Geometric Separation of Singularities using Combined Multiscale Dictionaries

Kanghui Guo[1] and Demetrio Labate[2]

January 25, 2014

Abstract

Several empirical results that have appeared in the literature during the last decade have shown that it is often possible to separate images and other multidimensional data into geometrically distinct constituents. A rigorous mathematical analysis of the *geometric separation* problem in the two-dimensional setting was recently introduced by Donoho and Kutyniok [4], who proposed a mathematical framework to separate point and smooth curve singularities in 2D images using a combined dictionary consisting of curvelets and wavelets. In this paper, we adapt their approach and introduce a novel argument to extend geometric separation to the three-dimensional setting. We show that it is possible to separate point and piecewise linear singularities in 3D using a combined dictionary consisting of shearlets and wavelets. Our new approach takes advantage of the microlocal properties of the shearlet transform and has the ability to handle singularities containing vertices and corner points, which could not be handled using the original arguments.

Key words and phrases: cluster coherence, geometric separation, ℓ^1 minimization, shearlets, sparse representations, wavelets.
AMS Mathematics Subject Classification: 42C15, 42C40.

1 Introduction

Data found in applications ranging from astronomy through remote sensing and biomedical imaging can be frequently modeled as superpositions of several distinct geometric components. Starck et al. [20, 21], in particular, proposed a very effective algorithmic approach, called Morphological Component Analysis (MCA), which assumes that a signal is the linear mixture of several constituents, the so-called *morphological components*, each one endowed with specific geometric properties. Under the assumption that the various morphological component are sufficiently distinct and that each one is sparsely represented in a specific basis but not in the other ones, MCA algorithms are very successful in separating the various components, as illustrated by several numerical demonstrations.

The idea of using combined-basis representations and take advantage of their sparsity properties has a long history in applied harmonic analysis and image processing. Some pioneering work about combined-basis representations can be found in the papers of Coifman and Wickerhauser [3] and Mallat and Zhang [18]. Another essential contribution to formalize these ideas was the introduction of Basis Pursuit [2], which established ℓ^1-norm minimization as an effective method to promote sparse representations from multiple bases. In more recent years, several other mostly empirical papers have further exploited this point of view and provided remarkable applications to problems from image processing. In addition to the work cited above, we recall, for example, the work in [5, 19, 23, 25][1].

[1]Department of Mathematics, Missouri State University, Springfield, Missouri 65804, USA. E-mail: KanghuiGuo@MissouriState.edu

[2]Department of Mathematics, University of Houston, Houston, Texas 77204, USA. E-mail: dlabate@math.uh.edu

[1]Note: as also remarked in [4], the geometric separation we consider here is rather different from the *separation of texture from smooth structure* where sparsity in the representation does not play a relevant role

1

In order to provide a rigorous mathematical understanding of the problem of data separation into geometrically distinct components, Donoho and Kutyniok [4] have recently proposed a theoretical framework for the geometric separation of point and curve singularities in 2D (cf. also related work in [15]). The underlying ansatz is that the success of MCA "stems from an interplay between geometric properties of objects to be separated and the harmonic analysis for singularities of various geometric types". Hence, as a mathematical idealization of a large class of two-dimensional objects, they consider images of the form $f = \mathcal{P} + \mathcal{T}$, where $\mathcal{P}$ denotes a collection of point-like singularities and $\mathcal{T}$ is a cartoon-like image, that is, a planar region enclosed by a smooth closed curve, producing an 'edge' singularity along a curve. What makes such images interesting is that wavelets provide very sparse representations of $\mathcal{P}$ while curvelets [1] (or sherlets [6]) provide very sparse representations of $\mathcal{T}$. However, neither wavelets nor curvelets (or sherlets) alone will provide a very sparse representation of $f = \mathcal{P} + \mathcal{T}$. In their paper, Donoho and Kutyniok consider the problem of deriving useful representations of f by combining both wavelets and curvelets and enforcing sparsity via ℓ_1 minimization. Their innovative viewpoint is derived from microlocal analysis, and, specifically, from the observation that while points and curves may overlap spatially, they are separated microlocally. In order to reveal this separation, the authors exploit the observation that the elements of the wavelet system are 'incoherent' to the curvelet system, meaning that they have limited overlap in phase space. This is used to derive an asymptotic estimate showing that, at very fine scales, the pointlike structure of f is essentially captured by the wavelet basis while curvelike structure of f is essentially captured by the curvelet basis.

In this paper, we extend the framework of geometric separation to the three-dimensional setting by using a combined dictionary of wavelets and shearlets [17] to separate point and piecewise linear singularities. In doing this, we adopt the important notion of *cluster coherence* introduced in [4]. However, to prove the geometric separation of the two types of singularities, we cannot use the same argument as [4] which maps the singularities in phase space, since this method does not generalize directly to 3D. Instead, we introduce a novel approach which is based on techniques previously developed by the authors and their collaborators for the geometric characterization of edge singularities in terms the shearlet transform [7, 12, 16]. In addition to solving the 3D geometric separation problem, an advantage of our new approach is the ability to deal with discontinuities including 3D edges and vertices, which could not be handled with the original arguments.

The rest of the paper is organized as follows. After setting some useful notation, we formulate the geometric separation problem and state our main theorem in Section 2. We present the proof of the main theorem in Section 3.

1.1 Notation.

In the following, we adopt the convention that $x \in \mathbb{R}^3$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^3$ (in the frequency domain) is a row vector, i.e., $\xi = (\xi_1, \xi_2, \xi_3)$. A vector x multiplying a matrix $A \in GL_3(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying A on the left is a row vector. Thus, $Ax \in \mathbb{R}^3$ and $\xi A \in \widehat{\mathbb{R}}^3$. The *Fourier transform* of $f \in L^1(\mathbb{R}^3)$ is defined as

$$\hat{f}(\xi) = \int_{\mathbb{R}^3} f(x) \, e^{-2\pi i \xi x} \, dx,$$

where $\xi \in \widehat{\mathbb{R}}^3$, and the inverse Fourier transform is

$$\check{f}(x) = \int_{\widehat{\mathbb{R}}^3} f(\xi) \, e^{2\pi i \xi x} \, d\xi.$$

We will use the notation $f \simeq g$ if there exist constants $0 < C_1 \leq C_2 < \infty$, independent of x, such that $C_1 \, g(x) \leq f(x) \leq C_2 \, g(x)$.

2 Geometric separation

In this paper, we show that a precise geometric separation of point and piecewise linear singularities in $\mathbb{R}^3$ can be achieved using an appropriate ℓ^1 minimization procedure over a combined dictionary of wavelets and shearlets.

Our setting of the problem extends to three-dimensions the setting from [4]. Namely, we consider point-like objects in $\mathbb{R}^3$, which we model as the distributions of the form $\mathcal{P} = |x|^{-1}$ and special cartoon-like images $\mathcal{T}$, which are characteristic functions of polyhedra, that is, compact regions in $\mathbb{R}^3$ with polygonal boundaries. The reason for choosing the exponent -1 in $\mathcal{P}$ is that we want to match the energies of $\mathcal{P}$ and $\mathcal{T}$ at each scale 2^{-2j}, $j \in \mathbb{Z}$. In fact, from the observation that $\widehat{\mathcal{P}}(\xi) \simeq |\xi|^{-2}$ (cf. [24, Ch.4]) it is easy to verify that $\int_{2^{2j}}^{2^{2j+2}} |\widehat{\mathcal{P}}(\xi)|^2 d\xi \simeq 2^{-2j}$. It will be shown below in Lemma 3.3 that the image $\mathcal{T}$ satisfies the same type of estimate, that is, $\int_{2^{2j}}^{2^{2j+2}} |\widehat{\mathcal{T}}(\xi)|^2 d\xi \simeq 2^{-2j}$.

We will use as analyzing system a combined dictionary of wavelets and shearlets. For the wavelet system, we will use a *Parseval frame of Lemariè-Meyer wavelets* $\Phi = \{\phi_\lambda : \lambda \in \Lambda\} \subset L^2(\mathbb{R}^3)$, for $\Lambda = \{\lambda = (j, k) \in \mathbb{Z} \times \mathbb{Z}^3\}$, where the functions $\phi_\lambda = \phi_{j,k}$ are defined in the Fourier domain by

$$\widehat{\phi}_{j,k}(\xi) = 2^{-3j} W(2^{-2j}\xi) e^{2\pi i 2^{-2j}\xi k},$$

$W \in C_0^\infty(\mathbb{R}^3)$ has support $\mathrm{supp}\,(W) \subset [-\frac{1}{2}, \frac{1}{2}]^3 \setminus [-\frac{1}{16}, \frac{1}{16}]^3$ and it satisfies $\sum_{j \in \mathbb{Z}} |W(2^{-2j}\xi)|^2 = 1$, for a.e. $\xi \in \widehat{\mathbb{R}}^3$ (cf. [14] for the construction of Lemariè-Meyer wavelets).

We recall below the construction of the 3D Parseval frame of shearlets (cf. [11]).

2.1 3D shearlets

A three-dimensional shearlet system is obtained by appropriately combining 3 systems of functions associated with the pyramidal regions

$$\mathcal{P}_1 = \left\{ (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\frac{\xi_2}{\xi_1}| \leq 1, |\frac{\xi_3}{\xi_1}| \leq 1 \right\},$$

$$\mathcal{P}_2 = \left\{ (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\frac{\xi_1}{\xi_2}| < 1, |\frac{\xi_3}{\xi_2}| \leq 1 \right\},$$

$$\mathcal{P}_3 = \left\{ (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3 : |\frac{\xi_1}{\xi_3}| < 1, |\frac{\xi_2}{\xi_3}| < 1 \right\},$$

in which the Fourier space $\widehat{\mathbb{R}}^3$ is partitioned. To define such systems, let ϕ be a C^∞ univariate function such that $0 \leq \hat{\phi} \leq 1$, $\hat{\phi} = 1$ on $[-\frac{1}{16}, \frac{1}{16}]$ and $\hat{\phi} = 0$ outside the interval $[-\frac{1}{8}, \frac{1}{8}]$. That is, ϕ is the scaling function of a Meyer wavelet, rescaled so that its frequency support is contained the interval $[-\frac{1}{8}, \frac{1}{8}]$. For $\xi = (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3$, we define

$$\widehat{\Phi}(\xi) = \widehat{\Phi}(\xi_1, \xi_2, \xi_3) = \hat{\phi}(\xi_1)\,\hat{\phi}(\xi_2)\,\hat{\phi}(\xi_3) \tag{2.1}$$

and

$$W(\xi) = \sqrt{\widehat{\Phi}^2(2^{-2}\xi) - \widehat{\Phi}^2(\xi)}.$$

It follows that

$$\widehat{\Phi}^2(\xi) + \sum_{j \geq 0} W^2(2^{-2j}\xi) = 1 \ \text{ for } \xi \in \widehat{\mathbb{R}}^3. \tag{2.2}$$

Notice that each function $W_j^2 = W^2(2^{-2j} \cdot)$ has support into the Cartesian corona

$$[-2^{-2j-1}, 2^{-2j-1}]^3 \setminus [-2^{-2j-4}, 2^{-2j-4}]^3 \subset \widehat{\mathbb{R}}^3,$$

3

and the functions W_j^2, $j \geq 0$, produce a smooth tiling of the frequency plane into Cartesian coronae, where

$$\sum_{j \geq 0} W^2(2^{-2j}\xi) = 1 \quad \text{for } \xi \in \widehat{\mathbb{R}}^3 \setminus [-\frac{1}{8}, \frac{1}{8}]^3.$$

Next, let $v \in C^\infty(\mathbb{R})$ be such that $\operatorname{supp} v \subset [-1, 1]$ and

$$|v(u-1)|^2 + |v(u)|^2 + |v(u+1)|^2 = 1 \quad \text{for } |u| \leq 1. \tag{2.3}$$

In addition, we can assume that $v(0) = 1$ and that $v^{(n)}(0) = 0$ for all $n \geq 1$.

Hence, for $d = 1, 2, 3$, $\ell = (\ell_1, \ell_2) \in \mathbb{Z}^2$, the 3D *shearlet systems associated with the pyramidal regions* $\mathcal{P}_d$ are defined as the collections

$$\{\psi_{j,\ell,k}^{(d)} : j \geq 0, -2^j \leq \ell_1, \ell_2 \leq 2^j, k \in \mathbb{Z}^3\}, \tag{2.4}$$

where

$$\hat{\psi}_{j,\ell,k}^{(d)}(\xi) = |\det A_{(d)}|^{-j/2} \, W(2^{-2j}\xi) \, V_{(d)}(\xi A_{(d)}^{-j} B_{(d)}^{[-\ell]}) \, e^{2\pi i \xi A_{(d)}^{-j} B_{(d)}^{[-\ell]} k}, \tag{2.5}$$

$V_{(1)}(\xi_1, \xi_2, \xi_3) = v(\frac{\xi_2}{\xi_1})v(\frac{\xi_3}{\xi_1})$, $V_{(2)}(\xi_1, \xi_2, \xi_3) = v(\frac{\xi_1}{\xi_2})v(\frac{\xi_3}{\xi_2})$, and $V_{(3)}(\xi_1, \xi_2, \xi_3) = v(\frac{\xi_1}{\xi_3})v(\frac{\xi_2}{\xi_3})$; the anisotropic dilation matrices $A_{(d)}$ are given by

$$A_{(1)} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \, A_{(2)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \, A_{(3)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

and the *shear matrices* are defined by

$$B_{(1)}^{[\ell]} = \begin{pmatrix} 1 & \ell_1 & \ell_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \, B_{(2)}^{[\ell]} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_1 & 1 & \ell_2 \\ 0 & 0 & 1 \end{pmatrix}, \, B_{(3)}^{[\ell]} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \ell_1 & \ell_2 & 1 \end{pmatrix}.$$

Notice that $(B_{(d)}^{[\ell]})^{-1} = B_{(d)}^{[-\ell]}$.

Due to the support conditions on W and v, the elements of the system of shearlets (2.4) have compact support in Fourier domain. In particular, for $d = 1$, the shearlets $\hat{\psi}_{j,\ell,k}^{(1)}(\xi)$ can be written explicitly as

$$\hat{\psi}_{j,\ell_1,\ell_2,k}^{(1)}(\xi) = 2^{-2j} \, W(2^{-2j}\xi) \, v\left(2^j \frac{\xi_2}{\xi_1} - \ell_1\right) v\left(2^j \frac{\xi_3}{\xi_1} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell_1, -\ell_2]} k}, \tag{2.6}$$

showing that their supports are contained inside the trapezoidal regions

$$\begin{aligned} U_{j,\ell} = U_{j,\ell_1,\ell_2} &= \{(\xi_1, \xi_2, \xi_3) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \\ &\quad |\frac{\xi_2}{\xi_1} - \ell_1 2^{-j}| \leq 2^{-j}, |\frac{\xi_3}{\xi_1} - \ell_2 2^{-j}| \leq 2^{-j}\}. \end{aligned} \tag{2.7}$$

A *Parseval frame of shearlets* is obtained by combining the shearlet systems associated with the cone-shaped regions $\mathcal{P}_\nu$ together with a coarse scale system and appropriate *boundary* shearlets. The boundary shearlets are slightly modified versions of the functions $\psi_{j,\ell_1,\ell_2,k}^{(d)}$, for $\ell_1, \ell_2 = \pm 2^j$, where the modification is needed to ensure that all elements of system are C_0^∞ in the Fourier domain (cf. [11] for details).

For brevity, in the following we will denote the Parseval frame of 3D shearlets as $\Psi = \{\psi_\eta : \eta \in M\} \subset L^2(\mathbb{R}^3)$, where the index set is $M = M_C \cup M_F$, $M_C = \{k \in \mathbb{Z}^3\}$ is the set of indices associated with coarse-scale shearlets and $M_F = \{\eta = (j, \ell, k, d) : j \geq 0, |\ell| \leq 2^j, k \in \mathbb{Z}^2, d = 1, 2, 3\}$ is the set of indices associated with fine-scale shearlets.

4

2.2 Main theorem

Both the wavelet and shearlet systems produce a multiscale decomposition associated with the family of band-pass filters $\widehat{F_j}(\xi) = W(2^{-2j}\xi)$, for $j \in \mathbb{Z}$. By applying these filters, for $j \in \mathbb{Z}$, we define

$$\mathcal{P}_j = \mathcal{P} * F_j, \quad \mathcal{T}_j = \mathcal{T} * F_j, \quad f_j = f * F_j, \tag{2.8}$$

where, as observed above, we have that $\|\mathcal{P}_j\|_2 \simeq 2^{-j}$, $\|\mathcal{T}_j\|_2 \simeq 2^{-j}$. Note that, the band-passed functions $\hat{f}_j$, $j \in \mathbb{Z}$, are supported in the Cartesian coronae $[-2^{2j-1}, 2^{2j-1}]^3 \setminus [-2^{2j-4}, 2^{2j-4}]^3$ in the Fourier domain $\widehat{\mathbb{R}}^3$. Furthermore, for $f \in L^2(\mathbb{R}^3)$, it follows from the properties of the function W that

$$f = \sum_j F_j * f_j, \tag{2.9}$$

with convergence in the L^2-norm.

Let $\mathcal{F}_j$ denote the range of the operator of convolution with F_j. It is a simple calculation to verify that the shearlets and wavelets at level j' are orthogonal to $\mathcal{F}_j$ unless $|j' - j| \leq 1$, that is, unless $j' = j-1, j, j+1$. It is useful to introduce the notation

$$\Lambda_j = \{\lambda = (j', k) : |j' - j| \leq 1, \, k \in \mathbb{Z}^3\} \tag{2.10}$$

and

$$M_j = \{\eta = (j', \ell, k, d) : |j' - j| \leq 1, |\ell_1| \leq 2^j, |\ell_1| \leq 2^j, \, k \in \mathbb{Z}^3, d = 1, 2, 3\}. \tag{2.11}$$

Hence, using this notation, we observe that the elements of $\mathcal{F}_j$ can be expanded using only the elements of the wavelet system in Λ_j or only the elements of the shearlet system in M_j. In other words, at the level j, we can use the wavelet system to represent f_j as

$$f_j = \sum_{j'=j-1}^{j'=j+1} \sum_{k' \in \mathbb{Z}^2} \langle f_j, \phi_{j',k'} \rangle \, \phi_{j',k'} = \sum_{\lambda \in \Lambda_j} \langle f_j, \phi_\lambda \rangle \, \phi_\lambda;$$

or we can use the shearlet system to represent f_j as

$$f_j = \sum_{d=1}^{3} \sum_{j'=j-1}^{j'=j+1} \sum_{|\ell_1| \leq 2^{j'}} \sum_{|\ell_2| \leq 2^{j'}} \sum_{k \in \mathbb{Z}^2} \langle f_j, \psi^{(d)}_{j',\ell_1,\ell_2,k} \rangle \, \psi^{(d)}_{j',\ell_1,\ell_2,k} = \sum_{\eta \in M_j} \langle f_j, \psi_\eta \rangle \, \psi_\eta.$$

Clearly, we can also consider a combined representation of the form

$$f_j = \sum_{\lambda \in \Lambda_j} w_\lambda \, \phi_\lambda + \sum_{\eta \in M_j} t_\eta \, \psi_\eta,$$

for an appropriate choice of coefficients $w = (w_\lambda)$ and $t = (t_\eta)$. Since, in this last expression, we are considering an overcomplete dictionary, there are many possible choices of coefficients w and t, some of which may provide sparser representations than either one of the two expansions above. Similar to [4], we seek a solution providing a *geometric separation*, that is, we consider the following dual-frame *component separation problem* based on ℓ_1 minimization:

$$(W_j, T_j) = \arg\min(\|w\|_1 + \|t\|_1), \quad \text{subject to } f_j = W_j + T_j, \tag{2.12}$$

where $w_\lambda = \langle W_j, \phi_\lambda \rangle$, $\lambda \in \Lambda_j$ and $t_\eta = \langle T_j, \psi_\eta \rangle$, $\eta \in M_j$. It follows from (2.9) that, if we let $\tilde{P} = \sum_j F_j * W_j$, $\tilde{T} = \sum_j F_j * T_j$, then we can express f as the superposition $f = \tilde{P} + \tilde{T}$.

The main result of our paper is the following theorem, which states that it is possible to achieve the *separation* of the distinct geometric objects $\mathcal{P}$ and $\mathcal{T}$, asymptotically at fine scales, through ℓ_1 minimization over a combined wavelet-shearlet dictionary.

Theorem 2.1. *Let Φ and Ψ be the Parseval frames of wavelets and shearlets, respectively, define above and denote $\|g\|_{1,\Phi} = \sum_{\lambda \in \Lambda_j} |\langle g, \phi_\lambda \rangle|$ and $\|g\|_{1,\Psi} = \sum_{\eta \in M_j} |\langle g, \psi_\eta \rangle|$, where Λ_j is given by (2.10) and M_j is given by (2.11). Let $\mathcal{P}_j$, $\mathcal{T}_j$ be given by (2.8). We have that*

$$\lim_{j \to \infty} \frac{\|W_j - \mathcal{P}_j\|_{1,\Phi} + \|T_j - \mathcal{T}_j\|_{1,\Psi}}{\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi}} = 0.$$

That is, asymptotically as the scale tends to zero, the pointlike component of f is captured by the Parseval frame of wavelets and the piecewise linear component of f is captured by the Parseval frame of shearlets.

Remark: The statement of Theorem 2.1 is different from the corresponding Theorem 1.1 in [4], that is valid in the two-dimensional case, since it is formulated using the ℓ^1–norm rather than the ℓ^2–norm used in [4]. This difference is crucial to be able to handle singularities containing edges and vertices. The main novel contributions of our paper are contained in the proofs of Lemma 3.5 and Lemma 3.6, which will be presented in the next section. The proof of Lemma 3.5 builds upon on a prior result (Lemma 3.4) and a self-improving process, which is valuable in itself. Lemma 3.6 gives the decay estimate for edges and corner points and is significantly more involved than the one used to analyze the smooth segment case in [4] (also compare our Lemma 3.7 in Section 3.2).

The rest of the paper is devoted mostly to the proof of Theorem 2.1.

3 Proof of main theorem

As we will show below, our proof follows the general structure from [4], which is centered around the notion of cluster coherence. However, the main and most difficult parts of the proof concerning the analysis of the piecewise linear singularities is completely new and does follows from the original paper.

Let $\Phi = \{\phi_\lambda : \lambda \in \Lambda\}$ and $\Psi = \{\psi_\mu : \mu \in M\}$ be the Parseval frames of 3D wavelets and 3D shearlets that we introduced above, respectively. For each level $j \in \mathbb{Z}$, we denote the indices for the cluster of significant wavelet coefficients as $S_{1,j} \subset \Lambda_j$ and the indices for the cluster of significant shearlet coefficients as $S_{2,j} \subset M_j$. We discuss below how to determine such sets.

Corresponding to the sets $S_{1,j}$ and $S_{2,j}$, we define the *wavelet approximation error* and the *shearlet approximation error* at the level j as

$$\delta_{1,j} = \sum_{\lambda \in S_{1,j}^c} |\langle P_j, \phi_\lambda \rangle|, \quad \delta_{2,j} = \sum_{\eta \in S_{2,j}^c} |\langle T_j, \psi_\eta \rangle|$$

respectively, and the *cluster coherences* as

$$\mu_c(S_{1,j}, \Phi; \Psi) = \max_\eta \sum_{\lambda \in S_{1,j}} |\langle \phi_\lambda, \psi_\eta \rangle|, \quad \mu_c(S_{2,j}, \Psi; \Phi) = \max_\lambda \sum_{\eta \in S_{2,j}} |\langle \phi_\lambda, \psi_\eta \rangle|.$$

The notion of cluster coherence was originally proposed in [4]. Unlike the more standard definition of *coherence*, given by $\mu(\Phi, \Psi) = \max_{\lambda,\eta} |\langle \phi_\lambda, \psi_\eta \rangle|$, the cluster coherence bounds coherence between a single member of a frame and a *cluster* of members of another frame. As we will see below, the cluster coherence is the 'right' notion for the purpose of geometric separation.

Let Φ be the matrix representation of the Parseval frame of wavelets and Ψ the matrix representation of the Parseval frame of shearlets. For a $g_j \in L^2(\mathbb{R}^3)$ such that $\operatorname{supp}(\hat{g}_j) \subset \mathcal{F}_j$, let

$$\|1_{S_{1,j}} \Phi^T g_j\|_1 = \sum_{\lambda \in S_{1,j}} |\langle g_j, \phi_\lambda \rangle|, \quad \|1_{S_{2,j}} \Psi^T g_j\|_1 = \sum_{\eta \in S_{2,j}} |\langle g_j, \psi_\eta \rangle|.$$

Next, we define the *joint concentration* by

$$k = k(S_{1,j}, S_{2,j}) = \sup_{g_j} \frac{\|1_{S_{1,j}} \Phi^T g_j\|_1 + \|1_{S_{2,j}} \Psi^T g_j\|_1}{\|\Phi^T g_j\|_{1,\Phi} + \|\Psi^T g_j\|_{1,\Psi}}.$$

The following useful observations illustrates the relationship between joint concentration and data separation. This result is similar to Proposition 2.1 from [4] and can be proved using the same argument.

Proposition 3.1. *Suppose that, for $j \in \mathbb{Z}$, $f_j = W_j + T_j$ so that each component of f_j is relatively sparse in Φ or Ψ, that is,*

$$\|1_{S_{1,j}}\Phi^T W_j\|_1 \leq \delta_{1,j}, \quad \|1_{S_{2,j}}\Psi^T T_j\|_1 \leq \delta_{2,j}.$$

If (W_j, T_j) solves (2.12), then

$$\|W_j - \mathcal{P}_j\|_{1,\Phi} + \|T_j - \mathcal{T}_j\|_{1,\Psi} \leq \frac{2(\delta_{1,j} + \delta_{2,j})}{1 - 2k}.$$

A related observation from [4] is that the joint concentration is bounded above by the maximum of the cluster coherences. We have (cf. [4, Lemma 2.1]):

Lemma 3.2.

$$k(S_{1,j}, S_{2,j}) \leq \max\{\mu_c(S_{1,j}, \Phi; \Psi), \ \mu_c(S_{2,j}, \Psi; \Phi)\}$$

It follows from Proposition 3.1 and Lemma 3.2 that Theorem 2.1 is proved if we can construct appropriate set of coefficients $S_{1,j}$ and $S_{2,j}$ such that $\delta_{1,j} = o(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi})$, $\delta_{2,j} = o(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi})$ and

$$\mu_c(S_{1,j}, \Phi; \Psi) \to 0, \quad \mu_c(S_{2,j}, \Psi; \Phi) \to 0, \quad \text{as } j \to \infty.$$

The rest of the proof is organized as follows. In Section 3.1, we will select an appropriate set $S_{1,j}$ and show that $\mu_c(S_{1,j}, \Phi; \Psi) \to 0$ and $\delta_{1,j} = o(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi})$. This is the easy part of the argument and it follows from an idea similar to [4]. For the 'hard' part of the proof, concerning the analysis of the piecewise linear singularities, the original argument from [4] cannot be extended directly and we will use a novel approach. Namely, in Section 3.2, we will choose an appropriate set $S_{2,j}$ and show that $\mu_c(S_{2,j}, \Psi; \Phi) \to 0$ and $\delta_{2,j} = o(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi})$. The most delicate part of our estimate is contained in the proof of Lemma 3.6 in Section 3.2, where it is highly nontrivial to control the size of the translation variable k corresponding to 3D edges and vertices. Our novel argument is inspired by the techniques developed for the analysis of singularities using the shearlet transform, which was developed to handle piecewise linear discontinuities in 3D (cf. [8, 10]).

In the following, for all our arguments, it will be sufficient to consider the shearlet system associated with the cone-shaped regions $\mathcal{P}_1 \subset \widehat{\mathbb{R}}^3$ only, since the properties of the similar system in $\mathcal{P}_2$ and $\mathcal{P}_3$ are the same. In the Fourier domain, the elements (2.6) of such shearlet system can be written[2] as

$$\hat{\psi}_{j,\ell_1,\ell_2,k}^{(1)}(\xi) = 2^{-2j}\Gamma_{j,\ell_1,\ell_2}(\xi)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell_1,-\ell_2]}k},$$

where

$$\Gamma_{j,\ell_1,\ell_2}(\xi) = W(2^{-2j}\xi)\, v\Big(2^j\frac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(2^j\frac{\xi_3}{\xi_1} - \ell_2\Big).$$

Note that $A_{(1)}^{-j}B_{(1)}^{[-\ell_1,-\ell_2]}k = (2^{-2j}(k_1 - \ell_1 k_2 - \ell_2 k_3), 2^{-j}k_2 + 2^{-j}k_3)$. Each function Γ_{j,ℓ_1,ℓ_2} is supported inside the set U_{j,ℓ_1,ℓ_2}, given by (2.7). It is easy to verify that its measure satisfies $|U_{j,\ell_1,\ell_2}| \leq C\, 2^{4j}$.

3.1 Estimate for the point singularities

In this section, we will select the set $S_{1,j}$ and prove that $\mu_c(S_{1,j}, \Phi; \Psi) \to 0$ and $\delta_{1,j} = o(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi})$, asymptotically as $j \to \infty$. As indicated above, the rather simple argument that we use is similar to [4].

---

[2]Here we ignore the fact that the boundary elements corresponding to $\ell = \pm 2^j$ need to be slightly modified. However this is irrelevant to the arguments in the rest of the paper.

Let $\phi_{j',k'}$ and $\psi_{j,\ell_1,\ell_2,k}$ be generic elements from the Parseval frames of wavelets and shearlets, respectively. Due to the frequency support of W, for any ℓ_1, ℓ_2, k and k' we have that $\langle \widehat{\psi_{j,\ell,k}}, \widehat{\phi_{j',k'}} \rangle = 0$ if $|j - j'| > 1$. Thus for all large j' and $j = j' - 1, j', j' + 1$, a direct computation shows that

$$
\begin{aligned}
|\langle \widehat{\psi_{j,\ell_1,\ell_2,k}}, \widehat{\phi_{j',k'}} \rangle| &= \left| \int_{\mathbb{R}^2} \left(2^{-2j} \Gamma_{j,\ell_1,\ell_2}(\xi) \, e^{-2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell_1,-\ell_2]} k} \right) \left(2^{-3j'} W(2^{-2j'}\xi) e^{2\pi i 2^{-3j'}\xi \cdot k'} \right) d\xi \right| \\
&\leq 2^{-2j} 2^{-2j'} \int_{\mathbb{R}^3} |\Gamma_{j,\ell_1,\ell_2}(\xi) \, W(2^{-2j'}\xi)| \, d\xi \\
&\leq C \, 2^{-2j} 2^{-3j'} \int_{\Omega_{j,\ell_1,\ell_2}} d\xi \leq C \, 2^{-2j} 2^{-3j'} 2^{4j} \leq C \, 2^{-j},
\end{aligned}
$$

where C is independent of ℓ_1, ℓ_2, k, k' and j.

For a fixed $0 < \epsilon < 1$, we set $S_{1,j} = \{(j', k') : j' = j - 1, j, j + 1; \ |k'| \leq 2^{\epsilon j'}\}$. Then, using the calculation above, we have that

$$
\mu_c(S_{1,j}, \Phi; \Psi) \leq C \max_{\ell_1,\ell_2,k} \sum_{j'=j-1}^{j+1} \sum_{|k'| \leq 2^{\epsilon j'}} |\langle \widehat{\psi_{j,\ell_1,\ell_2,k}}, \widehat{\phi_{j',k'}} \rangle| \leq C 2^{(-1+\epsilon)j}.
$$

It follows that $\mu_c(S_{1,j}, \Phi; \Psi) \to 0$, as $j \to \infty$.

We also observe that $\langle \widehat{\phi_{j',k'}}, \widehat{\mathcal{P}_j} \rangle = 0$ for all k' if $|j' - j| > 1$. For $|j' - j| \leq 1$, we have that

$$
\begin{aligned}
\langle \widehat{\phi_{j',k'}}, \widehat{\mathcal{P}_j} \rangle &= 2^{-3j'} C \int_{\mathbb{R}^3} W(2^{-2j'}\xi) \, e^{2\pi i 2^{-2j'}\xi \cdot k'} W(2^{-2j}\xi) |\xi|^{-2} \, d\xi \\
&= C 2^{-j'} \int_{\mathbb{R}^3} W(\xi) W(2^{2(j'-j)}\xi) \, e^{2\pi i \xi \cdot k'} |\xi|^{-2} \, d\xi.
\end{aligned}
$$

Hence, for $|k'| \geq 2^{\epsilon j}$, integration by parts gives that

$$
|\langle \widehat{\phi_{j',k'}}, \widehat{\mathcal{P}_j} \rangle| \leq C_N \, 2^{-j} (1 + |k'|)^{-N} \leq C_N \, 2^{-(1+N\epsilon)j}.
$$

It follows that, by choosing N sufficiently large, we have:

$$
\delta_{1,j} = \sum_{\lambda \in S_{1,j}^c} |\langle \phi_\lambda, P_j \rangle| \leq C \, 2^{-2j} = o(2^{-j}) = o\left(\|\mathcal{P}_j\|_{1,\Phi} + \|\mathcal{T}_j\|_{1,\Psi} \right).
$$

3.2 Estimate for the piecewise linear singularities

We first recall the Divergence Theorem in $\mathbb{R}^3$. Let $\vec{F}$ be a smooth vector field in $\mathbb{R}^2$ and S be a compact region in the plane with a piecewise smooth simple boundary ∂S. Then

$$
\int_S \nabla \cdot \vec{F} \, dA = \int_{\partial S} \vec{F} \cdot \vec{n} \, ds,
$$

where $\vec{n}(x)$ is the outer normal direction at $x \in \partial S$.

Let $\mathcal{T}$ be the characteristic function of a polyhedron of M faces in $\mathbb{R}^3$. Without loss of generality, we may assume that the polyhedron is contained inside the cube $[-1,1]^3$ (if not, the polyhedron can be rescaled by dilation on the space variables). Let $S = \cup_{m=1}^M S_m$ be the boundary of the polyhedron, where for each $1 \leq m \leq M$, S_m is a polygon in $\mathbb{R}^2$. For $\xi \in \widehat{\mathbb{R}}^3$, Using the divergence theorem we can express the Fourier

transform of $\mathcal{T}$ as follows

$$
\begin{aligned}
\widehat{\mathcal{T}}(\xi) &= -\frac{1}{2\pi i |\xi|} \int_S e^{-2\pi i \xi \cdot x} \frac{\xi}{|\xi|} \cdot \vec{n}(x)\, d\sigma(x) \\
&= \sum_1^M \left(-\frac{1}{2\pi i |\xi|}\right) \int_{S_m} e^{-2\pi i \xi \cdot x} \frac{\xi}{|\xi|} \cdot \vec{n}(x)\, d\sigma(x) \\
&= \sum_1^M \widehat{\mathcal{T}}^{(m)}(\xi),
\end{aligned}
$$

where $\widehat{\mathcal{T}}^{(m)}(\xi) = -\frac{1}{2\pi i |\xi|} \int_{S_m} e^{-2\pi i \xi \cdot x} \frac{\xi}{|\xi|} \cdot \vec{n}(x)\, d\sigma(x)$ for each $1 \le m \le M$.

In order to estimate $\widehat{\mathcal{T}}^{(m)}(\xi)$, we divide S_m into finite many sub-surfaces so that each sub-surface is the graph of a linear function on a triangle domain $D_m \in \mathbb{R}^2$. Without loss of generality, we may assume that each surface S_m can be written as $S_m = \{(Ax_2 + Bx_3, x_2, x_3) : (x_2, x_3) \in D_m\}$, where $D_m = \{(x_2, x_3) : 0 \le x_3 \le bx_2,\ a_1 \le x_2 \le a_2\}$ and A, B, b, a_1, a_2 are appropriate constants. Let $\eta = (\eta_1, \eta_2) = (A\xi_1 + \xi_2, B\xi_1 + \xi_3)$ so that, for $x \in S_m$, we have $\xi \cdot x = \xi_1(Ax_2 + Bx_2) + \xi_2 x_2 + \xi_3 x_3 = (A\xi_1 + \xi_2)x_2 + (B\xi_1 + \xi_3)x_3 = \eta \cdot (x_2, x_3)$.

A direction calculation shows that

$$
\begin{aligned}
\int_{S_m} e^{-2\pi i \xi \cdot x} \frac{\xi}{|\xi|} \cdot \vec{n}(x)\, d\sigma(x) &= \int_{D_m} e^{-2\pi i \, \eta \cdot (x_2, x_3)} \frac{\xi}{|\xi|} \cdot (-1, A, B)\, dx_3 dx_2 \\
&\simeq \frac{1}{1 + |B\xi_1 + \xi_3|} \frac{1}{1 + |A\xi_1 + \xi_2 + b(B\xi_1 + \xi_3)|}.
\end{aligned}
$$

It follows that

$$
\widehat{\mathcal{T}}^{(m)}(\xi) \simeq \frac{1}{|\xi|} \frac{1}{1 + |B\xi_1 + \xi_3|} \frac{1}{1 + |A\xi_1 + \xi_2 + b(B\xi_1 + \xi_3)|}. \tag{3.13}
$$

Let K be the characteristic function of the unit ball in $\mathbb{R}^3$. It is known that $\widehat{K}(\xi) = |\xi|^{-\frac{3}{2}} J_{\frac{3}{2}}(|\xi|)$, where $J_{\frac{3}{2}}$ is the Bessel function of order $3/2$. Using the observation that the asymptotic decay of $J_{\frac{3}{2}}(|\xi|)$, for large $|\xi|$, is of the order $|\xi|^{-\frac{1}{2}}$ (cf. [22, Ch.8]), it follows that $\|K_j\|_2 = \|K * F_j\|_2 \simeq 2^{-j}$. The following lemma shows that the same estimate holds for $\mathcal{T}_j$.

Lemma 3.3. *For $j \in \mathbb{N}$ and $\mathcal{T}_j$ defined above, the following estimate holds:*

$$
\|\mathcal{T}_j\|_2 \simeq 2^{-j}.
$$

Proof. Using the notation introduced above, let $\widehat{\mathcal{T}}_j^{(m)}(\xi) = \widehat{\mathcal{T}}^{(m)}(\xi)\, W(2^{-2j}\xi)$. Since $\widehat{\mathcal{T}}_j = \sum_1^M \widehat{\mathcal{T}}_j^{(m)}$ and $2^{2j} \le |\xi| \le 2^{2j+2}$ inside the support of $W(2^{-2j}\xi)$, an easy calculation using spherical coordinates gives that

$$
\begin{aligned}
&\|\mathcal{T}_j^{(m)}\|_2^2 \\
&\simeq \int_{2^{2j} \le |\xi| \le 2^{2j+2}} \frac{1}{|\xi|^2} \frac{1}{1 + |B\xi_1 + \xi_3|^2} \frac{1}{1 + |A\xi_1 + \xi_2 + b(B\xi_1 + \xi_3)|^2}\, d\xi \\
&\simeq 2^{-2j}.
\end{aligned}
$$

Hence, $\|\mathcal{T}_j\|_2 = \|\sum_1^M \mathcal{T}_j^{(m)}\|_2 \simeq 2^{-j}$. $\qquad\square$

The following lemma is a special case of Proposition 4.7 in [13].

Lemma 3.4. *Let $\beta = (\beta_i)$ be a sequence of non-negative numbers and let $|\beta|_{(N)}$ be the Nth-largest element in the decreasing rearrangement of β. Then*

$$
\left(\sum_{\beta_i \le |\beta|_{(N)}} \beta_i^2 \right)^{\frac{1}{2}} \le C\, N^{-\frac{1}{2}} \|\beta\|_1,
$$

9

where C is independent of N.

Using the Lemma 3.4, we derive the following observation.

Lemma 3.5. *For a fixed large j, let $\beta_j = \{\beta_j(j', \ell_1, \ell_2, k) = \langle \mathcal{T}_j, \psi_{j', \ell_1, \ell_2, k} \rangle : |\ell_1| \leq 2^{j'}, |\ell_1| \leq 2^{j'}, j - 1 \leq j' \leq j + 1, k \in \mathbb{Z}^3\}$ and, for $v = 1, 2$, define the norms*

$$\|\beta_j\|_v = \left(\sum_{j'=j-1}^{j+1} \sum_{|\ell_1| \leq 2^{j'}} \sum_{|\ell_2| \leq 2^{j'}} \sum_{k \in \mathbb{Z}^3} |\langle \mathcal{T}_j, \psi_{j', \ell_1, \ell_2, k} \rangle|^v \right)^{\frac{1}{v}}.$$

Then there is a constant $C > 0$ such that $\|\beta_j\|_1 \geq C \, 2^{-\frac{1}{4}j}$.

Proof. Let

$$\widehat{\alpha_{j', j, \ell}}(\xi A_{(1)}^{-j'} B_{(1)}^{[-\ell]}) = 2^{-2j'} W(2^{-2j'}\xi) \, \Gamma_{j', \ell}(\xi) \, e^{2\pi i \xi A_{(1)}^{-j'} B_{(1)}^{[-\ell]} k} W(2^{-2j}\xi)$$

so that

$$\beta_j(j', \ell_1, \ell_2, k) = 2^{2j'} \int_{[-1,1]^3} \mathcal{T}(x) \, \alpha_{j', j, \ell_1, \ell_2}(B_{(1)}^{[\ell]} A_{(1)}^{j'} x - k) \, dx.$$

Using the change of variable $y = B_{(1)}^{[\ell]} A_{(1)}^{j'} x$, it is easy to verify that there is a constact $C > 0$ such that $|\beta_j(j', \ell, k)| \leq C \, 2^{-2j'}$ since $\int_{\mathbb{R}^3} |\alpha_{j', j, \ell_1, \ell_2}(y - k)| dy \leq C$ uniformly for all j', j, ℓ_1, ℓ_2.

For a fixed $j \geq 0$, as in Lemma 3.4, we denote by $|\beta|_{(n)}$ the nth-largest element for $|\beta_j(j', \ell_1, \ell_2, k)|$. If we let $N = 2^j$, then we have the following estimate for the sum of the squares of the first 2^j largest terms:

$$\left(\sum_{n=1}^{2^j} |\beta|_{(n)}^2 \right)^{\frac{1}{2}} \leq C \, 2^{\frac{1}{2}j} 2^{-2j'} = C \, 2^{-\frac{1}{2}j} 2^{-j} \leq C \, 2^{-\frac{1}{2}j} \|\beta_j\|_2,$$

where we used the fact that $\|\beta_j\|_2 \simeq 2^{-j}$ since the shearlet system is a tight frame.

The last expression is controlled by $2^{-\frac{1}{2}j} \|\beta_j\|_1$ since $\|\beta_j\|_2 \leq \|\beta_j\|_1$. Thus, combining this fact with the estimate of Lemma 3.4 for $N = 2^j$, we have that

$$\|\beta_j\|_1 \geq 2^{\frac{1}{2}j} \|\beta_j\|_2 \geq 2^{-\frac{1}{2}j}.$$

Now we apply Lemma 3.4 again with $N = 2^{\frac{3}{2}j}$ to get that there is constant $C > 0$ such that

$$\left(\sum_{\beta_i \leq |\beta|_{(2^{3j/2})}} \beta_i^2 \right)^{\frac{1}{2}} \leq C \, 2^{-\frac{3}{4}j} \, \|\beta_j\|_1.$$

Again we observe that

$$\left(\sum_{n=1}^{2^{\frac{3}{2}j}} |\beta|_{(n)}^2 \right)^{\frac{1}{2}} \leq C \, 2^{\frac{3}{4}j} 2^{-2j'} = C \, 2^{-\frac{5}{4}j} = C \, 2^{-\frac{3}{4}j} 2^{-\frac{1}{2}j} \leq C \, 2^{-\frac{3}{4}j} \|\beta_j\|_1.$$

Combining the above two estimates, we conclude that there is constant $C > 0$ such that $\|\beta_j\|_1 \geq C 2^{\frac{3}{4}j} \|\beta_j\|_2 \geq C 2^{-\frac{1}{4}j}$. $\square$

Let $\beta_j^{(m)}(j', \ell_1, \ell_2, k) = \langle \mathcal{T}_j^{(m)}, \psi_{j', \ell_1, \ell_2, k} \rangle$, where (using the same notation as in the proof of Lemma 3.3) $\widehat{\mathcal{T}}_j^{(m)}(\xi) = \widehat{\mathcal{T}}^{(m)}(\xi) W(2^{-2j}\xi)$ and $\widehat{\mathcal{T}}^{(m)}(\xi)$ is given by (3.13). Since

$$\langle \mathcal{T}_j, \psi_{j', \ell_1, \ell_2, k} \rangle = \langle \widehat{\mathcal{T}}_j, \widehat{\psi_{j', \ell_1, \ell_2, k}} \rangle = \sum_{m=1}^{M} \langle \widehat{\mathcal{T}}_j^{(m)}, \widehat{\psi_{j', \ell_1, \ell_2, k}} \rangle,$$

we can write

$$\beta_j(j', \ell_1, \ell_2, k) = \sum_{m=1}^{M} \langle \mathcal{T}_j^{(m)}, \psi_{j', \ell_1, \ell_2, k} \rangle = \sum_{m=1}^{M} \beta_j^{(m)}(j', \ell_1, \ell_2, k).$$

A direct calculation gives that

$$
\begin{aligned}
\beta_j^{(m)}(j', \ell_1, \ell_2, k) &= \langle \widehat{\mathcal{T}}_j^{(m)}, \widehat{\psi_{j', \ell_1, \ell_2, k}} \rangle \\
&= 2^{-2j'} \int_{U_{j', \ell_1, \ell_2}} \widehat{\mathcal{T}}_j^{(m)}(\xi) \, \Gamma_{j', \ell_1, \ell_2}(\xi) \, e^{2\pi i \xi A_{(1)}^{-j'} B_{(1)}^{[-\ell]} k} \, d\xi.
\end{aligned}
$$

In order to deal with the edges and vertices in the singularity set and control the size of k, we will now decompose the functions $\widehat{\mathcal{T}}_j^{(m)}$ into their 'directional' components.

Let $g(t) = |V(t)|^2$, where V is the function introduced in Section 2 for the construction of the shearlets. Clearly, $g \in C_0^\infty(-1, 1)$ and, by (2.3), it follows that for all $j \geq 0$ it satisfies the equation

$$\sum_{p=-2^j}^{2^j} g(2^j t - p) = 1, \quad \text{for} \quad |t| \leq 1.$$

Hence, we obtain the following expansion of $\widehat{\mathcal{T}}_j^{(m)}$ into its the 'directional' components:

$$\widehat{\mathcal{T}_j^{(m)}}(\xi) = \sum_{p=-2^j}^{2^j} \sum_{q=-2^j}^{2^j} \widehat{\mathcal{T}_j^{(m,p,q)}}(\xi),$$

where

$$\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi) = -\frac{W(2^{-2j}\xi)}{(2\pi)^2 |\xi|} \frac{\xi}{|\xi|} \cdot (-1, A, B) \int_D g(2^j x_2 - p) g(2^j x_3 - q) e^{-2\pi i \eta \cdot (x_2, x_3)} dx_3 dx_2.$$

Let $\beta_j^{(m,p,q)}(j', \ell_1, \ell_2, k) = \langle \widehat{\mathcal{T}}_j^{(m,p,q)}, \widehat{\psi_{j', \ell_1, \ell_2, k}} \rangle$. For simplicity of notations, in the following we will assume $j' = j$. If $j' \neq j$, the only difference is that one needs to add up the terms corresponding to $j' = j-1, j, j+1$ and, as we have seen above, the effect of adding up $j' = j-1, j, j+1$ is irrelevant for the argument since it only yields a different uniform constants in the final estimates. Thus, in the following, we will ignore the sum over j' and only consider the terms

$$\beta_j^{(m,p,q)}(\ell_1, \ell_2, k) := \langle \widehat{\mathcal{T}}_j^{(m,p,q)}, \widehat{\psi_{j, \ell_1, \ell_2, k}} \rangle.$$

Let $\eta = (A\xi_1 + \xi_2, B\xi_1 + \xi_3) = \xi_1(A + \frac{\xi_2}{\xi_1}, B + \frac{\xi_3}{\xi_1})$. Due to the assumptions on the support of $\hat{\psi}$, we see that, on the support U_{j, ℓ_1, ℓ_2} of $\Gamma_{j, \ell_1, \ell_2}$, we have that $|2^j \frac{\xi_2}{\xi_1} - \ell_1| \leq 1$, $|2^j \frac{\xi_3}{\xi_1} - \ell_2| \leq 1$ and $2^{2j} \leq |\xi_1| \leq 2^{2j+2}$. Let $\ell_{1,m} = -A2^j, ; \ell_{2,m} = -B2^j$. It follows that

$$
\begin{aligned}
|\eta|^2 &= |\xi_1|^2 \left((A + \frac{\xi_2}{\xi_1})^2 + (B + \frac{\xi_3}{\xi_1})^2 \right) \\
&\simeq 2^j \left((2^j \frac{\xi_2}{\xi_1} + A2^j)^2 + (2^j \frac{\xi_3}{\xi_1} + B2^j)^2 \right) \\
&\simeq 2^j \left((\ell_1 - \ell_{1,m})^2 + (\ell_2 - \ell_{2,m})^2 \right).
\end{aligned}
$$

For each pair of $(p, q) \in \mathbb{Z} \times \mathbb{Z}$, it is easy to see that the support of the function $g(2^j x_2 - p) g(2^j x_3 - q)$ is contained inside the set $I_{p,q} = (2^{-j} p - 2^{-j}, 2^{-j} p + 2^{-j}) \times (2^{-j} q - 2^{-j}, 2^{-j} q + 2^{-j})$ and that $\bigcup_{p,q=-2^j}^{2^j} I_{p,q}$ is an open cover of $[-1, 1]^2$. We wil consider the following two types of integer pairs (p, q):

$$J_{p,q}^{(1)} = \{(p, q) \in \mathbb{Z} \times \mathbb{Z} : I_{p,q} \bigcap \partial D \neq \emptyset\}, \quad J_{p,q}^{(2)} = \{(p, q) \in \mathbb{Z} \times \mathbb{Z} : I_{p,q} \bigcap \partial D = \emptyset\}.$$

11

We observe that there are at most $C2^j$ pairs of (p,q) in $J_{p,q}^{(1)}$ and at most $C2^{2j}$ pairs of (p,q) in $J_{p,q}^{(1)}$. The following two lemmata are the key estimates for the proof of Theorem 2.1. The ideas used below are inspired by the arguments developed by the authors for the shearlet-based analysis of singularities in [7, 9].

Lemma 3.6. *For a given $\epsilon > 0$ and for (p,q) in $J_{p,q}^{(1)}$, we have that*

$$\sum_{(p,q)\in J_{p,q}^{(1)}} \sum_{m=1}^{M} \sum_{|\ell_1|\leq 2^j} \sum_{|\ell_2|\leq 2^j} \sum_{k\in\mathbb{Z}^3} |\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)| \leq C\, 2^{2\epsilon j} 2^{-\frac{1}{2}j}.$$

Proof. Let L be the differential operator :

$$L = \left(I - (\frac{2^{2j}}{2\pi})^2 \frac{\partial^2}{\partial\xi_1^2}\right) \left(1 - (\frac{2^j}{2\pi})^2 \frac{\partial^2}{\partial\xi_2^2}\right)^2 \left(1 - (\frac{2^j}{2\pi})^2 \frac{\partial^2}{\partial\xi_3^2}\right).$$

A direct computation shows that

$$\beta_j^{(m,p,q)}(\ell_1,\ell_2,k) = \langle \widehat{\mathcal{T}}_j^{(m,p,q)}, \widehat{\psi_{j,\ell_1,\ell_2,k}} \rangle$$

$$= \frac{2^{-2j}i}{2\pi} \int_D g(2^j x_2 - p)g(2^j x_3 - p) \int_{U_{j,\ell}} \frac{\xi \cdot (-1, A, B)}{|\xi|} W(2^{-2j}\xi)\, \Gamma_{j,\ell_1,\ell_2}(\xi)$$

$$\times e^{2\pi i \xi \cdot (2^{-2j}(k_1 - \ell_1 k_2 - \ell_2 k_3 - 2^{2j}(Ax_2 + Bx_3)), 2^{-j}(k_2 - 2^j x_2)), 2^{-j}(k_3 - 2^j x_2))}\, d\xi\, dx_3\, dx_2$$

$$= \frac{2^{-2j}i}{2\pi} \int_D g(2^j x_2 - p)g(2^j x_3 - p) \int_{U_{j,\ell}} L^N \left(\frac{\xi \cdot (-1, A, B)}{|\xi|} W(2^{-2j}\xi)\, \Gamma_{j,\ell_1,\ell_2}(\xi) \right)$$

$$\times L^{-N} \left(e^{2\pi i \xi \cdot (2^{-2j}(k_1 - \ell_1 k_2 - \ell_2 k_3 - 2^{2j}(Ax_2 + Bx_3)), 2^{-j}(k_2 - 2^j x_2)), 2^{-j}(k_3 - 2^j x_2))} \right) d\xi\, dx_3\, dx_2.$$

It is easy to verify that, for $\xi \in U_{j,\ell_1,\ell_2}$, we have:

$$\left| L^N \left(\frac{\xi \cdot (-1, A, B)}{|\xi|} W(2^{-2j}\xi)\, \Gamma_{j,\ell_1,\ell_2}(\xi) \right) \right| \leq C_N\, 2^{-2j}.$$

Since $|U_{j,\ell}| \leq C\, 2^{4j}$, it follows that

$$|\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)| \leq C_N \int_D \left(\left(1 + (k_1 - \ell_1 k_2 - \ell_2 k_3 - 2^{2j}(Ax_2 + Bx_3))^2\right) \right.$$

$$\times \left. \left(1 + (k_2 - 2^j x_2)^2\right) \left(1 + (k_3 - 2^j x_3)^2\right) \right)^{-N} dx_3 dx_2.$$

For any given $\epsilon > 0$, if $|k_2 - 2^j x_2| > 2^{\epsilon j}$ or $|k_3 - 2^j x_3| > 2^{\epsilon j}$ or $|k_1 - k_2\ell_1 - k_3\ell_2 - 2^{2j}(Ax_2 + Bx_3)| > 2^{\epsilon j}$ for all $(x_2, x_3) \in I_{p,q}$, then the above inequality with sufficient large N yields the lemma. Thus, let us examine the situation where instead we have that $|k_2 - 2^j x_2| \leq 2^{\epsilon j}$, $|k_3 - 2^j x_3| \leq 2^{\epsilon j}$ and $|k_1 - k_2\ell_1 - k_3\ell_2 - 2^{2j}(Ax_2 + Bx_3)| \leq 2^{\epsilon j}$ for some $(x_2, x_3) \in I_{p,q}$. Since $I_{p,q} = (2^{-j}p - 2^{-j}, 2^{-j}p + 2^{-j}) \times (2^{-j}q - 2^{-j}, 2^{-j}q + 2^{-j})$, it follows that $2^j x_2 = p + \alpha(x_2)$, $2^j x_3 = q + \beta(x_3)$, and $2^{2j}(Ax_2 + Bx_3) = 2^j(Ap + Bq) + \gamma(x_2, x_3)$ with $|\alpha(x_2)| \leq 1$, $|\beta(x_3)| \leq 1$ and $|\gamma(x_2, x_3)| \leq C2^j$ for all $(x_2, x_3) \in I_{p,q}$. Let $S_1 = \{k_2 : |k_2 - 2^j x_2| \leq 2^\epsilon\}$, $T_1 = \{k_3 : |k_3 - 2^j x_3| \leq 2^\epsilon\}$, for some $(x_2, x_3) \in I_{p,q}$. Let $S_2 = \{k_2 : |k_2 - \alpha(x_2)| \leq 2^\epsilon\}$, $T_2 = \{k_3 : |k_3 - \beta(x_3)| \leq 2^\epsilon\}$, for some $(x_2, x_3) \in I_{p,q}$. Since $2^j x_2 = p + \alpha(x_2)$, $2^j x_3 = q + \beta(x_3)$, it follows that $\text{card}\,(S_1) = \text{card}\,(S_2)$ and $\text{card}\,(T_1) = \text{card}\,(T_2)$. For any $k_2 \in S_2$, $k_3 \in T_2$, we have that $|k_2| \leq 2^{\epsilon j} + |\alpha(x_2)| \leq C2^{\epsilon j}$, $|k_3| \leq 2^{\epsilon j} + |\beta(x_3)| \leq C2^{\epsilon j}$. Thus, $\text{card}\,(S_1) = \text{card}\,(S_2) \leq C2^{\epsilon j}$ and $\text{card}\,(T_1) = \text{card}\,(T_2) \leq C2^{\epsilon j}$ uniformly for $(x_2, x_3) \in I_{p,q}$. Note that, for $k_2 \in S_1$, $k_3 \in T_1$, we have that $k_2 = k_2 - p + p = k_2' + p$, $k_3 = k_3 - q + q = k_3' + q$. Since $k_2' \in S_2$, $k_3' \in T_2$, we have $|k_2'| \leq C2^{\epsilon j}$, $|k_3'| \leq C2^{\epsilon j}$.

Similarly, for fixed k_2, k_3, let $J_1 = \{k_1 : |k_1 - k_2\ell_1 - k_3\ell_2 - 2^{2j}(Ax_2 + Bx_3)| \leq 2^{\epsilon j}\}$, $J_2 = \{k_1 : |k_1 - k_2'\ell_1 - k_3'\ell_2 - 2^{2j} - \gamma(x_2, x_3)| \leq 2^{\epsilon j}\}$, for some $(x_2, x_3) \in I_{p,q}$ so that $\text{card}\,(J_1) = \text{card}\,(J_2)$. For $k_1 \in J_2$, we have $|k_1 - k_2'\ell_1 - k_3'\ell_2 - \gamma(x_2, x_3)| \leq 2^{\epsilon j}$. It follows that

$$|k_1| \leq |k_2'\ell_1| + |k_3'\ell_2| + \gamma(x_2, x_3) + 2^{\epsilon j} \leq C2^{(1+\epsilon)j},$$

for all $x_2, x_3 \in I_{p,q}$ such that $|\ell_1| \leq 2^j, |\ell_2| \leq 2^j$. We have shown that $\mathrm{card}\,(J_1) = \mathrm{card}\,(J_2) \leq C2^{(1+\epsilon)j}$, uniformly for all $(x_2, x_3) \in I_{p,q}$ and all $|\ell_1| \leq 2^j, |\ell_2| \leq 2^j$. Thus, for the rest of the proof of the lemma, for $k = (k_1, k_2, k_3)$, we can assume that $\mathrm{card}\,(k) \leq C\,2^{(1+3\epsilon)j}$.

We recall that

$$
\begin{aligned}
\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi) &= \frac{-W(-2^j\xi)}{(2\pi)^2|\xi|}\frac{\xi}{|\xi|}\cdot(-1, A, B)\int_D g(2^j x_2 - p)g(2^j x_3 - q)e^{-2\pi i\eta\cdot(x_2, x_3)}\,dx_3\,dx_2 \\
&= \frac{-2^{-2j}W(-2^j\xi)}{(2\pi)^2|\xi|}\frac{\xi}{|\xi|}\cdot(-1, A, B)\int_{2^j D} g(x_2 - p)g(x_3 - q)e^{-2\pi i2^{-j}\eta\cdot(x_2, x_3)}\,dx_3\,dx_2.
\end{aligned}
$$

Similarly to (3.13), we have the following estimate:

$$
|\widehat{\mathcal{T}_j^{(m,p,q)}}(\xi)| \leq C\frac{2^{-2j}}{|\xi|}\frac{1}{1 + 2^{-j}|B\xi_1 + \xi_3|}\frac{1}{1 + 2^{-j}|A\xi_1 + \xi_2 + b(B\xi_1 + \xi_3)|}.
$$

It follows that, for $\xi \in U_{j,\ell_1,\ell_2}$, we have that

$$
|\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi)| \leq C\,2^{-4j}(1 + |\ell_2 - \ell_{2,m}|)^{-1}\left(1 + |(\ell_1 - \ell_{1,m}) + b(\ell_2 - \ell_{2,0})|\right)^{-1}.
$$

Using this last observation, together with the fact that the family

$$
\{2^{-2j}e^{2\pi i\xi A_{(1)}^{-j}B_{(1)}^{-\ell_1,\ell_2}k} : k \in \mathbb{Z}^3\}
$$

is an orthonormal basis for $L^2(U_{j,\ell_1,\ell_2})$, it follows that

$$
\begin{aligned}
\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)|^2 &= \int_{U_{j,\ell_1,\ell_2}}|W(2^{-2j}\xi)\,\Gamma_{j,\ell_1,\ell_2}(\xi)|^2\,|\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi)|^2\,d\xi \\
&\leq 2^{4j}2^{-8j}(1 + |\ell_2 - \ell_{2,m}|)^{-2}(1 + |(\ell_1 - \ell_{1,m}) + b(\ell_2 - \ell_{2,m})|)^{-2}.
\end{aligned}
$$

Since the cardinality of the $k \in \mathbb{Z}^3$ involved in the above sum is of order $O(2^{1+3\epsilon)j})$, Hölder's inequality yields that

$$
\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)| \leq C\,2^{\frac{1+3\epsilon}{2}j}2^{-2j}(1 + |\ell_2 - \ell_{2,m}|)^{-1}(1 + |(\ell_1 - \ell_{1,m}) + b(\ell_2 - \ell_{2,m})|)^{-1}.
$$

It follows that

$$
\sum_{|\ell_1|\leq 2^j}\sum_{|\ell_2|\leq 2^j}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)| \leq C\,j^2\,2^{-\frac{3}{2}j}2^{\frac{3}{2}j} \leq C\,2^{2\epsilon j}2^{-\frac{3}{2}j}.
$$

Hence, computing the sum over the indices m and the sum over (p, q) in $J_{p,q}^{(1)}$, we have that

$$
\sum_{(p,q)\in J_{p,q}^{(1)}}\sum_{m=1}^M\sum_{|\ell_1|\leq 2^j}\sum_{|\ell_2|\leq 2^j}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)| \leq C\,2^{2\epsilon j}2^{-\frac{1}{2}j}.
$$

This completes the proof of the lemma. $\square$

Lemma 3.7. *Using the notation that we have introduced above, for a given $\epsilon > 0$ and for (p, q) in $J_{p,q}^{(2)}$, we have that:*

$$
\sum_{(p,q)\in J_{p,q}^{(2)}}\sum_{m=1}^M\sum_{|\ell_2 - \ell_{2,0}|\geq 2^{\epsilon j}}\sum_{|\ell_2|\leq 2^j}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)| \leq C\,2^{2\epsilon j}2^{-\frac{1}{2}j}
$$

$$
\sum_{(p,q)\in J_{p,q}^{(2)}}\sum_{m=1}^M\sum_{|\ell_1|\leq 2^j}\sum_{|(\ell_1 - \ell_{1,0}) + b(\ell_2 - \ell_{2,0})|\geq 2^{\epsilon j}}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1, \ell_2, k)| \leq C2^{2\epsilon j}2^{-\frac{1}{2}j}
$$

Proof. We will only prove the first of the two estimates since the second one can be proved using a similar argument.

As in the proof of Lemma 3.6, we use the following identity:

$$\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi) = -\frac{2^{-2j}W(2^{-2j}\xi)}{(2\pi)^2|\xi|}\frac{\xi}{|\xi|}\cdot(-1,A,B)\int_{2^jD}g(x_2-p)g(x_3-q)e^{-2\pi i2^{-j}\eta\cdot(x_2,x_3)}\,dx_3\,dx_2.$$

Since $I_{p,q}\bigcap\partial D = \emptyset$, we see that $\mathrm{supp}\,(g(x_2-p)g(x_3-q))\bigcap\partial 2^jD = \emptyset$. We also have that $|\xi|\simeq 2^{2j}$ on U_{j,ℓ_1,ℓ_2}. Therefore, using integration by parts N times with respect to the variable x_2 and x_3 in the integral above, we have that

$$|\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi)| \le C\,2^{-4j}(1+|\ell_2-\ell_{2,m}|)^{-N}\,(1+|(\ell_1-\ell_{1,m})+b(\ell_2-\ell_{2,m})|)^{-N}\,.$$

It follows that

$$
\begin{aligned}
\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)|^2 &\le \int_{\Omega_{j,\ell_1,\ell_2}}|\Gamma_{j,\ell_1,\ell_2}(\xi)\,W(2^{-2j}\xi)|^2|\,\widehat{\mathcal{T}}_j^{(m,p,q)}(\xi)|^2\,d\xi \\
&\le 2^{4j}2^{-8j}(1+|\ell_2-\ell_{2,m}|)^{-2N}\,(1+|(\ell_1-\ell_{1,m})+b(\ell_2-\ell_{2,m})|)^{-2N}\,.
\end{aligned}
$$

As in the proof of Lemma 3.6, we have that card $(k) = O(2^{1+3\epsilon})$ uniformly for all (p,q). Hence, from the last inequality, we have that

$$\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)| \le C_N\,2^{\frac{1+3\epsilon}{2}l}2^{-2j}(1+|\ell_2-\ell_{2,m}|)^{-N}\,(1+|(\ell_1-\ell_{1,m})+b(\ell_2-\ell_{2,m})|)^{-N}\,.$$

This implies that

$$\sum_{|\ell_2-\ell_{2,m}|\ge 2^{\epsilon j}}\sum_{|\ell_1|\le 2^j}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)| \le C_N\,2^{\frac{1+3\epsilon}{2}j}2^{-2j}2^{(N-1)\epsilon j}.$$

It follows that, for N large enough, we have the estimate:

$$
\begin{aligned}
\sum_{(p,q)\in J_{p,q}^{(2)}}\sum_{m=1}^M\sum_{|\ell_2-\ell_{2,m}|\ge 2^{\epsilon j}}\sum_{|\ell_1|\le 2^j}\sum_{k\in\mathbb{Z}^3}|\beta_j^{(m,p,q)}(\ell_1,\ell_2,k)| &\le C_N\,2^{2j}2^{\frac{1+3\epsilon}{2}j}2^{-2j}2^{(N-1)\epsilon j} \\
&\le C\,2^{2\epsilon j}2^{-\frac{1}{2}j}.\qquad\square
\end{aligned}
$$

We are now ready to prove Theorem 2.1.

We choose any $0<\epsilon<\frac{1}{8}$ and we introduce the notation

$$S_{2,j}^{(m)} = \{(j,\ell_1,\ell_2,m,k): |\ell_1-\ell_{1,m}|\le C2^{\epsilon j},\,|\ell_2-\ell_{2,m}|\le 2^{\epsilon j},\,k\in\mathbb{Z}^3\}.$$

We have that $S_{2,j} = \bigcup_{m=1}^M S_{2,j}^{(m)}$. Since $|\ell_2-\ell_{2,m}|\le 2^{\epsilon j}$ and $|(\ell_1-\ell_{1,m})+b(\ell_2-\ell_{2,m})|\le 2^{\epsilon j}$ yield $|\ell_1-\ell_{1,m}|\le C2^{\epsilon j}$, it follows from Lemma 3.6 and Lemma 3.7 that

$$\delta_{2,j} = \sum_{\eta\in S_{2,j}^c}|\langle\mathcal{T}_j,\psi_\eta\rangle| \le C\,M\,2^{2\epsilon j}\,2^{-\frac{1}{2}j}.$$

Thus, since $0<\epsilon<\frac{1}{8}$, Lemma 3.5 implies that $\delta_{2,j} = o(\|\mathcal{P}_j\|_{1,\Phi}+\|\mathcal{T}_j\|_{1,\Psi})$.

As we pointed out above, to complete the proof of Theorem 2.1, it remains to show that $\mu_c(S_{2,j},\Psi;\Phi)\to 0$, as $j\to\infty$. Since

$$\mu_c(S_{2,j},\Psi;\Phi) = \max_\lambda\sum_{\eta\in S_{2,j}}|\langle\phi_\lambda,\psi_\eta\rangle| \le \max_\lambda\sum_{m=1}^M\sum_{\eta\in S_{2,j}^{(m)}}|\langle\phi_\lambda,\psi_\eta\rangle|,$$

it is sufficient to show that

$$\max_{\lambda} \sum_{\eta \in S_{2,j}^{(m)}} |\langle \phi_\lambda, \psi_\eta \rangle| \to 0, \quad \text{as } j \to \infty.$$

This means that the proof of Theorem 2.1 is completed once we show the following result.

Proposition 3.8. *Using the notation from above, for a given $\epsilon > 0$, we have that*

$$\max_{k' \in \mathbb{Z}^3} \sum_{|\ell_1 - \ell_{1,m}| \leq C 2^{\epsilon j}} \sum_{|\ell_2 - \ell_{2,m}| \leq 2^{\epsilon j}} \sum_{k \in \mathbb{Z}^3} |\langle \phi_{j,k'}, \psi_{j,\ell_1,\ell_2,k} \rangle| \to 0, \quad \text{as } j \to \infty.$$

Proof. Let L be the differential operator

$$L_1 = \left(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial \xi_1^2} \right) \left(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial \xi_2^2} \right) \left(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial \xi_3^2} \right).$$

For brevity, let

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = B_{(1)}^{[\ell_1, \ell_2]} A_{(1)}^j (2^{-2j} k') = (k_1' + 2^{-j} \ell_1 k_2' + 2^{-j} \ell_2 k_3', 2^{-j} k_2', 2^{-j} k_3').$$

By direct calculation, we have that

$$
\begin{aligned}
\langle \widehat{\psi_{j,\ell,k}}, \widehat{\phi_{j,k'}} \rangle &= \int_{\mathbb{R}^3} \left(2^{-2j} \Gamma_{j,\ell_1,\ell_2}(\xi) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k} \right) \left(2^{-3j} W(2^{-2j}\xi) e^{-2\pi i 2^{-2j}\xi \cdot k'} \right) d\xi \\
&= 2^{-5j} \int_{R^3} \Gamma_{j,\ell_1,\ell_2}(\xi) W(2^{-2j}\xi) e^{2\pi i \xi [A_{(1)}^{-j} B_{(1)}^{[-\ell]}(k-\alpha)]} d\xi \\
&= 2^{-j} \int_{\mathbb{R}^3} \hat{\psi}_2(\tfrac{\eta_2}{\eta_1}) \hat{\psi}_2(\tfrac{\eta_3}{\eta_1}) W^2(\eta_1, 2^{-j}(\ell_1\eta_1 + \eta_2), 2^{-j}(\ell_2\eta_1 + \eta_3)) e^{2\pi i \eta \cdot (k-\alpha)} d\eta \\
&= 2^{-j} \int_{\mathbb{R}^3} L_1 \left(\hat{\psi}_2(\tfrac{\eta_2}{\eta_1}) \hat{\psi}_2(\tfrac{\eta_3}{\eta_1}) W^2(\eta_1, 2^{-j}(\ell_1\eta_1 + \eta_2), 2^{-j}(\ell_2\eta_1 + \eta_3)) \right) L_1^{-1} \left(e^{2\pi i \eta \cdot (k-\alpha)} \right) d\eta.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\sum_{|\ell_1 - \ell_{1,m}| \leq C 2^{\epsilon j}} \sum_{|\ell_2 - \ell_{2,m}| \leq 2^{\epsilon j}} \sum_{k \in \mathbb{Z}^3} |\langle \phi_{j,k'}, \psi_{j,\ell_1,\ell_2,k} \rangle| \\
&\leq C 2^{2\epsilon j} 2^{-j} \sum_{k \in \mathbb{Z}^3} (1 + (k_1 - \alpha_1)^2)^{-1} (1 + (k_2 - \alpha_2)^2)^{-1} (1 + (k_3 - \alpha_3)^2)^{-1} \\
&\leq C 2^{(-1+2\epsilon)j},
\end{aligned}
$$

where the constant $C > 0$ is independent of k'. Since $\epsilon < \frac{1}{8}$, it follows that $-1 + 2\epsilon < 0$, and this implies that

$$\mu_c(S_{2,j}, \Psi; \Phi) \to 0 \quad \text{as } j \to \infty.$$

This completes the proof of the proposition. $\square$

This also completes the proof of Theorem 2.1.

3.3 Remarks

The approach for geometric separation originally introduced in [4] deals with 2D images containing point-like and smooth curve-like singularities, but does not handle curve-like singularities containing corner points. The approach presented in our paper can be easily adapted to the two-dimensional case so that Theorem 2.1 is indeed valid also in the case of 2D images containing point-like and piecewise linear singularities (including corner points).

Concerning the extension of our 3D result to the situation of singularities along general piece-wise smooth surfaces, we remark that our arguments does not work for general surfaces (e.g., a section of a sphere). In fact, the proof of Lemma 3.6 and Lemma 3.7 relies on the crucial observation that all values of p and q from 2^{-j} to 2^j correspond to the same $\ell_{1,m}$ and $\ell_{2,m}$, and this requires the assumption that the singularity surfaces are flat, i.e., linear or piece-wise linear. As a consequence, a different argument would be needed to deal with singularities along more general surfaces.

References

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

[2] S. S. Chen, D. L. Donoho and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Rev.*, 43:129–159, 2001.

[3] R. R. Coifman and M. V. Wickerhauser. Wavelets and adapted waveform analysis. A toolkit for signal processing and numerical analysis, Different perspectives on wavelets (San Antonio, TX, 1993), 119–153. *Proc. Sympos. Appl. Math.*, 47, Amer. Math. Soc., Providence, RI, 1993.

[4] Donoho and Kutyniok. Microlocal analysis of the geometric separation problem. *Comm. Pure Appl. Math.*, 66:1–47, 2013.

[5] R. Gribonval and E. Bacry. Harmonic decomposition of audio signals with matching pursuit. *IEEE Trans. Signal Proc.*, 51:1001-111, 2003.

[6] K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM J Math. Anal.*, 39:298–318, 2007.

[7] K. Guo and D. Labate. Characterization and analysis of edges using the continuous shearlet transform. *SIAM J. Imaging Sci.*, 2:959–986, 2009.

[8] K. Guo and D. Labate. Analysis and Detection of Surface Discontinuities using the 3D Continuous Shearlet Transform. *Appl. Comput. Harmon. Anal.*, 30:231–242, 2010.

[9] K. Guo and D. Labate. Optimally sparse representations of 3D data with C2 surface singularities using Parseval frames of shearlets. *SIAM J Math. Anal.*, 44:851–886, 2012.

[10] K. Guo and D. Labate. Characterization of Piecewise Smooth Surfaces using the 3D Continuous Shearlet Transform. *J. Fourier Anal. Appl.*, 18:488–516, 2012.

[11] K. Guo and D. Labate. The construction of smooth Parseval frames of shearlets. *Math. Model. Nat. Phenom.*, 8:82-105, 2013.

[12] K. Guo, D. Labate and W. Lim. Edge analysis and identification using the continuous shearlet transform. *Appl. Comput. Harmon. Anal.*, 27:24–46, 2009.

[13] G. Garrigòs and E. Hernandez. Sharp Jackson and Bernstein inequalities for N-term approximation in sequence spaces with applications. *Indiana Univ. Math. J.*, 53:1739–1762, 2004.

[14] E. Hernandez, and G. Weiss. *A first course on wavelets.* CRC Press,Boca Raton, FL, 1996

[15] G. Kutyniok, Geometric Separation by Single-Pass Alternating Thresholding. *Appl. Comput. Harmon. Anal.*, 36:23–50, 2014.

[16] G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Trans. Amer. Math. Soc.*, 361:2719–2754, 2009.

[17] D. Labate, W. Lim, G. Kutyniok and G. Weiss. Sparse multidimensional representation using shearlets. *SPIE Proc.*, 5914: 254–262, 2005.

[18] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Proc.*, 41:3397–3415, 1993.

[19] F. G. Meyer, A. Averbuch and R. R. Coifman. Multi-layered Image Representation: Application to Image Compression. *IEEE Trans. Image Process.*, 11:1072–1080, 2002.

[20] J.-L. Starck, M. Elad and D.L. Donoho. Redundant Multiscale Transforms and their Application for Morphological Component Analysis. *Adv. Imag. Elect. Phys.*, 132:287–348, 2004.

[21] J.-L. Starck, M. Elad and D.L. Donoho. Image Decomposition Via the Combination of Sparse Representation and a Variational Approach. *IEEE Trans. Image Process.*, 14:1570–1582, 2005.

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

[23] J.-L. Starck, M. Nguyen and F. Murtagh, Wavelets and curvelets for image deconvolution: A combined approach. *Signal Process.*, 83:2279–2283, 2003.

[24] R. S. Strichartz. *A guide to distribution theory and Fourier transforms*. World Scientific Pub., 2003.

[25] G. Teschke. Multi-frame representations in linear inverse problems with mixed multi-constraints. *Appl. Comput. Harmon. Anal.*, 22:43–60, 2007.