$$

This observation together with part (ii) of Proposition 4.18 implies that

$$|\partial^\gamma \beta_{\mu'}(x)| \le C_\gamma.$$

We will now examine the frequency decay of $\rho_{\mu'} \circ \phi_\mu$. Using the change of variables $x = \phi_\mu^{-1}(y)$ (hence: $dx = \frac{dy}{|J(\phi_\mu^{-1}(y))|}$), we have that

$$\begin{aligned} (\rho_{\mu'} \circ \phi_\mu)^\wedge(\xi) &= \int e^{-2\pi i x\xi}\rho_{\mu'}(\phi_\mu(x))\,dx \\ &= \int e^{-2\pi i \phi_\mu^{-1}(y)\xi}\rho_{\mu'}(y)\,\frac{dy}{|J(\phi_\mu^{-1}(y))|} \\ &= \int S_\xi(y)\,\rho_{\mu'}(y)\,dy, \end{aligned} \tag{4.61}$$

where $S_\xi(y) = \frac{e^{-2\pi i \phi_\mu^{-1}(y)\xi}}{|J(\phi_\mu^{-1}(y))|}$. Recall that the function $\rho_{\mu'}(y)$ is supported over a compact set, uniformly for all j', ℓ'. Thus, by the assumptions on Φ, over the support of $\rho_{\mu'}(y)$ we have that

$$|S_\xi(y)| \le C,$$

and, more generally, for all $m \ge 0$,

$$|\partial_1^m S_\xi(y)| \le C_m\,(1+|\xi|)^m.$$

By expanding $S_\xi(y_1, y_2)$ as a Taylor series about the point $y_1 = 0$, from the equation (4.61) we have

$$(\rho_{\mu'} \circ \phi_\mu)^\wedge(\xi) = \sum_{k=0}^{N-1}\int \frac{\partial_1^k S_\xi(0,y_2)}{k!}\left(\int y_1^k\,\rho_{\mu'}(y_1,y_2)\,dy_1\right)\,dy_2 + \int\!\!\int E_\xi(y_1,y_2)\,\rho_{\mu'}(y_1,y_2)\,dy_1\,dy_2, \tag{4.62}$$

where $E_\xi(y_1,y_2) = S_\xi(y_1,y_2) - \sum_{k<N}\frac{\partial_1^k S_\xi(0,y_2)}{k!}\,y_1^k$. By part (iii) of Proposition 4.18, we have that

$$\int y_1^k\,\rho_{\mu'}(y_1,y_2)\,dy_1 = 0, \quad k = 0,1,\dots,R.$$

In fact $\rho_{\mu'}(y) = 2^{\frac{3}{2}j'}\,\alpha_\mu(B^{\ell'}A^{j'}(y - k_{\mu'}))$ and the upper triangular matrix $B^{\ell'}A^{j'}$ does not affect the vanishing moments. Next observe that, on the support of $\rho_{\mu'}$ (where $|y_1| \le 2^{-2j'}$):

$$|E_\xi(y_1,y_2)| \le C_N\,(1+|\xi|)^N\,2^{-2j'N}.$$

Thus, using these observations, from (4.62) we have:

$$|(\rho_{\mu'} \circ \phi_\mu)^\wedge(\xi)| \le C_N\,(1+|\xi|)^N\, 2^{-2j'N} 2^{\frac{3}{2}j'}\, 2^{-3j'} \le C_N\, 2^{-\frac{3}{2}j'}\; 2^{-2j'N}\,(1+|\xi|)^N. \qquad (4.63)$$

Thus, we have shown that, for $L_{\mu,\mu'} = I$, the operator $T_\mu^{(2)}$ maps the atom $\rho_{\mu'}$ into another atom of same regularity having the same scale and orientation. Its original location $k_{\mu'}$ has been changed to $y_{\mu,\mu'}$.

Let us examine now what happens when $L_{\mu,\mu'}$ is included in the computation. In this case (still assuming $y_{\mu,\mu'} = 0$) we have that

$$\rho_{\mu'}(\phi_\mu(x)) = m_{\mu'}(L_{\mu,\mu'}x),$$

where

$$m_{\mu'}(x) = 2^{\frac{3}{2}j'}\,\alpha_{\mu'}\left(B^{\ell'} A^{j'}\,(x+\tilde g(x))\right), \quad \tilde g(x) = g(L^{-1}_{\mu,\mu'}x).$$

Observe that L_μ^{-1} is uniformly bounded by (3.9). Therefore it follows from the estimates deduced above in the case $L_{\mu,\mu'} = I$ that $m_{\mu'}$ is also a shearlet atom.

Thus, the operator $T_\mu^{(2)}$ maps the atom $\rho_{\mu'}$, centered at $(k_{\mu'}, \xi_{\mu'})$ in the phase space into another atom of same regularity centered at $(y_{\mu,\mu'}, \xi_\mu\, L_{\mu,\mu'})$. This induces the bijective index mapping h_μ on $\mathcal{M}$

$$T_\mu^{(2)}\,\rho_{\mu'}(x) = \rho_{\mu'}(\phi_\mu(x)) = \rho_{h_\mu(\mu')}(x). \quad \square$$

Remark 4.1. Recall the canonical transformation associated with the phase $\Phi(x,\xi)$ of the Fourier Integral Operator T:

$$y = \nabla_\xi \Phi(x,\xi), \quad \eta = \nabla_x \Phi(x,\xi).$$

Using the notation introduced above, let $\phi_\mu(x) = \nabla_\xi \Phi(x,\xi_\mu)$ and $\phi_\mu(y_{\mu,\mu'}) = k_{\mu'}$. Since $\Phi(x,\xi)$ is homogeneous of degree one in ξ, then

$$\Phi(x,\xi) = \xi\,\nabla_\xi \Phi(x,\xi)$$

and

$$\nabla_x \Phi(x,\xi) = \xi\,\nabla_x \nabla_\xi \Phi(x,\xi).$$

Using these observations and the definition of $L_{\mu,\mu'}$ we have that

$$\begin{aligned} \nabla_\xi \Phi(k_{\mu'},\xi_\mu) &= \phi_\mu(y_{\mu,\mu'}) = k_{\mu'}, \\ \nabla_x \Phi(k_{\mu'},\xi_\mu) &= \xi_\mu\,\nabla_x\nabla_\xi \Phi(k_{\mu'},\xi_\mu) = \xi_\mu\,L_{\mu,\mu'}. \end{aligned}$$

This shows that the action of the operator $T_\mu^{(2)}$ on the phase space coordinates of the shearlet atom $\rho_{\mu'}$ corresponds in fact to the canonical transformation:

$$(k_{\mu'}, \xi_{\mu'}) \to (y_{\mu,\mu'}, \xi_\mu\, L_{\mu,\mu'}).$$

That is, the index mapping h_μ on $\mathcal{M}$ is a bijective mapping induced by the canonical transformation. As shown by H. Smith in [19, Lemma 2.2], this map preserves the parabolic pseudo-distance.

Also observe that the dependence of ϕ_μ upon μ is not essential in the proof of Theorem 4.10. The only property that was used is that the derivatives of ϕ_μ are bounded uniformly with respect to μ.

4.4 Proof of Main Theorem

We can now prove Theorem 3.1.

Let ψ_{μ_0} and ψ_{μ_1} be two fixed shearlets. Without loss of generality, let us assume that they are both horizontal shearlets.

By Theorem 4.2, $m_{\mu_0} = T^{(1)}_{\mu_0}\, \psi_{\mu_0}$ is a shearlet molecule. Also observe that, using the atomic decomposition (4.59), we can expand ψ_{μ_1} as a superposition of (horizontal) shearlet atoms $\rho_{\mu'}$:

$$\psi_{\mu_1} = \sum_{\mu'} c_{\mu',\mu_1}\, \rho_{\mu'},$$

where

$$c_{\mu',\mu_1} = \left(\iiint_{Q_{\mu'}} |\langle q(D)\psi_{\mu_1}, \psi_{ast}\rangle|^2 \, \frac{da}{a^3}\, ds\, dt \right)^{1/2}. \tag{4.64}$$

Therefore, using these observations and Theorem 4.10 we have that

$$\begin{aligned}
\langle \psi_{\mu_1}, T\, \psi_{\mu_0}\rangle &= \langle \psi_{\mu_1}, T^{(2)}_{\mu_0} T^{(1)}_{\mu_0}\, \psi_{\mu_0}\rangle \\
&= \langle (T^{(2)}_{\mu_0})^*\, \psi_{\mu_1}, T^{(1)}_{\mu_0}\, \psi_{\mu_0}\rangle \\
&= \sum_{\mu'} c_{\mu',\mu_1}\, \langle (T^{(2)}_{\mu_0})^*\, \rho_{\mu'}, m_{\mu_0}\rangle \\
&= \sum_{\mu'} c_{\mu',\mu_1}\, \langle m_{\tilde{h}_{\mu_0}(\mu')}, m_{\mu_0}\rangle,
\end{aligned}$$

where $m_{\tilde{h}(\mu')}$ is a shearlet molecule and $\tilde{h} = h^{-1}$ is the inverse of the bijective index mapping h. Next observe that for every $N > 0$, there is a constant C_N such that:

$$|c_{\mu',\mu_1}| \le C_N\, \omega(\mu', \mu_1)^{-N}.$$

This follows by discretizing the integral (4.64) and noticing that, since $q(\xi)$ and all its derivatives are bounded (see Proposition 4.16), the argument of Proposition 4.8 can be applied to this integral.

By Proposition 4.7, the distance ω is invariant under the bijective index mapping h_{μ_0} induced by the canonical transformation: $\omega(\mu, \mu') \sim \omega(h_{\mu_0}(\mu), h_{\mu_0}(\mu'))$, uniformly over $\mu_0 \in \mathcal{M}$. Therefore, using Propositions 4.7 and 4.8, we have that for every $N > 0$, there is a constant C_N such that:

$$\begin{aligned}
|\langle \psi_{\mu_1}, T\, \psi_{\mu_0}\rangle| &\le \sum_{\mu'} |c_{\mu',\mu_1}| \left| \langle m_{\tilde{h}_{\mu_0}(\mu')}, m_{\mu_0}\rangle \right| \\
&\le C_N \sum_{\mu'} \omega(\mu', \mu_1)^{-N}\, \omega(\tilde{h}_{\mu_0}(\mu'), \mu_0)^{-N} \\
&\le C_N \sum_{\mu'} \omega(\tilde{h}_{\mu_0}(\mu'), \tilde{h}_{\mu_0}(\mu_1))^{-N}\, \omega(\tilde{h}_{\mu_0}(\mu'), \mu_0)^{-N} \\
&\le C_N \sum_{\mu'} \omega(\tilde{h}_{\mu_0}(\mu_1), \tilde{h}_{\mu_0}(\mu'))^{-N}\, \omega(\tilde{h}_{\mu_0}(\mu'), \mu_0)^{-N} \\
&\le C_N\, \omega(\tilde{h}_{\mu_0}(\mu_1), \mu_0)^{-N+1} \\
&\sim C_N\, \omega(\mu_1, h_{\mu_0}(\mu_0))^{-N+1}. \qquad \square
\end{aligned}$$

Acknowledgments

This authors are indebted to E. Candès, L. Demanet and G. Weiss for many helpful discussions.

References

[1] G. Beylkin, R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms, *Comm. Pure Appl. Math.* **44** (1992), 141–183.

[2] E. J. Candès, and L. Demanet, The curvelet representation of wave propagators is optimally sparse, *Comm. Pure Appl. Math.* **58** (2005), 1472–1528.

[3] E. J. Candès and D. L. Donoho, New tight frames of curvelets and optimal representations of objects with C^2 singularities, *Comm. Pure Appl. Math.* **56** (2004), 219–266.

[4] P. G. Casazza, The art of frame theory, *Taiwanese J. Math.*, **4** (2000), 129–201.

[5] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.

[6] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, and H.-G. Stark, The Uncertainty Principle Associated with the Continuous Shearlet Transform, to appear in Int. J. Wavelets Multiresolut. Inf. Process., (2007).

[7] G. Easley, D. Labate, and W. Lim Sparse Directional Image Representations using the Discrete Shearlet Transform, to appeat in *Appl. Comput. Harmon. Anal.* (2007).

[8] Frazier, M., B. Jawerth, and G. Weiss, *Littlewood-Paley Theory and the Study of Function Spaces*, CBMS Regional Conference Ser., 79, American Math. Society, 1991.

[9] L. Grafakos *Classical and Modern Fourier Analysis*, Prentice Hall, 2004.

[10] K. Gröchenig, *Foundations of Time–Frequency Analysis*, Birkäuser, Boston (2001).

[11] K. Guo, G. Kutyniok, and D. Labate, Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators, in: *Wavelets and Splines*, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), 189–201.

[12] K. Guo and D. Labate, Optimally Sparse Multidimensional Representation using Shearlets, SIAM J. Math. Anal., **39** (2007), 298–318.

[13] K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, Wavelets with composite dilations, *Electr. Res. Announc. of AMS* **10** (2004), 78–87.

[14] Guo, K., W. Lim, D. Labate, G. Weiss and E. Wilson, The theory of wavelets with composite dilations, in: *Harmonic Analysis and Applications*, C. Heil (ed.), pp. 231-249, Birkauser, 2006.

[15] K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, Wavelets with composite dilations and their MRA properties, *Appl. Comput. Harmon. Anal.* **20** (2006), 220–236.

[16] G. Kutyniok and D. Labate, Resolution of the Wavefront Set using Continuous Shearlets, preprint (2006).

[17] P. Lax, Asymptotic solutions of oscillatory initial value problems, *Duke Math. J.* **24** (1957) 627–646.

[18] Y. Meyer, R. Coifman *Wavelets, Calderón-Zygmund Operators and Multilinear Operators*, Cambridge Univ. Press, Cambridge, 1997.

[19] H.F. Smith, A Hardy space for Fourier integral operators, *J. Geom. Anal.* **8** (1998), 629–653.

[20] H.F. Smith, A parametrix construction for wave equations with $C_{1,1}$ coefficients, *Ann. Inst. Fourier, Grenoble* **48** (1998), 797–835.

[21] C. D. Sogge, *Fourier Integrals in Classical Analysis,* Cambridge University Press, Cambridge, 1993.

[22] E. M. Stein, *Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993.

[23] M. E. Taylor, *pseudodifferential Operators*, Princeton University Press, Princeton, NJ, 1981.

Submitted January 13, 2007; revised August 6, 2007.

Department of Mathematics, Missouri State University,
Springfield, MO 65804, USA
e-mail:kag026@smsu.edu

Department of Mathematics, North Carolina State University, Campus Box 8205
Raleigh, NC 27695, USA
e-mail: dlabate@math.ncsu.edu