5-28-2012

Optimally sparse representations of 3D data with C^2 surface singularities using parseval frames of shearlets

Kanghui Guo
Missouri State University

Demetrio Labate

OPTIMALLY SPARSE REPRESENTATIONS OF 3D DATA WITH C^2 SURFACE SINGULARITIES USING PARSEVAL FRAMES OF SHEARLETS

KANGHUI GUO* AND DEMETRIO LABATE†

Abstract. This paper introduces a Parseval frame of shearlets for the representation of 3D data, which is especially designed to handle geometric features such as discontinuous boundaries with very high efficiency. This system of 3D shearlets forms a multiscale pyramid of well-localized waveforms at various locations and orientations, which become increasingly thin and elongated at fine scales. We prove that this 3D shearlet construction provides essentially optimal sparse representations for functions on $\mathbb{R}^3$ which are C^2-regular away from discontinuities along C^2 surfaces. As a consequence, we show that within this class of functions the N-term approximation f_N^S obtained by selecting the N largest coefficients of the shearlet expansion of f satisfies the asymptotic estimate

$$\|f - f_N^S\|_2^2 \asymp N^{-1}(\log N)^2, \qquad \text{as } N \to \infty.$$

This asymptotic behavior significantly outperforms wavelet and Fourier series approximations which only yield an approximation rate of $O(N^{-1/2})$ and $O(N^{-1/3})$, respectively. This result extends to the 3D setting the (essentially) optimally sparse approximation results obtained by the authors using 2D shearlets and by Candès and Donoho using curvelets and is the first nonadaptive construction to provide provably (nearly) optimal representations for a large class of 3-dimensional data.

Key words. Affine systems, curvelets, nonlinear approximations, shearlets, sparsity, wavelets.

AMS subject classifications. 42C15, 42C40

1. Introduction.

Sparse representations of multidimensional data have gained more and more prominence in recent years as a variety of applied problems require to process massive and multi-dimensional data sets in a timely and effective manner. This is a major challenge in applications such as remote sensing, satellite imagery, scientific simulations and electronic surveillance. Sparse representations enable not only to accurately and reliably compress data and expedite their transmission and storage, but also to develop more effective algorithms for tasks such as feature extraction and pattern recognition. In fact, constructing sparse representations for data in a certain class entails the intimate understanding of their true nature and structure [10].

Wavelets and other traditional multiscale methods have been extremely successful during the past 20 years because of their ability to provide optimally sparse representations for data with point singularities. This property was exploited to develop a number of impressive applications in signal and image processing. Wavelets, however, are not equally efficient when dealing with distributed discontinuities, and this is a major limitation in multidimensional applications where edges and discontinuous boundaries are frequently the dominant features of the objects to be analyzed. This inefficiency of wavelets in dealing with distributed singularities is due to their isotropic nature, which hampers the ability to really capture the geometry of edges and other essential features of multidimensional data. To overcome these limitations, a new generation of multiscale systems was introduced in recent years, most notably the *curvelets* [2], the *contourlets* [8] and the *shearlets* [13, 14], which are especially designed to represent efficiently anisotropic features in images. The intuitive idea behind their construction is that, in order to deal efficiently with the edges and the other anisotropic features which are prominent in most images of practical interest, the analyzing elements must be defined not only at various locations and scales, as traditional wavelets, but also at various orientations and with highly anisotropic shapes. Thanks to their geometrical properties, the curvelet and shearlet representations turn out to be essentially as good as an adaptive representation from the point of view of their ability to approximate images containing edges. Specifically, for functions f which are C^2 away from C^2 edges, the N term approximation f_N^S obtained from the N largest coefficients in its curvelet or shearlet expansion, obeys

$$\|f - f_N^S\|_2^2 \asymp N^{-2}(\log N)^3, \qquad \text{as } N \to \infty. \tag{1.1}$$

*Department of Mathematics, Missouri State University, Springfield, Missouri 65804 (KanghuiGuo@MissouriState.edu).
†Department of Mathematics, University of Houston, Houston, Texas 77204 (dlabate@math.uh.edu)

Ignoring the loglike factor, this is the optimal approximation rate for this class of functions, as claimed in [2]; in comparison, the wavelet and Fourier representations only achieve approximation rate N^{-1} and $N^{-1/2}$, respectively.

The goal of this paper is to extend to the 3D setting the remarkable optimal approximation result achieved for piecewise smooth functions of 2 variables. This new 3D result will be derived using a system of 3D shearlets which forms a Parseval frame of well-localized functions defined at various locations, scale and orientations.

Notice that a number of 3D multiscale directional constructions have been already proposed in the literature, including the *3D curvelets* in [1] and the *surfacelets* in [26]. In all these cases, the focus of these constructions is the numerical implementation. Also 3D shearlets have been already introduced in the literature (e.g., [5, 6, 14]) and some of their microlocal properties have been recently analyzed by the authors in [16]. However, no rigorous analysis of the sparsity properties of curvelets or shearlets or any other similar system in the 3D setting has been published so far and, in particular, there is no proof of the analogue of estimate (1.1) for the 3D setting. The extension of this result to 3D is highly nontrivial since the proof of the (almost) optimal sparsity does not follow directly from the arguments used in the 2-variable case and, as it will be apparent from our presentation below, this analysis requires to introduce some fundamentally new tools.

In this paper, we prove that Parseval frames of 3D shearlets provide essentially optimal sparse representations for piecewise smooth function of 3 variables. As a consequence of our result we show that, denoting by f_N^S the shearlet approximation of f which is obtained from the N largest coefficients of its shearlet representation, the approximation error satisfies

$$\|f - f_N^S\|_2^2 \asymp N^{-1}(\log N)^2, \qquad \text{as } N \to \infty. \tag{1.2}$$

This is the first published proof for a result of this type and it is the analogue of estimate (1.1) in the 3D setting. In a certain sense which will be made precise below, the rate N^{-1} is the best rate achievable. Notice, in particular, that the approximation error rate (1.2) obtained using the shearlet representation significantly outperforms wavelet and Fourier approximations, whose asymptotic approximation error rates are of the order of $N^{-1/2}$ and $N^{-1/3}$, respectively.

As an additional remark, it is important to emphasize that the approach presented in this paper is purely non-adaptive. Adaptive approximations of multidimensional piecewise smooth functions can be found in [3, 25, 27, 31, 32]. Remarkably, for the class of functions considered in this paper, the shearlet approach is as effective as an adaptive representation with respect to its ability to approximate 3D data with discontinuous boundaries.

Concerning the comparison of shearlet and curvelet representations, one prominent difference is that shearlets use shear matrices rather than rotations to control the directional features of the representation system. This is more 'natural' in discrete implementations, since shear matrices, unlike rotations, preserve the integer lattice. In fact, it is useful to recall that the so-called *digital curvelets* introduced in [1] to derive a digital implementation of the curvelet transform use shearing rather than rotations. In this respect, the shearlet approach ensures a unified framework for both the continuum to the digital setting [19, 20, 22]. Digital implementations of the shearlet representation which are faithful its continuous-domain counterparts are found in [12, 23] for the 2D setting and in [29] for the 3D one.

Remark. During the final editing of this paper, a similar (essentially) optimal sparsity result was announced, without proof, by Kutyniok, Lemvig and Lim, based on a new remarkable construction of compactly supported shearlet frames [24]. This approach considers frames which are not Parseval frames and extends the corresponding 2D approach introduced by the same authors in [21]. Despite the fact that such frames are not tight, the ability to have compactly supported analyzing functions is an advantage in some applications. While the proof of this result is not available at this time, we expect it to be very different from the one found in this paper, due to the very different construction of the analyzing system (cf. the 2D proof in [21]).

1.1. Outline. The paper is organized as follows. The construction of the 3D Parseval frame of shearlets is presented in Section 2. The main results of the paper are given in Section 3. The technical constructions needed for the proofs are collected in Section 4. Finally, Section 5 discusses the general issue of optimally sparse approximations and the theoretical best error approximation rate that can be achieved in 3D.

2. The shearlet representation. The shearlet representation, originally derived from the framework of wavelets with composite dilations [18, 19], provides a general method for the construction of function systems made up of waveforms ranging not only at various scales and locations, as traditional wavelets, but also at various orientations. Thanks to the ability of the shearlet systems to deal with directionality and anisotropy, the geometric content of multivariate functions and data is captured much more efficiently than using wavelets or other traditional representation methods. In addition, as mentioned above, the use of shear matrices enables shearlets to provide a unified treatment of the continuum and digital setting. These properties and the special flexibility of the shearlet framework made shearlets very successful in several imaging applications [4, 11, 12, 15, 30, 34].

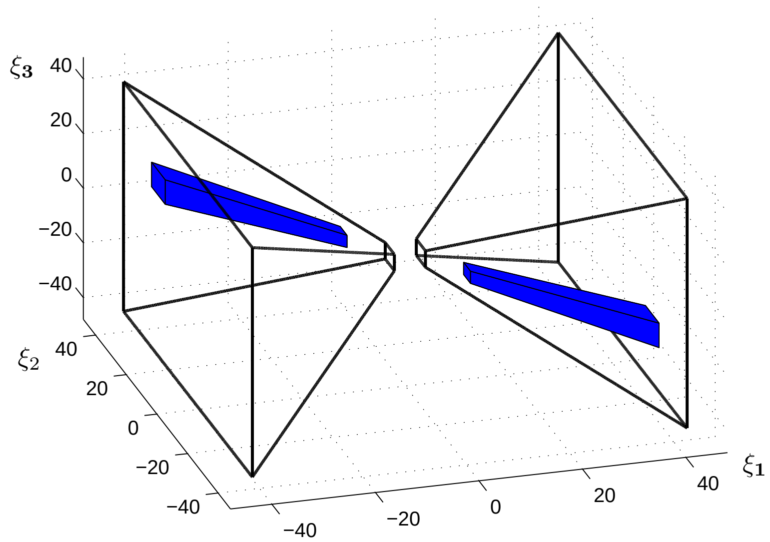

Fig. 2.1. *Frequency support of a representative shearlet function $\psi^{(1)}_{j,\ell,k}$, inside the pyramidal region $\mathcal{P}_1$. The orientation of the support region is controlled by $\ell = (\ell_1, \ell_2)$; its shape is becoming more elongated as j increases ($j = 4$ in this plot)*

The construction of 3D shearlets presented below is similar to the digital curvelets from [1]. An alternative way to construct smooth Parseval frames of shearlets is discussed in Sec. 5.2

In dimension $D = 3$, a shearlet system is obtained by appropriately combining 3 systems of functions associated with the pyramidal regions

$$\mathcal{P}_1 = \left\{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : |\tfrac{\xi_2}{\xi_1}| \le 1, |\tfrac{\xi_3}{\xi_1}| \le 1\right\},$$
$$\mathcal{P}_2 = \left\{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : |\tfrac{\xi_1}{\xi_2}| < 1, |\tfrac{\xi_3}{\xi_2}| \le 1\right\},$$
$$\mathcal{P}_3 = \left\{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : |\tfrac{\xi_1}{\xi_3}| < 1, |\tfrac{\xi_2}{\xi_3}| < 1\right\},$$

in which the 3D Fourier space is partitioned.

To define such systems, let ϕ be a C^∞ univariate function such that $0 \le \hat{\phi} \le 1$, $\hat{\phi} = 1$ 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 $\xi=(\xi_1,\xi_2,\xi_3)\in\mathbb{R}^3$, 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

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

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

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

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

$$\sum_{j\geq 0}W^2(2^{-2j}\xi)=1 \quad \text{for } \xi\in\mathbb{R}^3\setminus[-\tfrac{1}{8},\tfrac{1}{8}]^3. \tag{2.3}$$

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

$$|v(u-1)|^2+|v(u)|^2+|v(u+1)|^2=1 \qquad \text{for } |u|\leq 1. \tag{2.4}$$

In addition, we will assume that $v(0)=1$ and that $v^{(n)}(0)=0$ for all $n\geq 1$. It was shown in [14] that there are several examples of functions satisfying these properties. It follows from equation (2.4) that, for any $j\geq 0$,

$$\sum_{m=-2^j}^{2^j}|v(2^j\,u-m)|^2=1, \qquad \text{for } |u|\leq 1. \tag{2.5}$$

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

$$\{\psi^{(d)}_{j,\ell,k}:\, j\geq 0, -2^j\leq \ell_1,\ell_2\leq 2^j, k\in\mathbb{Z}^3\}, \tag{2.6}$$

where

$$\hat{\psi}^{(d)}_{j,\ell,k}(\xi)=|\det A_{(d)}|^{-j/2}\,W(2^{-2j}\xi)\,V_{(d)}(\xi A_{(d)}^{-j}B_{(d)}^{[-\ell]})\,e^{2\pi i\xi A_{(d)}^{-j}B_{(d)}^{[-\ell]}k}, \tag{2.7}$$

$V_{(1)}(\xi_1,\xi_2,\xi_3)=v(\frac{\xi_2}{\xi_1})v(\frac{\xi_3}{\xi_1})$, $V_{(2)}(\xi_1,\xi_2,\xi_3)=v(\frac{\xi_1}{\xi_2})v(\frac{\xi_3}{\xi_2})$, $V_{(3)}(\xi_1,\xi_2,\xi_3)=v(\frac{\xi_1}{\xi_3})v(\frac{\xi_2}{\xi_3})$, 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

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

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

Due to the support conditions on W and v, the elements of the system of shearlets (2.6) have compact support in Fourier domain. In particular, for $d = 1$, the shearlets $\hat{\psi}^{(1)}_{j,\ell,k}(\xi)$ can be written more explicitly as

$$\hat{\psi}^{(1)}_{j,\ell_1,\ell_2,k}(\xi) = 2^{-2j}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(2^j \frac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi A^{-j}_{(1)} B^{[-\ell_1,-\ell_2]}_{(1)} k}, \tag{2.8}$$

showing that their supports are contained inside the trapezoidal regions

$$U_{j,\ell} = \{(\xi_1,\xi_2,\xi_3) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\frac{\xi_2}{\xi_1} - \ell_1 2^{-j}| \le 2^{-j}, |\frac{\xi_3}{\xi_1} - \ell_2 2^{-j}| \le 2^{-j}\}.$$

These support regions [1] become increasingly more elongated at fine scales, with the orientations controlled by ℓ_1, ℓ_2, as illustrated in Fig. 2.1.

A simple computation shows that the elements of the shearlets systems (2.7) can be written in space domain as

$$\psi^{(d)}_{j,\ell,k}(x) = |\det A_{(d)}|^{j/2}\, \psi^{(d)}_{j,\ell}(B^{[\ell]}_{(d)} A^j_{(1)} x - k), \tag{2.9}$$

for $j \ge 0$, $\ell = (\ell_1, \ell_2)$ with $\ell_1, \ell_2 \le 2^j, k \in \mathbb{Z}^3, d = 1,2,3$, where

$$\hat{\psi}^{(d)}_{j,\ell}(\xi) = W(2^{-2j} \xi B^{[\ell]}_{(d)} A^j_{(d)})\, V_{(d)}(\xi),$$

showing that the systems (2.7) are not affine-like. However, the functions $\psi^{(d)}_{j,\ell}$ depend very little on j, ℓ. Indeed, thanks to the support and regularity conditions on W and $V_{(d)}$, one can show [17] that for each $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in \mathbb{N}^3$ and each $N \ge 0$ there is a constant $C_{\gamma,N,d} > 0$ such that,

$$\left| \partial^\gamma_x \psi^{(d)}_{j,\ell}(x) \right| \le C_{\gamma,N,d}\, (1 + |x|)^{-N}, \tag{2.10}$$

with $C_{\gamma,N,d}$ independent of j, ℓ.

2.1. A smooth Parseval frame of shearlets for $L^2(\mathbb{R}^3)$. A Parseval frame of shearlets for $L^2(\mathbb{R}^3)$ is obtained by using an appropriate combination of the systems of shearlets associated with the 3 pyramidal regions $\mathcal{P}_d$, $d = 1,2,3$, together with a coarse scale system, which will take care of the low frequency region. In order to build such system in a way that all its elements are C^∞_c in the Fourier domain, the elements of the shearlet systems overlapping the boundaries of the pyramidal regions $\mathcal{P}_d$ in the Fourier domain have to be modified. Hence, we define the *3D shearlet systems for $L^2(\mathbb{R}^3)$* as the collections

$$\Big\{\widetilde{\psi}_{-1,k} : k \in \mathbb{Z}^3\Big\} \bigcup \Big\{\widetilde{\psi}_{j,\ell,k,d} : j \ge 0, |\ell_1| < 2^j, |\ell_2| \le 2^j, k \in \mathbb{Z}^3, d = 1,2,3\Big\}$$
$$\bigcup \Big\{\widetilde{\psi}_{j,\ell,k} : j \ge 0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3\Big\} \tag{2.11}$$

consisting of:

- the *coarse-scale shearlets* $\{\widetilde{\psi}_{-1,k} = \Phi(\cdot - k) : k \in \mathbb{Z}^3\}$, where Φ is given by (2.1);
- the *interior shearlets* $\{\widetilde{\psi}_{j,\ell,k,d} = \psi^{(d)}_{j,\ell,k} : j \ge 0, |\ell_1||\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$, where $\psi^{(d)}_{j,\ell,k}$ are given by (2.7);
- the *boundary shearlets* $\{\widetilde{\psi}_{j,\ell,k,d} : j \ge 0, |\ell_1| < 2^j, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$ and $\{\widetilde{\psi}_{j,\ell,k} : j \ge 0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3\}$, obtained by joining together slightly modified versions of $\psi^{(1)}_{j,\ell,k}$, $\psi^{(2)}_{j,\ell,k}$ and $\psi^{(3)}_{j,\ell,k}$, for $\ell_1, \ell_2 = \pm 2^j$. Their precise definition is given below.

[1] Notice that, since $|\ell_1|, |\ell_2| \le 2^j$, each support region $U_{j,\ell}$ is contained in a box of size $\approx 2^{2j} \times 2^j \times 2^j$ in the Fourier domain. Since the functions $\hat{\psi}^{(1)}_{j,\ell_1,\ell_2,k}$ are C^∞_c, it follows that in space domain their supports are essentially contained inside boxes of size $\approx 2^{-2j} \times 2^{-j} \times 2^{-j}$.

For $j \geq 1$, $|\ell_1| < 2^j$, $\ell_2 = \pm 2^j$, we define

$$(\widetilde{\psi}_{j,\ell_1,\ell_2,k,1})^\wedge(\xi) = \begin{cases} 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(2^j \frac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_1, \\ 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_2} - \ell_1\Big)\, v\Big(2^j \frac{\xi_3}{\xi_2} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_2; \end{cases} \tag{2.12}$$

$$(\widetilde{\psi}_{j,\ell_1,\ell_2,k,2})^\wedge(\xi) = \begin{cases} 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_2} - \ell_2\Big)\, v\Big(2^j \frac{\xi_3}{\xi_2} - \ell_1\Big)\, e^{2\pi i \xi 2^{-2} A_{(2)}^{-j} B_{(2)}^{[-(\ell_2,\ell_1)]} k}, & \text{if } \xi \in \mathcal{P}_2, \\ 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_3} - \ell_2\Big)\, v\Big(2^j \frac{\xi_2}{\xi_3} - \ell_1\Big)\, e^{2\pi i \xi 2^{-2} A_{(2)}^{-j} B_{(2)}^{[-(\ell_2,\ell_1)]} k}, & \text{if } \xi \in \mathcal{P}_3; \end{cases}$$

$$(\widetilde{\psi}_{j,\ell_1,\ell_2,k,3})^\wedge(\xi) = \begin{cases} 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_2}{\xi_1} - \ell_2\Big)\, v\Big(2^j \frac{\xi_3}{\xi_1} - \ell_1\Big)\, e^{2\pi i \xi 2^{-2} A_{(3)}^{-j} B_{(3)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_1, \\ 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_3} - \ell_1\Big)\, v\Big(2^j \frac{\xi_2}{\xi_3} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(3)}^{-j} B_{(3)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_3. \end{cases}$$

For $j \geq 1$, $\ell_1, \ell_2 = \pm 2^j$, we define

$$(\widetilde{\psi}_{j,\ell_1,\ell_2,k})^\wedge(\xi) = \begin{cases} 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(2^j \frac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_1, \\ 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_2} - \ell_1\Big)\, v\Big(2^j \frac{\xi_3}{\xi_2} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_2, \\ 2^{-2j-3}\, W(2^{-2j}\xi)\, v\Big(2^j \frac{\xi_1}{\xi_3} - \ell_1\Big)\, v\Big(2^j \frac{\xi_2}{\xi_3} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_3. \end{cases}$$

Similarly, for $j = 0$, $\ell_1, \ell_2 = 0, \ \pm 1$, we define

$$(\widetilde{\psi}_{0,\ell_1,\ell_2,k})^\wedge(\xi) = \begin{cases} W(\xi)\, v\Big(\frac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(\frac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_1, \\ W(\xi)\, v\Big(\frac{\xi_1}{\xi_2} - \ell_1\Big)\, v\Big(\frac{\xi_3}{\xi_2} - \ell_2\Big)\, e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_2, \\ W(\xi)\, v\Big(\frac{\xi_1}{\xi_3} - \ell_1\Big)\, v\Big(\frac{\xi_2}{\xi_3} - \ell_2\Big)\, e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_3. \end{cases}$$

Notice that the boundary shearlet functions are compactly supported in the Fourier domain by construction. In addition, it can be shown that they are C^∞ in the Fourier domain. In fact, let us consider the function $(\widetilde{\psi}_{j,2^j,\ell_2,k})^\wedge$, given by (2.12). To show that it is continuous, it is easy to verify that the two terms of the piecewise defined function are equal when $\xi_1 = \xi_2$ and $\xi_1 = \xi_3$. The smoothness is verified that checking the derivatives of these functions on the plane $\xi_1 = \xi_2 = \xi_3$. Specifically, we have that

$$\begin{aligned} &\frac{\partial}{\partial \xi_1}[W(2^{-2j}\xi)\, v\Big(2^j(\frac{\xi_2}{\xi_1} - 1)\Big)\, v\Big(2^j \frac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-2^j,-\ell_2]} k}]|_{\xi_3=\xi_2=\xi_1} \\ &= 2^{-2j} \frac{\partial W}{\partial \xi_1}(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v(0)\, v(2^j - \ell_2)\, e^{2\pi i 2^{-2j-2}\xi_1 k_1} \\ &- \frac{2^j}{\xi_1} W(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v'(0)\, v(2^j - \ell_2)\, e^{2\pi i 2^{-2j-2}\xi_1 k_1} \\ &- \frac{2^j}{\xi_1} W(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v(0)\, v'(2^j - \ell_2)\, e^{2\pi i 2^{-2j-2}\xi_1 k_1} \\ &+ 2\pi i(2^{-2j-2}k_1 - 2^{-j-2}k_2 - 2^{-2j-2}\ell_2 k_3)\, W(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v(0)\, v(2^j - \ell_2) \\ &\times e^{2\pi i(2^{-2j-2}\xi_1 k_1 + 2^{-j-2}(\xi_2-\xi_1)k_2 + 2^{-j-2}(\xi_3 - \ell_2 2^{-j}\xi_1)k_3)}; \end{aligned}$$

$$
\begin{aligned}
&\frac{\partial}{\partial \xi_1}[W(2^{-2j}\xi)\, v\Big(2^j(\frac{\xi_1}{\xi_2}-1)\Big)\, v\Big(2^j\frac{\xi_3}{\xi_2}-\ell_2\Big)\, e^{2\pi i \xi 2^{-2} A_{(1)}^{-j} B_{(1)}^{[-2^j,-\ell_2]}k}]|_{\xi_3=\xi_2=\xi_1} \\
&= 2^{-2j}\frac{\partial W}{\partial \xi_1}(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v(0)\, v(2^j-\ell_2)\, e^{2\pi i 2^{-2j-2}\xi_1 k_1} \\
&+\frac{2^j}{\xi_1}\, W(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v'(0)\, v(2^j-\ell_2)\, e^{2\pi i 2^{-2j-1}\xi_1 k_1} \\
&+2\pi i(2^{-2j-2}k_1-2^{-j-2}k_2-2^{-2j-2}\ell_2 k_3)\, W(2^{-2j}(\xi_1,\xi_1,\xi_1))\, v(0)\, v(2^j-\ell_2) \\
&\times e^{2\pi i(2^{-2j-2}\xi_1 k_1+2^{-j-2}(\xi_2-\xi_1)k_2+2^{-j-2}(\xi_3-\ell_2 2^{-j}\xi_1)k_3)}.
\end{aligned}
$$

Since $v'(0)=0$ and $v'(2^j-\ell_2)=0$ (this is due to the fact that $v'(0)=0$ and v' vanishes outside its support in $(-1,1)$), the two partial derivatives agree for $\xi_1=\xi_2=\xi_3$. A very similar calculation shows that also the partial derivatives with respect to ξ_2 and ξ_3 agree for $\xi_1=\xi_2=\xi_3$. This observation can be repeated for higher order derivatives since $v^{(n)}(0)=0$ for all $n\geq 1$, implying that the functions $(\widetilde{\psi}_{j,\ell_1,\ell_2,k})^\wedge(\xi)$, given by (2.12), are infinitely differentiable. A similar computation shows that all boundary shearlets are infinitely differentiable. Notice that this idea for constructing regular boundary shearlets by matching the shearlet elements from different pyramidal regions is similar to an idea [1] (even though no details are provided there).

For brevity, in the following it will be convenient to denote the system of shearlets (2.11) using the compact notation:

$$\{\widetilde{\psi}_\mu,\ \mu\in M\}, \tag{2.13}$$

where $M=M_C\cup M_I\cup M_B$ are the indices associated with the coarse-scale shearlets, the interior shearlets, and the boundary shearlets, respectively, given by

- $M_C=\{\mu=(j,k): j=-1, k\in\mathbb{Z}^3\}$ (coarse-scale shearlets)
- $M_I=\{\mu=(j,\ell_1,\ell_2,k,d):\ j\geq 0, |\ell_1|\&|\ell_2|<2^j, k\in\mathbb{Z}^3, d=1,2,3\}$ (interior shearlets)
- $M_B=\{\mu=(j,\ell_1,\ell_2,k,d):\ j\geq 0, |\ell_1|<2^j, \ell_2=\pm 2^j, k\in\mathbb{Z}^3, d=1,2,3\}\cup\{\mu=(j,\ell_1,\ell_2,k):\ j\geq 0, \ell_1,\ell_2=\pm 2^j, k\in\mathbb{Z}^3\}$ (boundary shearlets)

We have the following result, whose proof is found in [17].

Theorem 2.1. *The 3D shearlet system* (2.13) *is a Parseval frame for* $L^2(\mathbb{R}^3)$. *In addition, the elements of this systems are* C^∞ *and compactly supported in the Fourier domain.*

2.2. Significance. Before presenting the proof of the main sparsity result, a simple heuristic argument can be used to explain why the 3D Parseval frame of shearlets constructed above is expected to be particularly effective in providing very sparse representations for functions of 3 variables with discontinuous boundaries. In fact, let us consider a bounded function f, defined on a bounded domain, which is smooth away from a discontinuity along a smooth surface. We will examine the behavior of the *shearlet coefficients* of f, which are given by $s_\mu(f)=\langle f,\widetilde{\psi}_\mu\rangle$, where the shearlet elements $\widetilde{\psi}_\mu$ are given by (2.13). For simplicity, we will only consider the interior shearlets $\psi^{(d)}_{j,\ell,k}$, given by (2.6). The first observation is that, thanks to their localization properties, at scale 2^{-2j}, the elements $\psi^{(d)}_{j,\ell,k}$, are essentially supported on a parallelepiped of size $2^{-2j}\times 2^{-j}\times 2^{-j}$, with locations controlled by k, and orientations controlled by ℓ. Also, using (2.9) it follows that

$$\int_{\mathbb{R}^3}|\psi^{(d)}_{j,\ell,k}(x)|\,dx = 2^{2j}\int|\psi^{(d)}_{j,\ell}(B^{[\ell]}_{(d)}A^j_{(d)}x-k)|\,dx = 2^{-2j}\int_{\mathbb{R}^3}|\psi^{(d)}_{j,\ell}(y)|\,dy,$$

so that, at scale 2^{-2j}, all these shearlet coefficients are controlled by

$$|s_{j,\ell,k}(f)|\leq\|f\|_\infty\,\|\psi^{(d)}_{j,\ell}\|_{L^1}\leq C\,2^{-2j}. \tag{2.14}$$

At sufficiently fine scales (for j sufficiently large), it is reasonable to assume that the only significant coefficients are those corresponding to the shearlet elements which are tangent to the surface of discontinuity. Since there are $O(2^{2j})$ coefficients of this type and they are bounded by (2.14), it follows that the N-th

largest shearlet coefficient in magnitude, denoted by $|s(f)|_{|N|}$ is bounded by $O(N^{-1})$. This implies that, if f is approximated by taking the N largest coefficients in the shearlets expansion, the L^2–error approximately obeys the estimate:

$$\|f - f_N\|_{L^2}^2 \leq \sum_{\ell > N} |s(f)|_{|\ell|}^2 \leq C\,N^{-1}. \tag{2.15}$$

A rigorous analysis of the behavior of the shearlet coefficients is the main goal of this paper and will be presented below. This analysis requires a careful examinations of the terms which were neglected in our heuristic argument and, as the detailed calculations below will show, this produces an additional logarithmic factor, finally yielding estimate (1.2).

3. Main Results. Before stating our main results, let us define the class of functions that will be considered in this paper. Fix a constant $A > 0$. We will consider a class $\mathcal{M}(A)$ of indicator functions of sets $B \subset [0,1]^3$ whose boundary $\Sigma = \partial B$ is a C^2 2-manifold which can be written as $\bigcup_\alpha \Sigma_\alpha$, where α ranges over a finite index set and $\Sigma_\alpha = \{(v,\ E_\alpha(v)),\ \ v \in V_\alpha \subset \mathbb{R}^2\}$, such that $\|E_\alpha\|_{C^2(V_\alpha)} \leq A$ for all α. Also, let $C_c^2([0,1]^3)$ be the collection of twice differentiable functions supported inside $[0,1]^3$. Hence, we define the set $\mathcal{E}^2(A)$ of *functions which are C^2 away from a C^2 surface* as the collection of functions of the form

$$f = f_0 + f_1\,\chi_B,$$

where $f_0, f_1 \in C_c^2([0,1]^3)$, $B \in \mathcal{M}(A)$ and $\|f\|_{C^2} = \sum_{|\alpha| \leq 2} \|D^\alpha f\|_\infty \leq 1$.

Let $\{\widetilde{\psi}_\mu : \mu \in M\}$ denote the Parseval frame of shearlets for $L^2(\mathbb{R}^3)$ given by (2.13). The *shearlet coefficients* of a function f are the elements of the sequence $\{s_\mu(f) = \langle f, \widetilde{\psi}_\mu \rangle : \mu \in M\}$. We denote by $|s(f)|_{(N)}$ the N-th largest entry, in magnitude, in this sequence. We can now state the main results of this paper.

THEOREM 3.1. *Let $f \in \mathcal{E}^2(A)$ and $\{s_\mu(f) = \langle f, \widetilde{\psi}_\mu \rangle : \mu \in M\}$ be the sequence of shearlet coefficients associated with f. Then*

$$\sup_{f \in \mathcal{E}^2(A)} |s_\mu(f)|_{(N)} \leq C\,N^{-1}\,(\log N). \tag{3.1}$$

Let us comment on the significance of this result. It shows that, up to the loglike factor, the shearlet representation provides the *optimal degree of sparsity* for functions in $\mathcal{E}^2(A)$. In fact, as discussed in more detail in Sec. 5.1 (extending a classical 2D result by Donoho in [9]), there is no representation

$$f = \sum c_i(f)\,\phi_i$$

satisfying polynomial depth search that can provide approximations for $f \in \mathcal{E}^2(A)$ where the coefficients $(c_i(f))$ are bounded in weak ℓ^p norm for $p < 1$. That is, the rate N^{-1} in (3.1) is the optimal that can be achieved using not only orthogonal bases or frames but even considering larger dictionary, as long as they satisfy a polynomial depth search condition.

Using Theorem 3.1, we are just one step away from our other main result about shearlet approximations. Indeed, let f_N^S be the N–term approximation of f obtained from the N largest coefficients of its shearlet expansion, namely

$$f_N^S = \sum_{\mu \in I_N} \langle f, \widetilde{\psi}_\mu \rangle\, \widetilde{\psi}_\mu,$$

where $I_N \subset M$ is the set of indices corresponding to the N largest entries of the sequence $\{|s_\mu(f)| = |\langle f, \widetilde{\psi}_\mu \rangle|^2 : \mu \in M\}$. The approximation error satisfies the estimate:

$$\|f - f_N^S\|_2^2 \leq \sum_{m > N} |s(f)|_{(m)}^2.$$

Therefore, from (3.1) we immediately have:

THEOREM 3.2. *Let $f \in \mathcal{E}^2(A)$ and f_N^S be the approximation to f defined above. Then*

$$\|f - f_N^S\|_2^2 \le C\, N^{-1}\, (\log N)^2.$$

This result extends to the 3D setting the essentially optimal approximation result given by (1.1).

3.1. Arguments and constructions. The general structure of the proof of Theorem 3.1 follows the overall structure of the corresponding 2-dimensional sparsity result in [14]. However, as it will be clear below, the core of the proof requires a fundamentally new approach which is significantly different from the 2D case.

As in [14], it will be convenient to introduce the *weak–ℓ^p quasi–norm* $\|\cdot\|_{w\ell^p}$ to measure the sparsity of the shearlet coefficients (cf. [7] for an overview of the weak–ℓ^p spaces). For a sequence (s_μ), this is defined by

$$\|s_\mu\|_{w\ell^p} = \sup_{N>0} N^{\frac{1}{p}}\, |s_\mu|_{(N)},$$

wheret $|s_\mu|_{(N)}$ is the N-th largest entry in the sequence $\{s_\mu\}$. One can show (cf. [33, Sec.5.3]) that this definition is equivalent to

$$\|s_\mu\|_{w\ell^p} = \left(\sup_{\epsilon>0} \#\{\mu : \, |s_\mu| > \epsilon\}\, \epsilon^p \right)^{\frac{1}{p}}.$$

To analyze the decay properties of the shearlet coefficients $\{\langle f, \widetilde{\psi}_\mu\rangle \colon \mu \in M\}$ at a given scale 2^{-j}, $j \ge 0$, we will smoothly localize the function f near dyadic cubes. Namely, for a scale parameter $j \ge 0$ fixed, let $\mathcal{Q}_j$ be the collection of dyadic cubes of the form $Q = [\frac{k_1}{2^j}, \frac{k_1+1}{2^j}] \times [\frac{k_2}{2^j}, \frac{k_2+1}{2^j}] \times [\frac{k_3}{2^j}, \frac{k_3+1}{2^j}]$, with $k_1, k_2, k_3 \in \mathbb{Z}$. For w a nonnegative C^∞ function with support in $[-1,1]^3$, we define a smooth partition of unity

$$\sum_{Q \in \mathcal{Q}_j} w_Q(x) = 1, \quad x \in \mathbb{R}^3,$$

where, for each dyadic square $Q \in \mathcal{Q}_j$, $w_Q(x) = w(2^j x_1 - k_1, 2^j x_2 - k_2, 2^j x_3 - k_3)$. We will then examine the shearlet coefficients of the localized function $f_Q = f\, w_Q$, i.e., $\{\langle f_Q, \widetilde{\psi}_\mu\rangle : \mu \in M_j\}$, where M_j denotes the collection of the $\mu \in M$ such that j is fixed (for example, in the case of the indices associated with the interior shearlets, $M_j = \{(j, \ell_1, \ell_2, k, d) : -2^j \le \ell_1, \ell_2 \le 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$).

As it will be shown below, for $f \in \mathcal{E}^2(A)$, the shearlet coefficients $\{\langle f_Q, \widetilde{\psi}_\mu\rangle : \mu \in M_j\}$ exhibit a different decay behavior depending on whether the surface intersects the support of w_Q or not. Let $\mathcal{Q}_j = \mathcal{Q}_j^0 \cup \mathcal{Q}_j^1$, where the union is disjoint and $\mathcal{Q}_j^0$ is the collection of those dyadic cubes $Q \in \mathcal{Q}_j$ such that the surface intersects the support of w_Q. Since each Q has sidelength $2 \cdot 2^{-j}$, then $\mathcal{Q}_j^0$ has cardinality $|\mathcal{Q}_j^0| \le C_0\, 2^{2j}$, where C_0 is independent of j. Similarly, since f is compactly supported in $[0,1]^3$, $|\mathcal{Q}_j^1| \le 2^{3j} + 6 \cdot 2^{2j}$.

Using this notation, we can now state the basic results that are needed to prove Theorem 3.1. For simplicity, in the following, we will use the same letter C to denote different uniform constants.

THEOREM 3.3. *Let $f \in \mathcal{E}^2(A)$. For $Q \in \mathcal{Q}_j^0$, with $j \ge 0$ fixed, the sequence of shearlet coefficients $\{\langle f_Q, \widetilde{\psi}_m u\rangle : \mu \in M_j\}$ obeys*

$$\|\langle f_Q, \widetilde{\psi}_\mu\rangle\|_{w\ell^1} \le C\, 2^{-2j},$$

for some constant C independent of Q and j.

THEOREM 3.4. *Let $f \in \mathcal{E}^2(A)$. For $Q \in \mathcal{Q}_j^1$, with $j \ge 0$ fixed, the sequence of shearlet coefficients $\{\langle f_Q, \widetilde{\psi}_\mu\rangle : \mu \in M_j\}$ obeys*

$$\|\langle f_Q, \widetilde{\psi}_\mu\rangle\|_{\ell^1} \le C\, 2^{-4j},$$

for some constant C independent of Q and j.

The proofs of Theorems 3.3 and 3.4 are rather involved. Theorems 3.3, in particular, is the "hardest" part of the new sparsity result, and is the result whose argument is most different with respect to the 2D case. Concerning Theorem 3.4, it also shows that 3D shearlets are as effective as traditional isotropic wavelets in dealing with smooth functions. [2] Before presenting the proofs of Theorems 3.3 and 3.4, we show how these two theorems are used to prove Theorem 3.1. Indeed, we have the following simple corollary.

Corollary 3.5. *Let $f \in \mathcal{E}^2(A)$ and, for $j \geq 0$, $s_j(f)$ be the sequence $s_j(f) = \{\langle f, \widetilde{\psi}_\mu \rangle \colon \mu \in M_j\}$. Then there is a constant C independent of j such that:*

$$\|s_j(f)\|_{w\ell^1} \leq C.$$

Proof. Using Theorems 3.3 and 3.4, by the triangle inequality for weak ℓ^1 spaces, we have

$$\begin{aligned}
\|s_j(f)\|_{w\ell^1} &\leq \sum_{Q \in \mathcal{Q}_j} \|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{w\ell^1} \\
&\leq \sum_{Q \in \mathcal{Q}_j^0} \|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{w\ell^1} + \sum_{Q \in \mathcal{Q}_j^1} \|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{\ell^1} \\
&\leq C\, |\mathcal{Q}_j^0|\, 2^{-2j} + C\, |\mathcal{Q}_j^1|\, 2^{-4j} \\
&\leq C(2^{2j}\, 2^{-2j} + 2^{3j}\, 2^{-4j}) \leq C.
\end{aligned}$$

Here we used the facts that $|\mathcal{Q}_j^0| \leq C\, 2^{2j}$, where C is independent of j, and $|\mathcal{Q}_j^1| \leq 2^{3j} + 6 \cdot 2^{2j}$.

We can now prove Theorem 3.1

Proof of Theorem 3.1. By Corollary 3.5, we have that

$$R(j, \epsilon) = \#\{\mu \in M_j : |\langle f, \widetilde{\psi}_\mu \rangle| > \epsilon\} \leq C\, \epsilon^{-1}. \tag{3.2}$$

Next observe that, for an interior shearlet $\psi^{(d)}_{j,\ell,k}$, given by (2.6), using (2.9) and (2.10), a direct calculation gives

$$\begin{aligned}
|\langle f, \psi^{(d)}_{j,\ell,k} \rangle| &= \left| \int_{\mathbb{R}^3} f(x)\, 2^{2j}\, \psi^{(d)}(B^{[\ell]}_{(d)} A^j_{(d)} x - k)\, dx \right| \\
&\leq 2^{2j}\, \|f\|_\infty \int_{\mathbb{R}^3} |\psi^{(d)}_{j,\ell}(B^{[\ell]}_{(d)} A^j_{(d)} x - k)|\, dx \\
&= 2^{-2j}\, \|f\|_\infty \int_{\mathbb{R}^3} |\psi^{(d)}_{j,\ell}(y)|\, dy < C'\, 2^{-2j}.
\end{aligned} \tag{3.3}$$

A very similar computation holds for the boundary shearlets. As a consequence, there is a scale j_ϵ such that $|\langle f, \widetilde{\psi}_\mu \rangle| < \epsilon$ for each $j \geq j_\epsilon$. Specifically, it follows from (3.3) that $R(j, \epsilon) = 0$ for $j > 2\,(\log_2(\epsilon^{-1}) + \log_2(C')) > 2\, \log_2(\epsilon^{-1})$. Thus, using (3.2), we have that

$$\#\{\mu \in M : |\langle f, \widetilde{\psi}_\mu \rangle| > \epsilon\} \leq \sum_{j \geq 0} R(j, \epsilon) = \sum_{j=0}^{2 \log_2(\epsilon^{-1})} R(j, \epsilon) \leq C\, \epsilon^{-1}\, \log_2(\epsilon^{-1}),$$

and this implies (3.1). □

4. Proofs of Main Theorems.

[2] Furthermore, an argument similar to Theorem 8.2 in [2] can be used to analyze the estimate the Sobolev norm of a smooth function using shearlet coefficients.

4.1. Proof of Theorem 3.3. Let us consider a function $f \in \mathcal{E}^2(A)$ which contains a C^2 surface of discontinuity. For $j > j_0$ sufficiently large, the scale 2^{-j} is small enough, so that, over a cube of side 2^{-j}, the surface of discontinuity can be parametrized as $x_1 = E(x_2, x_3)$ or $x_2 = E(x_1, x_3)$ or $x_3 = E(x_1, x_2)$. For simplicity, we will assume that this surface, denoted by Σ, satisfies the equation

$$x_1 = E(x_2, x_3), \quad -2^{-j} \le x_2, x_3 \le 2^{-j}.$$

Also we assume that the surface contains the origin $(0,0,0)$ and the normal direction of the surface at $(0,0,0)$ is $(1,0,0)$, which is equivalent to assuming that $E(0,0) = E_{x_2}(0,0) = E_{x_3}(0,0) = 0$. As we will show in Section 4.5, there is no loss in generality in analyzing this case only, since the situation where the surface does not contain the origin or has a different normal direction can be easily converted into the case where $E(0,0) = E_{x_2}(0,0) = E_{x_3}(0,0) = 0$. To further simplify the notation, throughout the remainder of the paper, for a function $g(x)$ with $x \in \mathbb{R}^2$ and $m = (m_1, m_2)$ with $0 \le |m| = m_1 + m_2 \le 2$, we will write $\frac{\partial^m}{\partial x^m} g$ as g_m.

From Taylor's Theorem we have that $E(x_2, x_3) = \frac{1}{2}(E_{(2,0)}(c)x_2^2 + 2E_{(1,1)}(c)x_2x_3 + E_{(0,2)}(c)x_3^2)$, where $c = (c_2, c_3)$ is some point in $[-2^{-j}, 2^{-j}]^2$. It follows that

$$|E(x_2, x_3)| \le 2^{-2j}(\|E_{(2,0)}\|_\infty + \|E_{(1,1)}\|_\infty + \|E_{(0,2)}\|_\infty).$$

Thus, the surface is locally nearly flat near the origin. Notice that this only holds for $j > j_0$. The situation when $j \le j_0$ is much simpler and will be handled separately in Section 4.6.

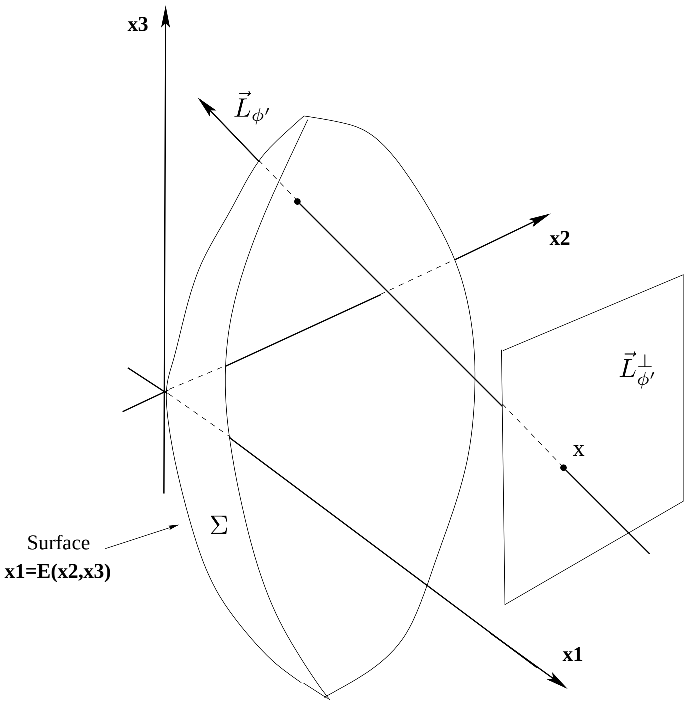

FIG. 4.1. *The surface of discontinuity Σ of equation $x_1 = E(x_2, x_3)$. A line with direction $\vec{L}_{\phi'}$ through the point x intersects the surface at most in one point.*

The key step in the following argument is based on the estimate of the decay of the function f near the surface of discontinuity. In order to define this localized version of f, let w_0 be a nonnegative C^∞ window function with support in $[-1,1]^3$. Hence, for $j \in \mathbb{Z}$, a *surface fragment* is a function of the form:

$$f(x) = w_0(2^j x)\, g(x)\, \chi_{[x_1 > E(x_2, x_3)]}(x), \quad x \in [-2^{-j}, 2^{-j}]^3, \tag{4.1}$$

where $g \in C_0^2((-1,1)^3)$. After re-scaling, we have

$$F(x) = f(2^{-j}x) = w_0(x)\, g(2^{-j}x)\, \chi_{[x_1 > E^{(j)}(x_2,x_3)]}(x), \quad x \in [-1,1]^3, \tag{4.2}$$

where $E^{(j)}(x_2,x_3) = 2^j\, E(2^{-j}x_2, 2^{-j}x_3)$. In particular, we have that $\widehat{F}(\xi) = 2^{3j}\, \hat{f}(2^j\xi)$, and, thus, writing $\xi \in \mathbb{R}^3$ in spherical coordinates as $\lambda\Theta$, where $\lambda \geq 0$, $\Theta \in S^2$, we have that

$$\int_{\lambda \in [a,b]} |\hat{f}(\lambda\Theta)|^2\, d\lambda = 2^{-5j} \int_{\lambda \in 2^{-j}[a,b]} |\hat{F}(\lambda\Theta)|^2\, d\lambda. \tag{4.3}$$

For simplicity of notation, without loss of generality we may assume that $(\|E_{(2,0)}\|_\infty + \|E_{(1,1)}\|_\infty + \|E_{(0,2)}\|_\infty) = 1$, which yields that $|E(x_2,x_3)| \leq 2^{-2j}$ and $|E_m(x_2,x_3)| \leq 2^{-j}$ for $|m| \leq 2$ for all $(x_2,x_3) \in [-1,1]$.

4.2. Analysis of the Surface Fragment. The main goal of this section is to obtain an L^2 estimate for the elements of the Parseval frame of shearlets against the surface fragment (4.2). For this, it will be sufficient to consider the interior shearlets (2.6) associated with the pyramidal region $\mathcal{P}_1$. In fact, the boundary shearlets satisfy similar support and regularity conditions, except for the fact that they are piecewise defined, so that the estimates involving these functions against the surface fragment can be handled in the same way. It is also clear that the analysis for the shearlets associated with the other pyramidal regions $\mathcal{P}_2$ and $\mathcal{P}_3$ can be handled in exactly the same way.

In the following, it will be convenient to express $\xi \in \mathbb{R}^3$ using spherical coordinates, so that we will write $\xi = (\rho\cos\theta\sin\phi, \rho\cos\theta\sin\phi, \rho\cos\phi)$, where $\rho > 0$, $\theta \in [0,2\pi)$ and $\phi \in [0,\pi]$. Since we are only dealing with the frequency region contained in $\mathcal{P}_1$, we will assume that $\phi \in [\frac{\pi}{4}, \frac{3\pi}{4}]$ and $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$. Also notice that, since the variables ξ_2, ξ_3 are symmetric in the construction of the shearlets in $\mathcal{P}_1$, we may assume that $|\ell_1| \leq |\ell_2|$.

For $\xi = (\xi_1,\xi_2,\xi_3) \in \mathcal{P}_1$, $j \geq 0$, $|\ell_1| \leq |\ell_2| \leq 2^j$, let

$$\Gamma_{j,\ell}(\xi) = W\big(2^{-2j}\,\xi\big)\; v\Big(2^j\,\frac{\xi_2}{\xi_1} - \ell_1\Big)\; v\Big(2^j\,\frac{\xi_3}{\xi_1} - \ell_2\Big). \tag{4.4}$$

Using this notation, the interior shearlets (2.8) associated with the pyramidal region $\mathcal{P}_1$ can be written as

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

We have the following useful result:

Theorem 4.1. *Let f be the surface fragment given by expression (4.1). Then, for each $j \geq 0$, and $|\ell_1|, |\ell_2| \leq 2^j$, the following estimate holds:*

$$\int_{\widehat{\mathbb{R}}^3} |\hat{f}(\xi)|^2\, |\Gamma_{j,\ell}(\xi)|^2\, d\xi \leq C\, 2^{-4j} (1+|\ell_2|)^{-5}. \tag{4.5}$$

The proof of these results is based on the computation of the ray transform of the surface fragment f which is presented below.

4.3. Ray Transform And Fourier Slice Theorem. While the Radon and ray transforms of bivariate functions are equivalent, this is not true in the three-dimensional setting [28]. Namely, the 3-dimensional *ray transform* maps a function on $\mathbb{R}^3$ into the sets of its line integrals; this is different from the *Radon transform* which maps a function on $\mathbb{R}^3$ into the sets of its integrals over planes in $\mathbb{R}^3$. More precisely, if $\Theta \in S^2$ and $x \in \mathbb{R}^3$, then the ray transform of $g \in \mathcal{S}(\mathbb{R}^3)$ is defined by

$$Pg(\Theta, x) = \int_{\mathbb{R}} g(t\Theta + x)\, dt.$$

This is the integral of g over the straight line through x with direction Θ (see Figure 4.2). Notice that $Pg(\Theta, x)$ does not change if x is moved in the direction Θ. Hence, x is normally restricted to $\Theta^\perp$ so that Pf

is a function on the tangent bundle $\{(\Theta, x) : \Theta \in S^2, x \in \Theta^\perp\}$. It is useful to recall the *Fourier Slice Theorem* which establishes that following relationship between the ray transform of g and its Fourier transform:

$$\mathcal{F}_2[Pg](\Theta, \eta) = \int_{\Theta^\perp} Pg(\Theta, x)\, e^{-2\pi i \eta x}\, dx = \hat{g}(\eta), \quad \eta \in \Theta^\perp,$$

where $\mathcal{F}_2$ denotes the Fourier transform over the second variable. We refer the reader to [28] for this and additional properties of the ray transform.

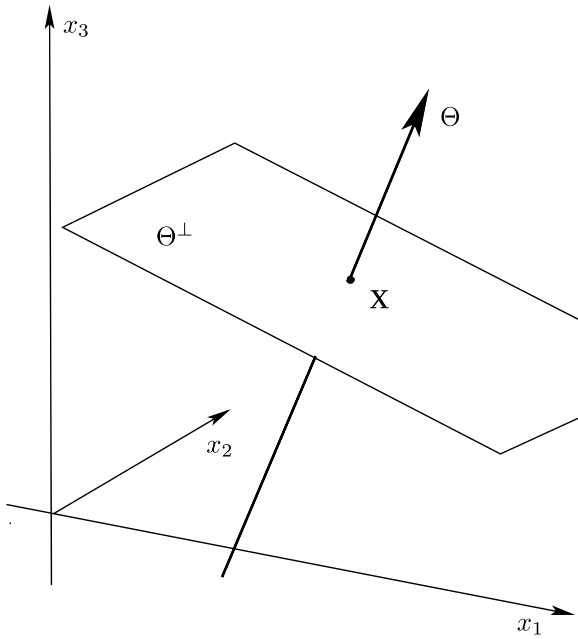

FIG. 4.2. *The ray transform is defined by integration over the lines through the point x with direction Θ.*

In order to deduce an estimate for the integral of the surface fragment given by the expression (4.3), we will analyze the ray transform of the surface fragment F, given by (4.2). Let $\phi' \in [-\frac{\pi}{4}, \frac{\pi}{4}]$. The ray transform of F in the direction $\vec{L}_{\phi'} = (\sin\phi', 0, \cos\phi')$ is given by

$$PF(\phi', x) = \int_{\mathbb{R}} F(tL_{\phi'} + x)\, dt \tag{4.6}$$

where $x \in \mathbb{R}^3$. This is the integral of F over the straight line through x with direction $L_{\phi'}$. Notice that $PF(\phi', x)$ does not change if x moves along the direction $\vec{L}_{\phi'}$. Hence, x is effectively restricted to $\vec{L}_{\phi'}^\perp$ so that PF is a function on the tangent bundle $\{(\vec{L}_{\phi'}, x) : \vec{L}_{\phi'} \in S^2, x \in \vec{L}_{\phi'}^\perp\}$. By introducing the vectors $\vec{L}_1 = (0, -1, 0)$ and $\vec{L}_2 = (\cos\phi', 0, -\sin\phi')$, we can express $x \in L_{\phi'}^\perp$ as

$$\{x \in \vec{L}_{\phi'}^\perp\} = \{s\vec{L}_1 + w\vec{L}_2 : \; s, w \in \mathbb{R}\}. \tag{4.7}$$

It follows that

$$PF(\phi', s, w) = \int_{\mathbb{R}} F\left(\rho_{\phi'} \begin{pmatrix} t \\ s \\ w \end{pmatrix}\right) dt, \tag{4.8}$$

where $\rho_{\phi'} = \begin{pmatrix} \sin\phi' & 0 & \cos\phi' \\ 0 & -1 & 0 \\ \cos\phi' & 0 & -\sin\phi' \end{pmatrix}$.

By the Fourier Slice Theorem, we have that

$$\mathcal{F}_2[PF](\phi',\eta) = \int_{\vec{L}^{\perp}_{\phi'}} PF(\phi',s,w)\, e^{-2\pi i \eta\cdot(s,w)}\, ds\, dw = \hat{F}(\eta,\phi'), \quad \eta \in \vec{L}^{\perp}_{\phi'}.$$

Hence, by the properties of the Fourier transform (Plancherel and differentiation theorems), we obtain the following identity:

$$\|(PF)_{ss}\|^2 + 2\|(PF)_{sw}\|^2 + \|(PF)_{ww}\|^2 = (2\pi)^4 \int_{\mathbb{R}^2} |\eta|^4\, |\hat{F}(\eta,\phi')|^2\, d\eta, \tag{4.9}$$

where $\eta = \eta_1\, \vec{L}_1 + \eta_2\, \vec{L}_2$.

4.3.1. Ray Transform of the Surface Fragment. For brevity, let us introduce the following notation:

$$F^{\phi'}(t,s,w) = F\left(\rho_{\phi'}\begin{pmatrix} t \\ s \\ w\end{pmatrix}\right), \quad g^{\phi'}(t,s,w) = g\left(2^{-j}\rho_{\phi'}\begin{pmatrix} t \\ s \\ w\end{pmatrix}\right), \quad w^{\phi'}(t,s,w) = w\left(\rho_{\phi'}\begin{pmatrix} t \\ s \\ w\end{pmatrix}\right).$$

Using this notation, we will rewrite the ray transform of the surface fragment, given by expression (4.8), as

$$PF(\phi',s,w) = \int_{\mathbb{R}} F^{\phi'}(t,s,w)\, dt. \tag{4.10}$$

As described above, this is an integral over the lines $\Lambda_{s,w,\phi'} = \{y \in \mathbb{R}^3 : \, y \cdot \vec{L}_1 = s \quad \& \quad y \cdot \vec{L}_2 = w\}$, where $\vec{L}_1$ and $\vec{L}_2$, given by (4.7), depend on ϕ'. Depending on the values of (s, w, ϕ'), the lines $\Lambda_{s,w,\phi'}$ may or may not intersect the surface $\Sigma = \{(E^{(j)}(u,v), u, v) : \, |u|, |v| \le 1\}$. In the following, we will analyze the two situations separately.

Case 1: No Intersection.

When the line $\Lambda_{s,w,\phi'}$ does not intersect the surface Σ, the ray transform of F takes the form:

$$PF(\phi',s,w) = \int_{\mathbb{R}} g^{\phi'}(t,s,w)\, w^{\phi'}(t,s,w)\, dt. \tag{4.11}$$

In this case we have the following result.

PROPOSITION 4.2. *The function PF is twice differentiable as a function of s and w and admits the decomposition*

$$(PF(\phi',s,w))_{ss}(\phi',s,w) + (PF(\phi',s,w))_{sw}(\phi',s,w) + (PF(\phi',s,w))_{ww}(\phi',s,w) = F^0(\phi',s,w) + F^1(\phi',s,w),$$

where

$$\|F^0(\phi',s,w)\|^2 \le C\, 2^{-2j},$$

$$\|\left(F^1(\phi',s,w)\right)_s\|^2 + \|\left(F^1(\phi',s,w)\right)_w\|^2 \le C.$$

Proof. With an abuse of notation, in the following we will write g for $g^{\phi'}$ and w_0 for w_0^{ϕ}. By direct computation we have:

$$(PF)_{ss}(\phi,s,w) = \int_{\mathbb{R}} \frac{\partial^2}{\partial s^2}\left(g(t,s,w)w(t,s,w)\right)\, dt = F^0(\phi',s,w) + F^1(\phi',s,w),$$

where $F^0(\phi', s, w) = \int_{\mathbb{R}} (g_{ss}\, w_0 + 2g_s\, w_{0s})\, dt$ and $F^1(\phi', s, w) = \int_{\mathbb{R}} g\, w_{0ss}\, dt$.

Recalling that $g(t, s, w) = g^{\phi'}(t, s, w) = g\left(2^{-j} \rho_{\phi'} \left(\begin{smallmatrix} t \\ s \\ w \end{smallmatrix}\right)\right)$, a direct computation yields that $|g_s| \le C\, 2^{-j}$ and $|g_{ss}| \le C\, 2^{-2j}$. It follows that $|g_s w_{0s}| \le C\, 2^{-j}$ and $|g_{ss} w_0| \le C\, 2^{-2j}$. Since w_0 (and hence PF) has compact support, it follows that $\int_{\mathbb{R}} |g_{ss}\, w_0| dt \le C\, 2^{-2j}$, and $\int_{\mathbb{R}} |g_s\, w_{0s}| dt \le C\, 2^{-j}$. This implies that

$$\|F^0(\phi', s, w)\|^2 \le C\, 2^{-2j}.$$

For $F^1(\phi', s, w)$, we have

$$\frac{\partial}{\partial s}(F^1(\phi', s, w)) = \int_{\mathbb{R}} \frac{\partial}{\partial s}(g\, w_{0ss})\, dt = \int_{\mathbb{R}} (g_s\, w_{ss} + g\, w_{sss})\, dt.$$

Using the same argument as the one used for $F^0(\phi', s, w)$, it follows that $\|\big(F^1(\phi', s, w)\big)_s\|^2 \le C$. Similarly it follows that $\|\big(F^1(\phi', s, w)\big)_w\|^2 \le C$. The proof is completed by repeating the same argument for $(PF)_{sw}(\phi, s, w)$ and $(PF)_{ww}(\phi, s, w)$.

From Proposition 4.2, using the Fourier Slice Theorem for the ray transform and the Plancherel theorem, it follows that

$$\int_0^\infty \int_0^{2\pi} |\hat{F}(r, \theta', \phi')|^2\, r^5\, d\theta'\, dr \le C\, 2^{-2j}$$

and, hence, that

$$\int_{2^{j-2}}^{2^{j+1}} \int_0^{2\pi} |\hat{F}(r, \theta', \phi')|^2\, d\theta'\, dr \le C\, 2^{-7j}.$$

Since $F(x) = f(2^{-j}x)$, we have $\hat{F}(\xi) = 2^{3j} \hat{f}(2^j \xi)$. Thus, the above inequality implies the following one:

$$\int_{2^{2j-4}}^{2^{2j+2}} \int_0^{2\pi} |\hat{f}(r, \theta', \phi')|^2\, d\theta'\, dr \le C\, 2^{-12j}. \tag{4.12}$$

This completes the analysis in the case where there is no intersection.

Case 2: Intersection.

In order to find the intersection of the line $\Lambda_{s,w,\phi'}$ and the surface Σ, one has to solve the equation

$$\rho'_\phi \begin{pmatrix} t \\ s \\ w \end{pmatrix} = \begin{pmatrix} E^{(j)}(u) \\ u \\ v \end{pmatrix},$$

which leads to the system:

$$t = E^{(j)}(u, v) \sin\phi' + v \cos\phi', \tag{4.13}$$
$$s = -u, \tag{4.14}$$
$$w = E^{(j)}(u, v) \cos\phi' - v \sin\phi'. \tag{4.15}$$

To compute the solution of this system, we will use the Implicit Function Theorem to express t as a function of s and w. In order to do that, we first check that the conditions of the Implicit Function Theorem are satisfied. A direct computation gives:

$$s_u = -1, s_v = 0,$$
$$w_u = E^{(j)}_u(u, v) \cos\phi', w_v = E^{(j)}_v(u, v) \cos\phi' - \sin\phi',$$

and

$$\Delta(\phi') = \det \begin{pmatrix} s_u & s_v \\ w_u & w_v \end{pmatrix} = \sin\phi' - E_v^{(j)} \cos\phi' \tag{4.16}$$

The following proposition deals with the case when $|\sin\phi'| \le 2^{1-j}$.

PROPOSITION 4.3. *Assume that $|\sin\phi'| \le 2^{1-j}$. Then, for each fixed j and ϕ', we have that*

$$\int_{2^{2j-4}}^{2^{2j+2}} \int_0^{2\pi} |\hat{f}(r,\theta',\phi')|^2 \, d\theta' \, dr \le C\, 2^{-7j},$$

where where C is independent of j and ϕ'.

Proof. Since $|E_v^{(j)}| \le 2^{-j}$ (from the assumption that $\|E''\|_{L^\infty} = 1$), it follows that $|\Delta(\phi')| \le C 2^{-j}$ with C independent of j and ϕ'. Let A be the region defined by $\{(s(u,v), w(u,v)) : (u,v) \in [-1,1]^2\}$. Since $\int_A ds\, dw = \int_{-1}^1 \int_{-1}^1 |\Delta(\phi')| \, du\, dv \le C|\sin\phi'|$ and F is bounded (and hence PF is bounded), it follows from a direct calculation that $\|(PF)\|_{L^2}^2 \le C \int_{-1}^1 \int_{-1}^1 |\Delta(\phi')| \, du\, dv \le C 2^{-j}$. Using the Fourier Slice Theorem for the ray transform and the Plancherel theorem, we have that

$$\int_0^\infty \int_0^{2\pi} |\hat{F}(r,\theta',\phi')|^2 \, r \, d\theta' \, dr \le C\, 2^{-j}$$

and, hence, that

$$\int_{2^{j-2}}^{2^{j+1}} \int_0^{2\pi} |\hat{F}(r,\theta',\phi')|^2 \, d\theta' \, dr \le C\, 2^{-2j}.$$

Since $F(x) = f(2^{-j}x)$, we have $\hat{F}(\xi) = 2^{3j} \hat{f}(2^j \xi)$. Thus the above inequality gives

$$\int_{2^{2j-4}}^{2^{2j+2}} \int_0^{2\pi} |\hat{f}(r,\theta',\phi')|^2 \, d\theta' \, dr \le C\, 2^{-7j}.$$

This finishes the proof of Proposition 4.3.

For the case when $|\sin\phi'| \ge 2^{1-j}$, we have that $2^{-j} \le \frac{1}{2}|\sin\phi'| \le |\Delta(\phi')| \le 2|\sin\phi'|$. Thus, we can apply the Inverse Function Theorem and use equations (4.14) and (4.15) to derive the functions $u = u(s,w)$ and $v = v(s,w)$. Inserting these functions into (4.13), we obtain the intersection point in terms of t as

$$t_0(s,w,\phi') = E^{(j)}(u(s,w), v(s,w)) \sin\phi' + v(s,w)\cos\phi'. \tag{4.17}$$

This shows that there is at most one point of intersection for each fixed (s,w) and ϕ'.

We can write $\eta \in \vec{L}_{\phi'}^{\perp}$ as $\eta = (\eta_2 \cos\phi', -\eta_1, -\eta_2 \sin\phi') = (r\sin\theta' \cos\phi', -r\cos\theta', -r\sin\theta' \sin\phi')$, where $\eta_1 = r\cos\theta'$, $\eta_2 = r\sin\theta'$. Then (4.9) can be rewritten as

$$\|(PF)_{ss}\|^2 + 2\|(PF)_{sw}\|^2 + \|(PF)_{ww}\|^2 = \int_0^\infty \int_0^{2\pi} r^5 |\hat{F}(r,\theta',\phi')|^2 \, d\theta' \, dr. \tag{4.18}$$

Since the same η can also be expressed in spherical coordinates as $\eta = (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \cos\phi)$, it follows that we must have $\rho = r$ and

$$\begin{aligned} \sin\theta' \cos\phi' &= \cos\theta \sin\phi, \\ \cos\theta' &= \sin\theta \sin\phi, \\ -\sin\theta' \sin\phi' &= \cos\phi. \end{aligned}$$

From the first and the third identities, we have $\tan\phi' = -\cot\phi\sec\theta$, which implies that ϕ' is *equivalent* to $\phi - \frac{\pi}{2}$, that is, there is are constants $0 < C_1(\theta) \le C_2(\theta) < \infty$ such that $C_1(\theta)\,\phi' \le \phi - \frac{\pi}{2} \le C_2(\theta)\,\phi'$. Also since $|\phi - \frac{\pi}{2}| \le \frac{\pi}{4}$ and $|\theta| \le \frac{\pi}{4}$, we see that $|\frac{\partial\phi'}{\partial\phi}| \le C$ and $|\frac{\partial\phi'}{\partial\theta}| \le C$ and hence

$$|\phi_1' - \phi_2'| \le C\,(|\phi_1 - \phi_2| + |\theta_1 - \theta_2|). \tag{4.19}$$

Also, we have

$$u_s = \frac{E_v^{(j)}(u,v)\cos\phi - \sin\phi}{\Delta(\phi')}, u_w = 0, \tag{4.20}$$

$$v_s = -\frac{E_u^{(j)}(u,v)\cos\phi}{\Delta(\phi')}, v_w = -\frac{1}{\Delta(\phi')}. \tag{4.21}$$

From (4.20) and (4.21), it is easy to verify the following proposition.

Proposition 4.4.

$$|u_s| \le C\frac{1}{|\sin\phi'|},\ |u_{s^2}| \le C\frac{2^{-j}}{|\sin\phi'|^3},\ |u_{sw}| \le C\frac{2^{-j}}{|\sin\phi'|^3},\ |u_{w^2}| \le C\frac{2^{-j}}{|\sin\phi'|^3}$$
$$|v_s| \le C\frac{1}{|\sin\phi'|},\ |v_{s^2}| \le C\frac{2^{-j}}{|\sin\phi'|^3},\ |v_{sw}| \le C\frac{2^{-j}}{|\sin\phi'|^3},\ |v_{w^2}| \le C\frac{2^{-j}}{|\sin\phi'|^3},$$

where the constant C is independent of $(u,v) \in [-1,1]^2$, $\phi' \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ with $|\sin\phi'| \ge 2^{1-j}$.

Using the expression (4.17) that was found for the intersection point, from (4.6) and (4.8) we obtain the following formulation of the ray transform $PF(\phi', s, w)$:

$$PF(\phi', s, w) = \int_{-\infty}^{t_0(s,w,\phi')} F\left(\rho_{\phi'}\begin{pmatrix} t \\ s \\ w \end{pmatrix}\right) dt. \tag{4.22}$$

From Proposition 4.4, one can use essentially the same argument as the 2-dimensional case (see Lemma 6.2 in [2]) to prove the following proposition. For completeness, a sketch of its proof is provided below.

Proposition 4.5. *The ray transform of F is twice differentiable as a function of s and w and admits the decomposition*

$$(PF(\phi',s,w))_{ss}(\phi',s,w) + (PF(\phi',s,w))_{sw}(\phi',s,w) + (PF(\phi',s,w))_{ww}(\phi',s,w) = F^0(\phi',s,w) + F^1(\phi',s,w),$$

where

$$\|F^0(\phi',s,w)\|^2 \le C\,2^{-2j}|\sin\phi'|^{-5},$$

$$\|\left(F^1(\phi',s,w)\right)_s\|^2 + \|\left(F^1(\phi',s,w)\right)_w\|^2 \le C|\sin\phi'|^{-5}.$$

Proof (Sketch). We will adopt the same notations as in Proposition 4.2.

From (4.22), we have that

$$PF(\phi', s, w) = \int_{-\infty}^{t_0(s,w,\phi')} F\left(\rho_{\phi'}\begin{pmatrix} t \\ s \\ w \end{pmatrix}\right) dt = \int_{-\infty}^{t_0(s,w,\phi')} g(t,s,w) w_0(t,s,w)\, dt.$$

This implies that

$$(PF)_s(\phi',s,w) = g(t_0,s,w)\,w_0(t_0,s,w)\,t_{0s} + \int_{-\infty}^{t_0(s,w,\phi')} (g_s(t,s,w)\,w_0(t,s,w) + g(t,s,w)\,w_{0s}(t,s,w))\,dt$$
$$(PF)_{ss}(\phi',s,w) = T_1 + T_2 + T_3 + T_4,$$

where $T_1 = g_t\, w_0\, (t_{0s})^2 + g_s\, w_0\, t_{0s} + g\, w_0\, t_{0ss}$, $T_2 = g w_{0t}\, (t_{0s})^2 + g\, w_{0s}\, t_{0s}$, $T_3 = \int_{-\infty}^{t_0(s,w,\phi')} (g_{ss}\, w_0 + 2 g_s\, w_{0s})\, dt$, $T_4 = \int_{-\infty}^{t_0(s,w,\phi')} g\, w_{0ss}\, dt$.

From $t_0(s,w,\phi') = E^{(j)}\,(u(s,w), v(s,w))\, \sin\phi' + v(s,w)\, \cos\phi'$, using Proposition 4.4, it is easy to verify that $|t_{0s}| \le C\, \frac{1}{|\sin\phi'|}$, $|t_{0ss}| \le C\, \frac{2^{-j}}{|\sin\phi'|^3}$. It follows that $|T_1| \le C \frac{2^{-j}}{|\sin\phi'|^3}$ and, hence, $\|T_1\|^2 \le C\, \frac{2^{-2j}}{|\sin\phi'|^6} \int_A ds dw \le \frac{2^{-2j}}{|\sin\phi'|^5}$ since $\int_A ds dw \le C|\sin\phi'|$. Using the assumption that $|\sin\phi'| \ge 2^{2-j}$, one can verify that $|(T_2)_s| \le C\, \frac{1}{|\sin\phi'|^3}$. Similarly one can verify that $|T_3| \le C 2^{-j}$, and $(T_4)_s| \le C$. Thus, it follows that $\|T_3\|^2 \le C\, \frac{2^{-2j}}{|\sin\phi'|^5}$, and $\|(T_4)s\|^2 \le C\,|\sin\phi'|^5$ since $|\sin\phi'| \le 1$.

Now the argument is completed by letting $F^0(\phi', s, w) = T_1 + T_3$ and $F^1(\phi', s, w) = T_2 + T_4$. □

As a direct corollary of Proposition 4.5, it follows that

$$\int_0^\infty \int_0^{2\pi} r^5 |\hat{F}(r,\theta',\phi')|^2\, d\theta'\, dr \le C\, 2^{-2j} |\sin\phi'|^{-5},$$

which implies that

$$\int_{2^{j-2}}^{2^{j+1}} \int_0^{2\pi} |\hat{F}(r,\theta',\phi')|^2\, d\theta'\, dr \le C\, 2^{-7j} |\sin\phi'|^{-5}. \tag{4.23}$$

Using again the identity $\hat{F}(\xi) = 2^{3j} f(2^j \xi)$, from (4.23) it follows that

$$\int_{2^{2j-4}}^{2^{2j+2}} \int_0^{2\pi} |\hat{f}(r,\theta',\phi')|^2\, d\theta'\, dr \le C\, 2^{-12j} |\sin\phi'|^{-5}. \tag{4.24}$$

We can now prove Theorem 4.1

Proof of Theorem 4.1.

We only need to consider $\xi = (\xi_1, \xi_2, \xi_3)$ inside the support of $\Gamma_{j,\ell}$ By the assumptions on the support of W and v if follows that

$$\begin{aligned} \operatorname{supp} W(2^{-2j}\xi) &\subset \big\{\xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}]\big\}, \\ \operatorname{supp} v(2^j \tfrac{\xi_2}{\xi_1} - \ell) &\subset \big\{(\xi_1,\xi_2,\xi_3) : |2^j \frac{\xi_2}{\xi_1} - \ell_1| \le 1\big\}, \\ \operatorname{supp} v(2^j \tfrac{\xi_3}{\xi_1} - \ell) &\subset \big\{(\xi_1,\xi_2,\xi_3) : |2^j \frac{\xi_3}{\xi_1} - \ell_2| \le 1\big\}. \end{aligned}$$

By representing (ξ_1, ξ_2, ξ_3) using spherical coordinates as $(\lambda\cos\theta\sin\phi, \lambda\sin\theta\sin\phi, \lambda\cos\phi)$, we can write the last two expressions as

$$\begin{aligned} \operatorname{supp} v(2^j \tfrac{\xi_2}{\xi_1} - \ell) &\subset \big\{(\lambda,\theta,\phi) :\, 2^{-j}(\ell_1 - 1) \le \tan\theta \le 2^{-j}(\ell_1 + 1)\big\}, \\ \operatorname{supp} v(2^j \tfrac{\xi_3}{\xi_1} - \ell) &\subset \big\{(\lambda,\theta,\phi) :\, 2^{-j}(\ell_2 - 1) \le \frac{\cot\phi}{\cos\theta} \le 2^{-j}(\ell_2 + 1)\big\}. \end{aligned}$$

Notice that $|\theta| \le \frac{\pi}{4}$, so that $1 \le |\cos\theta| \le \frac{\sqrt{2}}{2}$.

Since $\lambda^2 = \xi_1^2 + \xi_2^2 + \xi_3^2 = \xi_1^2\,(1 + (\tan\theta)^2 + \frac{(\cot\phi)^2}{(\cos\theta)^2})$ and $|\ell_1| \le |\ell_2| \le 2^j$, it is easy to verify that

$$2^{2j-4} \le |\lambda| \le 2^{2j+2}.$$

Thus, using the fact that $\tan\phi \ge 2^{-j}\, \cos\theta\, (\ell_2 - 1)$, it follows that the support of the function $\Gamma_{j,\ell}$, given by (4.4), the set:

$$\begin{aligned} U_{j,\ell} = \{(\lambda,\theta,\phi) :\, & 2^{2j-4} \le |\lambda| \le 2^{2j+2},\ \tan^{-1}(2^{-j}(\ell_1 - 1)) \le \theta \le \tan^{-1}(2^{-j}(\ell_1 + 1)), \\ & \cot^{-1}(2^{-j}(\ell_2 - 1)) \le \phi \le \cot^{-1}(2^{-j}(\ell_2 + 1))\}. \end{aligned} \tag{4.25}$$

When (λ, θ, ϕ) is contained in the set $U_{j,\ell}$, the variables θ and ϕ are contained in intervals of length $C\,2^{-j}$, which, in the following, will be denoted by I_θ and I_ϕ, respectively. Hence, from (4.19), it follows that ϕ' is contained in an interval $I_{\phi'}$ of length $C\,2^{-j}$. Furthermore, if $(\lambda, \theta, \phi) \in U_{j,\ell}$ and $|\sin\phi'| \geq 2^{1-j}$, then $2^j|\sin\phi'|$ is equivalent to $|\ell_2|$, so that $\ell_2 \neq 0$.

Let $\xi_1 = r\sin\theta'\cos\phi'$, $\xi_2 = -r\cos\theta'$, $\xi_3 = -r\sin\theta'\sin\phi'$. A direct computation shows that the Jacobian of (ξ_1, ξ_2, ξ_3) with respect to (r, θ', ϕ') is $-r^2\sin^2\theta'$. It follows that

$$\begin{aligned}
\int_{\widehat{\mathbb{R}}^3} |\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi &\leq \int_{U_{j,\ell}} |\hat{f}(\xi)|^2\,d\xi \\
&\leq \int_{I_{\phi'}} \int_{2^{2j-2}}^{2^{2j+4}} \int_0^{2\pi} |\hat{f}(r,\theta',\phi')|^2 r^2 \sin^2\theta'\,dr\,d\phi' \\
&\leq C\,2^{4j} \int_{I_{\phi'}} \int_{2^{2j-2}}^{2^{2j+4}} \int_0^{2\pi} |\hat{f}(r,\theta',\phi')\xi)|^2\,d\theta'\,dr\,d\phi' \qquad (4.26)
\end{aligned}$$

We can now use the estimates from Propositions 4.2, 4.3 and 4.5 to complete the proof. Namely, in the non-intersection case, inequality (4.12) gives that

$$\int_{\widehat{\mathbb{R}}^3} |\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi \leq C\,2^{-9j}. \qquad (4.27)$$

For the intersection case, with the assumption that $|\sin\phi'| \leq 2^{1-j}$, Proposition 4.3 gives that

$$\int_{\widehat{\mathbb{R}}^3} |\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi \leq C\,2^{-4j}$$

Finally, for the intersection case, with the assumption that $|\sin\phi'| \geq 2^{1-j}$, inequality (4.24) yields

$$\begin{aligned}
\int_{\widehat{\mathbb{R}}^3} |\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi &\leq C\,2^{-8j} \int_{I_{\phi'}} |\sin\phi'|^{-5}\,d\phi' \\
&\leq C\,2^{-4j}\,|\ell_2|^{-5}.
\end{aligned}$$

Since $|\ell_2| \leq 2^j$, the proof of Theorem 4.1 is completed by combining the three inequalities given above. □

Before proving Theorem 3.3, we need some additional estimates involving the derivatives of the surface fragment.

Let $m = (m_1, m_2, m_3)$ and, let us adopt the usual multi-index notation where $|m| = m_1 + m_2 + m_3$, $x^m = x_1^{m_1} x_2^{m_2} x_3^{m_3}$ and $\frac{\partial^m}{\partial\xi^m}\hat{f} = \frac{\partial^{m_1}}{\partial\xi_1^{m_1}}\frac{\partial^{m_2}}{\partial\xi_2^{m_2}}\frac{\partial^{m_3}}{\partial\xi_3^{m_3}}\hat{f}(\xi)$. For a surface fragment f, we may rewrite $x^m f(x)$ as

$$x^m f(x) = 2^{-j|m|} f_m(x),$$

where $f_m(x) = g(x)(2^j x)^m\,w(2^j x)\,\chi_{[x_1 \geq E(x_2,x_3)]}(x)$ is another surface fragment. Since the Fourier transform of $x^m f(x)$ is $i^m \frac{\partial^m}{\partial\xi^m}\hat{f}$, the inequalities (4.5) and (4.27) imply the following estimates:

$$\begin{aligned}
&\int_{\widehat{\mathbb{R}}^3} |\frac{\partial^m}{\partial\xi^m}\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi \leq C\,2^{-j|m|}\,2^{-4j}(1+|\ell_2|)^{-5}, \quad \text{if there is an intersection,} \\
&\int_{\widehat{\mathbb{R}}^3} |\frac{\partial^m}{\partial\xi^m}\hat{f}(\xi)|^2\,|\Gamma_{j,\ell}(\xi)|^2\,d\xi \leq C\,2^{-j|m|}\,2^{-9j}, \quad \text{if there is no intersection.}
\end{aligned}$$

Notice that, for the non-intersection case, the estimate $2^{-j|m|}2^{-9j}$ is the best possible one. However, for the intersection case, the estimate $2^{-j|m|}2^{-4j}(1+|\ell_2|)^{-5}$ can be improved if $m_1 > 0$. The reason is that, on the surface, $|x_1| = |E^j(x_2,x_3)| \leq C\,2^{-j}$. Indeed, using the argument of Proposition 4.5 for the surface fragment

$F_m(x)$ (recall that $F_m(x) = f_m(2^{-j}x)$), if the derivatives don't involve x_1, then one obtains the additional factor 2^{-jm_1}. On the other hand, when one takes one derivative with respect to $x_1^{m_1}$, this only produces a factor $2^{-j(m_1-1)}$. However, in this last case, one can compute one additional derivative with respect to the the remaining function in the expression of $F_m(x)$ so that the missing factor 2^{-j} can be compensated, thanks to Plancherel theorem and the observation that, in the Fourier domain, the domain is restricted to the region where $2^{j-1} \le |\xi| \le 2^{j+2}$. Indeed this is the key idea in the proof of Lemma 6.2 in [2] (and hence in the proof of Proposition 4.5).

Using these observations, we obtain the following refinement of Proposition 4.5 valid for $F_m(x)$, in the case where $m_1 = 2$. The behavior for other values of m_1 is similar.

PROPOSITION 4.6. *The ray transform of F_m is twice differentiable as a function of s and w and admits the decomposition*

$$\begin{aligned}&(PF(\phi', s, w))_{ss}(\phi', s, w) + (PF(\phi', s, w))_{sw}(\phi', s, w) + (PF(\phi', s, w))_{ww}(\phi', s, w)\\ &= F^0(\phi', s, w) + F^1(\phi', s, w) + F^2(\phi', s, w) + F^3(\phi', s, w),\end{aligned}$$

where, for $q = (q_1, q_2)$ and $|q| = q_1 + q_2$, we have that

$$\begin{aligned}&\|F^0(\phi', s, w)\|^2 \le C\, 2^{-2jm_1} 2^{-2j} |\sin\phi'|^{-5},\\ &\|\big(F^1(\phi', s, w)\big)_s\|^2 + \|\big(F^1(\phi', s, w)\big)_w\|^2 \le C 2^{-2jm_1} |\sin\phi'|^{-5},\\ &\sum_{|q|=2} \|\big(F^2(\phi', s, w)\big)_{s^{q_1} w^{q_2}}\|^2 \le C 2^{-2j(m_1-1)} |\sin\phi'|^{-5},\\ &\sum_{|q|=3} \|\big(F^3(\phi', s, w)\big)_{s^{q_1} w^{q_2}}\|^2 \le C 2^{-2j(m_1-2)} |\sin\phi'|^{-5}.\end{aligned}$$

Using the assumptions on the supports of W and v and the assumption that $|\ell_1| \le |\ell_2|$, one can easily verify the following inequality (see the proof of Lemma 2.5 in [14] for a similar argument):

$$\left|\frac{\partial^m}{\partial\xi^m}\Gamma_{j,\ell}(\xi)\right| \le C_m\, 2^{-m_1 j} 2^{-|m|j} (1 + |\ell_2|)^{m_1}.$$

Since the sets U_{j,ℓ_1,ℓ_2} and $U_{j,\ell_{1'},\ell_2}$ are essentially disjoint for $\ell_1 \ne \ell_{1'}$ (that is, each point lies in a finite number of sets U_{j,ℓ_1,ℓ_2}), using the last inequality we obtain that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \left|\frac{\partial^m}{\partial\xi^m}\Gamma_{j,\ell}(\xi)\right| \le C_m\, 2^{-m_1 j} 2^{-|m|j} (1 + |\ell_2|)^{m_1}. \tag{4.28}$$

Notice that, even for $|\ell_2| = 2^j$, the above estimate is uniform for all ξ in the interior of $\mathcal{P}_1$. Exactly the same type of estimate holds for the corresponding functions defined in the other pyramidal regions. Due to the regularity of the shearlet construction, this estimate also holds for the boundary shearlet elements, which are piecewise defined.

Finally, letting $m_f = (m_{f1}, m_{f2}, m_{f3})$, $m_\gamma = (m_{\gamma 1}, m_{\gamma 2}, m_{\gamma 3})$, using Proposition 4.6 and inequality (4.28) we obtain:

$$\begin{aligned}&\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \int_{\widehat{\mathbb{R}}^3} \left|\frac{\partial^{m_f}}{\partial\xi^{m_f}}\hat{f}(\xi)\right|^2 \left|\frac{\partial^{m_\gamma}}{\partial\xi^{m_\gamma}}\Gamma_{j,\ell}(\xi)\right|^2 d\xi\\ &\quad\le C\, 2^{-2j|m_f|} \left(2^{-2jm_{f1}} 2^{-4j} (1 + |\ell_2|)^{-5} + 2^{-9j}\right) 2^{-m_{\gamma 1} j} 2^{-|m_\gamma| j} (1 + |\ell_2|)^{m_{\gamma 1}}.\end{aligned} \tag{4.29}$$

Let L be the differential operator defined by:

$$L = \left(I - \left(\frac{2^{2j}}{2\pi(1 + |\ell_2|)}\right)^2 \frac{\partial^2}{\partial\xi_1^2}\right) \left(I - \left(\frac{2^j}{2\pi}\right)^2 \frac{\partial^2}{\partial\xi_2^2}\right) \left(I - \left(\frac{2^j}{2\pi}\right)^2 \frac{\partial^2}{\partial\xi_3^2}\right). \tag{4.30}$$

From inequality (4.29), a routine calculation gives the following theorem which extends the result in Theorem 4.1 (again using the fact that $|\ell_2| \le 2^j$).

THEOREM 4.7. *Let f be the surface fragment given by expression (4.1) and $\Gamma_{j,\ell}$ be given by (4.4). Then, for $j \ge 0$ and $-2^j \le \ell_2 \le 2^j$, the following estimate holds:*

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \int_{\widehat{\mathbb{R}}^3} \left| L\left(\hat{f}(\xi)\,\Gamma_{j,\ell}(\xi)\right)\right|^2 d\xi \le C\,2^{-4j}\,(1+|\ell_2|)^{-5}.$$

It is clear that the same type of result will hold for functions defined on the other pyramidal regions. In particular, for

$$\Gamma^{(2)}_{j,\ell}(\xi) = W\big(2^{-2j}\,\xi\big)\; v\Big(2^j\,\frac{\xi_1}{\xi_2} - \ell_1\Big)\; v\Big(2^j\,\frac{\xi_3}{\xi_2} - \ell_2\Big),$$

which is supported in the region $\mathcal{P}_2$, one obtains the following analogue of the estimate from Theorem 4.7:

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \int_{\widehat{\mathbb{R}}^3} \left| L_2\left(\hat{f}(\xi)\,\Gamma^{(2)}_{j,\ell}(\xi)\right)\right|^2 d\xi \le C\,2^{-4j}\,(1+|\ell_2|)^{-5},$$

where

$$L_2 = \left(I - \left(\frac{2^{2j}}{2\pi\,(1+|\ell_2|)}\right)^2 \frac{\partial^2}{\partial\xi_2^2}\right)\left(1 - \left(\frac{2^j}{2\pi}\right)^2 \frac{\partial^2}{\partial\xi_1^2}\right)\left(1 - \left(\frac{2^j}{2\pi}\right)^2 \frac{\partial^2}{\partial\xi_3^2}\right).$$

A similar result holds for the region $\mathcal{P}_3$.

In the following section, the estimates above will be used to analyze the shearlet coefficients $\langle f, \widetilde{\psi}_\mu\rangle$. In particular, the result of Theorem 4.7 relates directly to the analysis of the interior shearlets in $\mathcal{P}_1$. As shown above, the interior shearlets associated with the other pyramidal regions can be handled in a very similar way. For the boundary shearlets, the situation is as follows. Consider, for example, the boundary shearlets corresponding to the boundary of $\mathcal{P}_1$ and $\mathcal{P}_2$. In this case, we define (for both regions where the shearlets are piecewise defined) the differential operator

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

Since $\ell_2 = \pm 2^j$ for the boundary shearlets, it follows that the operator $L_{1,2}$ is equivalent to L_1 on $\mathcal{P}_1$ and to L_2 on $\mathcal{P}_2$, so that the analysis of the boundary shearlets is equivalent to the interiors ones. Thus, in the following section, it will be sufficient to consider the interior shearlets associated with $\mathcal{P}_1$ only.

4.4. Proof of Theorem 3.3. Using the preparatory work from the previous sections, we are now ready to prove Theorem 3.3.

Fix $j \ge 0$ and, for simplicity of notation, let $f = f_Q$. As discussed above, it will be sufficient to consider the system interior shearlets in the pyramidal region $\mathcal{P}_1$ only.

For $\mu \in M_j$, the shearlet coefficients of f associated with the interior shearlets in $\mathcal{P}_1$ can be expressed as

$$\langle f, \widetilde{\psi}_\mu\rangle = \langle f, \psi^{(1)}_{j,\ell,k}\rangle = |\det A_{(1)}|^{-j/2} \int_{\widehat{\mathbb{R}}^2} \hat{f}(\xi)\,\Gamma_{j,\ell}(\xi)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\, d\xi,$$

where $\Gamma_{j,\ell}$ is given by (4.4). By the equivalent definition of weak ℓ^1 norm, the theorem is proved provided we show that

$$\#\{\mu \in M_j : \left|\langle f, \widetilde{\psi}_\mu\rangle\right| > \epsilon\} \le C\,2^{-2j}\,\epsilon^{-1}. \qquad (4.31)$$

Observe that

$$\xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k = \begin{pmatrix}\xi_1 & \xi_2 & \xi_3\end{pmatrix} \begin{pmatrix} 2^{-2j} & 0 & 0 \\ 0 & 2^{-j} & 0 \\ 0 & 0 & 2^{-j} \end{pmatrix} \begin{pmatrix} 1 & -\ell_1 & -\ell_2 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$$
$$= (k_1 - k_2\ell_1 - k_3\ell_2)\, 2^{-2j}\xi_1 + k_2 2^{-j}\xi_2 + k_3 2^{-j}\xi_3. \tag{4.32}$$

Let L be the second order differential operator defined by (4.30). It is easy to check that

$$L\left(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\right) = \begin{cases} \left(1 + (\frac{|\ell_2|}{(1+|\ell_2|)})^2 (\frac{k_1}{|\ell_2|} - \frac{k_2\ell_1}{|\ell_2|} \pm k_3)^2\right)(1+k_2^2)(1+k_3^2)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k} & \text{if } \ell_2 \neq 0 \\ (1+k_1^2)(1+k_2^2)(1+k_3^2)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k} & \text{if } \ell_2 = 0, \end{cases} \tag{4.33}$$

where we have $\pm k_3$ depending on whether ℓ_2 is positive or negative. Using integration by parts, we have:

$$\langle f, \widetilde{\psi}_\mu \rangle = |\det A_{(1)}|^{-j/2} \int_{\widehat{\mathbb{R}}^3} L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) L^{-1}\left(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\right) d\xi.$$

To analyze this quantity, we will consider separately the case $\ell \neq 0$ and $\ell = 0$.

Case 1: $\ell_2 \neq 0$**.** In this case, using (4.33), we have that

$$L^{-1}\left(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\right) = G(k,\ell)^{-1}\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}, \tag{4.34}$$

where $G(k,\ell) = \left(1 + (\frac{|\ell_2|}{(1+|\ell_2|)})^2 (\frac{k_1}{|\ell_2|} - \frac{k_2\ell_1}{|\ell_2|} \pm k_3)^2\right)(1+k_2^2)(1+k_3^2)$. Thus, we have that

$$\langle f, \widetilde{\psi}_\mu \rangle = |\det A_{(1)}|^{-j/2}\, G(k,\ell)^{-1} \int_{\widehat{\mathbb{R}}^3} L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\, d\xi,$$

or, equivalently, that

$$G(k,\ell)\, \langle f, \widetilde{\psi}_\mu \rangle = |\det A_{(1)}|^{-j/2} \int_{\widehat{\mathbb{R}}^3} L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k}\, d\xi.$$

Let $K = (K_1, K_2, K_3) \in \mathbb{Z}^3$ and define $R_K = \{k = (k_1,k_2,k_3) \in \mathbb{Z}^3 : \frac{k_1}{|\ell_2|} \in [K_1, K_1+1],\ -\frac{k_2\ell_1}{|\ell_2|} \in [K_2, K_2+1],\ k_3 = K_3\}$. Since, for j,ℓ fixed, the set $\{|\det A_{(1)}|^{-j/2}\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell]} k} : k \in \mathbb{Z}^3\}$ is an orthonormal basis for the L^2 functions on $[-\frac{1}{2}, \frac{1}{2}] A_{(1)}^{j} B_{(1)}^{[\ell]}$, and the function $\Gamma_{j,\ell}(\xi)$ is supported on this set, then

$$\sum_{k \in R_K} G(k,\ell)^2\, |\langle f, \widetilde{\psi}_\mu \rangle|^2 \le \int_{\widehat{\mathbb{R}}^3} \left| L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) \right|^2 d\xi.$$

This implies that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \sum_{k \in R_K} G(k,\ell)^2\, |\langle f, \widetilde{\psi}_\mu \rangle|^2 \le \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \int_{\widehat{\mathbb{R}}^3} \left| L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) \right|^2 d\xi.$$

From the definition of R_K, it follows that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \sum_{k \in R_K} |\langle f, \widetilde{\psi}_\mu \rangle|^2 \le C\, \left(1 + (K_1 - K_2 \pm K_3)^2\right)^{-2} (1+K_2^2)^{-2} (1+K_3^2)^{-2} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \int_{\widehat{\mathbb{R}}^3} \left| L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) \right|^2 d\xi.$$

Thus, by Theorem 4.7, we have that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} \sum_{k \in R_K} |\langle f, \widetilde{\psi}_\mu \rangle|^2 \le C\, L_K^{-2}\, 2^{-4j} (1+|\ell_2|)^{-5}, \tag{4.35}$$

where $L_K = \left(1 + (K_1 - K_2 \pm K_3)^2\right)(1+K_2^2)(1+K_3^2)$.

For j, ℓ fixed, let $N_{j,\ell,K}(\epsilon) = \#\{k \in R_K : |\langle f, \psi^{(1)}_{j,\ell,k}\rangle| > \epsilon\}$. Since $|\ell_1| \le |\ell_2|$, it is clear that $N_{j,\ell,K}(\epsilon) \le C\,(1+|\ell_2|)^2$ (C is independent of ℓ_1) and, hence, $\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \le C\,(1+|\ell_2|)^3$. Using the new notation, from (4.35) we have that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \le C\, L_K^{-2}\, 2^{-4j}\, \epsilon^{-2} (1+|\ell_2|)^{-5}.$$

This implies that

$$\sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \le C\, \min\left((|\ell_2|+1)^3, L_K^{-2}\, 2^{-4j}\, \epsilon^{-2}(1+|\ell_2|)^{-5}\right). \tag{4.36}$$

Using (4.36) we will now show that:

$$\sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \le C\, L_K^{-1}\, 2^{-2j}\, \epsilon^{-1}. \tag{4.37}$$

In fact, let ℓ_2^* be defined by $(\ell_2^*+1)^3 = L_K^{-2}\, 2^{-4j}\, \epsilon^{-2}(1+\ell_2^*)^{-5}$. That is, $(\ell_2^*+1)^4 = L_K^{-1}\, 2^{-2j}\, \epsilon^{-1}$. Then

$$\begin{aligned}
\sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) &\le \sum_{|\ell_2| \le (\ell_2^*+1)} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) + \sum_{|\ell_2| > (\ell_2^*+1)} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \\
&\le \sum_{|\ell_2| \le (\ell_2^*+1)} (|\ell_2|+1)^3 + \sum_{|\ell_2| > (\ell_2^*+1)} L_K^{-2}\, 2^{-4j}\, \epsilon^{-2}(1+|\ell_2|)^{-5} \\
&\le C\ell_2^*+1)^4 + C L_K^{-2}\, 2^{-4j}\, \epsilon^{-2}(1+\ell_2^*)^{-4} \le C\,(\ell_2^*+1)^4,
\end{aligned}$$

which gives (4.37).

Since $\sum_{K \in \mathbb{Z}^3} L_K^{-1} < \infty$, using (4.37) we then have that

$$\#\{\mu \in M_j : |\langle f, \widetilde{\psi}_\mu\rangle| > \epsilon\} \le \sum_{K \in \mathbb{Z}^3} \sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-|\ell_2|}^{|\ell_2|} N_{j,\ell,K}(\epsilon) \le C\, 2^{-2j}\, \epsilon^{-1} \sum_{K \in \mathbb{Z}^3} L_K^{-1} \le C\, 2^{-2j}\, \epsilon^{-1},$$

and, thus, (4.31) holds.

Case 2: $\ell_2 = 0$**.** In this case, we also have $\ell_1 = 0$. It follows that

$$L^{-1}\left(e^{2\pi i \xi A_{(1)}^{-j} k}\right) = (1+k_1^2)^{-1}(1+k_2^2)^{-1}(1+k_3^2)^{-1}\, e^{2\pi i \xi A_{(1)}^{-j} k}.$$

Let $L_k = (1+k_1^2)\,(1+k_2^2)\,(1+k_3^2)$. It is clear that also in this case $\sum_{k \in \mathbb{Z}^3} L_k^{-1} < \infty$. We have

$$\langle f, \psi^{(1)}_{j,0,k}\rangle = |\det A_{(1)}|^{-j/2} L_k^{-1} \int_{\widehat{\mathbb{R}}^3} L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi)\right) e^{2\pi i \xi A_{(1)}^{-j} k}\, d\xi,$$

or, equivalently, that

$$\langle f, \psi^{(1)}_{j,0,k} \rangle L_k = |\det A_{(1)}|^{-j/2} \int_{\widehat{\mathbb{R}}^3} L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi) \right) e^{2\pi i \xi A^{-j}_{(1)} k}\, d\xi,$$

It follows that

$$\sum_{k \in \mathbb{Z}^3} L_k^2\, |\langle f, \psi^{(1)}_{j,0,k} \rangle|^2 = \int_{\widehat{\mathbb{R}}^3} \left| L\left(\hat{f}(\xi)\, \Gamma_{j,\ell}(\xi) \right) \right|^2 d\xi \le C 2^{-4j}.$$

In particular, for each $k \in \mathbb{Z}^3$, we have $|\langle f, \psi^{(1)}_{j,0,k} \rangle| \le C\, L_k^{-1}\, 2^{-2j}$ and hence $\sum k \in \mathbb{Z}^3 |\langle f, \psi^{(1)}_{j,0,k} \rangle| \le C\, 2^{-2j}$, or $\|\langle f, \psi^{(1)}_{j,0,k} \rangle\|_{l^1} \le C\, 2^{-2j}$ which implies $\|\langle f, \psi^{(1)}_{j,0,k} \rangle\|_{wl^1} \le C\, 2^{-2j}$.

This completes the proof of the theorem. □

4.5. Remark on the proof of Theorem 3.3. In the proof of Theorem 3.3, it was assumed that the boundary surface contains the origin and has normal direction $(1, 0, 0)$ at the origin. In general one can "transform" any given surface into the above special case by using a combination of translation and rotation. Obviously the translation has no impact on the proof which was given above. It only remains to explain the effect of rotations, since the shearlet system is not invariant with respect to rotations.

As in the proof of Theorem 4.1, let us consider $\xi = (\xi_1, \xi_2, \xi_3) \in \mathcal{P}_1$. Recall that the support of the function $\Gamma_{j,\ell}$, given by (4.4), is contained in a set $U_{j,\ell}$ which, using spherical coordinates, is given by:

$$\begin{aligned} U_{j,\ell} = \{(\lambda, \theta, \phi) :\ & 2^{2j-4} \le |\lambda| \le 2^{2j+2},\ \tan^{-1}(2^{-j}(\ell_1 - 1)) \le \theta \le \tan^{-1}(2^{-j}(\ell_1 + 1)), \\ & \cot^{-1}(2^{-j}(\ell_2 - 1)) \le \phi \le \cot^{-1}(2^{-j}(\ell_2 + 1))\}. \end{aligned}$$

Also, as in the proof of Theorem 4.1, for the special case considered, we notice that there are positive C_1, C_2 such that, on the set $U_{j,\ell}$, we have $C_1 2^{-j} \le |\theta| \le C_2 2^{-j}$, $C_1 2^{-j} \le |\phi| \le C_2 2^{-j}$ and $C_1 2^{-j} |\ell_2| \le |\cos\phi| \le C_2 2^{-j} |\ell_2|$ for all large j and all $|\ell_2| \le 2^j$ (recall that $|\phi - \frac{\pi}{2}| \le \frac{\pi}{4}$). Let θ' and ϕ' be the angles derived from the ray transform. In the proof of Proposition 4.3, we obtain the following identities.

$$\begin{aligned} \sin\theta' \cos\phi' &= \cos\theta \sin\phi, \\ \cos\theta' &= \sin\theta \sin\phi, \\ -\sin\theta' \sin\phi' &= \cos\phi. \end{aligned}$$

Next, we deduced that $|\sin\phi'|$ is equivalent to $|\cos\phi|$ on $U_{j,\ell}$ and the rest of the proof (for the special case) follows from there.

In spherical coordinates, a rotation can be realized by the mapping $(\lambda, \theta, \phi) \to (\lambda, \theta - \theta_0, \phi - \phi_0)$, where θ_0 and ϕ_0 are the two rotation angles. Let $\Gamma^0_{j,\ell}(\xi)$, $U^0_{j,\ell}$ and $\hat{f}^0$ be the images of $\Gamma_{j,\ell}(\xi)$, $U_{j,\ell}$ and $\hat{f}$, respectively, under the rotation by θ_0 and ϕ_0. Then we have

$$\begin{aligned} U^0_{j,\ell} = \{(\lambda, \theta, \phi) :\ & 2^{2j-4} \le |\lambda| \le 2^{2j+2},\ \tan^{-1}(2^{-j}(\ell_1 - 1)) \le \theta - \theta_0 \le \tan^{-1}(2^{-j}(\ell_1 + 1)), \\ & \cot^{-1}(2^{-j}(\ell_2 - 1)) \le \phi - \phi_0 \le \cot^{-1}(2^{-j}(\ell_2 + 1))\}. \end{aligned}$$

Now one can adapt the ray transform to the rotation angles by letting

$$\begin{aligned} \sin\theta' \cos\phi' &= \cos(\theta - \theta_0) \sin(\phi - \phi_0), \\ \cos\theta' &= \sin(\theta - \theta_0) \sin(\phi - \phi_0), \\ -\sin\theta' \sin\phi' &= \cos(\phi - \phi_0). \end{aligned}$$

It follows that $|\sin\phi'|$ is equivalent to $|\cos(\phi - \phi_0)|$. On $U^0_{j,\ell}$, we have $C_1 2^{-j} |\ell_2| \le |\cos(\phi - \phi_0)| \le C_2 2^{-j} |\ell_2|$. Hence it follows that

$$\int_{\widehat{\mathbb{R}}^3} |\hat{f}^0(\xi)|^2\, |\Gamma^0_{j,\ell}(\xi)|^2\, d\xi \le C\, 2^{-4j} (1 + |\ell_2|)^{-5}.$$

The rest of the argument is exactly the same as in the proof of Theorem 3.3, where $\hat{f}(\xi)$ and $\Gamma_{j,\ell}(\xi)$ are replaced by $\hat{f}^0(\xi)$ and $\Gamma^0_{j,\ell}(\xi)$.

4.6. Analysis of the coarse scale. At the beginning of Section 4.1, we assumed that the scale parameter j is large enough, i.e., $j > j_0$ for some $j_0 > 0$. The situation where $j \le j_0$ is much simpler. In fact, if f_Q is an edge fragment, then a trivial estimate shows that

$$\|f_Q\|_2 = \left(\int_{\operatorname{supp} w_Q} |f_Q(x)|^2 \, dx \right)^{1/2} \le C \, |\operatorname{supp} w_Q|^{1/2} = C \, 2^{-\frac{3}{2}j}.$$

It follows that $\|\langle f_Q, \psi_\mu \rangle\|_{\ell^2} \le C \, \|f_Q\|_2 \le C \, 2^{-\frac{3}{2}j}$. To deduce an ℓ^1 type estimate, we notice that

$$\|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{\ell^p} \le N^{\frac{1}{p} - \frac{1}{2}} \, \|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{\ell^2},$$

is valid for any sequence $\{\langle f_Q, \psi_\mu \rangle\}$ of N elements. Since, at scale 2^{-j}, there are about 2^{2j} shearlet elements in $\mathcal{Q}_j^0$, it follows that

$$\|\langle f_Q, \widetilde{\psi}_\mu \rangle\|_{\ell^1} \le C \, 2^j \, 2^{-\frac{3}{2}j} = C \, 2^{-\frac{1}{2}j}.$$

This satisfies Theorem 3.3 for $j \le j_0$.

4.7. Proof of Theorem 3.4. The proof of Theorem 3.4 follows essentially the idea from the 2-dimensional case in [14]. We start by proving the following lemmata which will be useful in the following.

LEMMA 4.8. *Let $f = g \, w_Q$, where $g \in \mathcal{E}^2(A)$ and $Q \in \mathcal{Q}_j^1$ and $U_{j,\ell}$ be given by (4.25). Then*

$$\int_{U_{j,\ell}} |\hat{f}(\xi)|^2 \, d\xi \le C \, 2^{-11j}. \tag{4.38}$$

Proof. The following proof adapts [14, Lemma 2.6].

The function f belongs to $C_c^2(\mathbb{R}^3)$ and its second partial derivative with respect to x_1 is

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{\partial^2 g}{\partial x_1^2} \, w_Q + 2 \, \frac{\partial \, g}{\partial x_1} \, \frac{\partial \, w_Q}{\partial x_1} + f \, \frac{\partial^2 w_Q}{\partial x_1^2} := h_1 + h_2 + h_3.$$

Using the fact that w_Q is supported in a square of sidelength $2 \cdot 2^{-j}$, we have

$$\int_{\widehat{\mathbb{R}}^3} |\hat{h}_1(\xi)|^2 \, d\xi = \int_{\mathbb{R}^3} |h_1(x)|^2 \, dx \le C \, 2^{-3j}.$$

Next, observe that $\|\frac{\partial}{\partial x_1} h_2\|_\infty \le C \, 2^{2j}$. Using again the condition on the support of w_Q it follows that

$$\int_{\widehat{\mathbb{R}}^3} |2\pi \xi_1 \, \hat{h}_2(\xi)|^2 \, d\xi = \int_{\mathbb{R}^3} \left| \frac{\partial}{\partial x_1} h_2(x) \right|^2 \, dx \le C \, 2^j,$$

and thus, for $\xi \in U_{j,\ell}$ (hence $\xi_1 \approx 2^{2j}$),

$$\int_{U_{j,\ell}} |\hat{h}_2(\xi)|^2 \, d\xi \le C \, 2^{-3j}.$$

Finally, observing that $\|\frac{\partial^2}{\partial x_1^2} h_3\|_\infty \le C \, 2^{4j}$, it follows that $\int_{\widehat{\mathbb{R}}^3} |\hat{h}_3(\xi)|^2 \, d\xi \le C \, 2^{5j}$ and, thus,

$$\int_{U_{j,\ell}} |\hat{h}_3(\xi)|^2 \, d\xi \le C \, 2^{-3j}.$$

Since $-(2\pi)^2 \, \xi_1^2 \, \hat{f}(\xi) = \hat{h}_1(\xi) + \hat{h}_2(\xi) + \hat{h}_3(\xi)$, it follows from the estimates above that

$$\int_{U_{j,\ell}} |\hat{f}(\xi)|^2 \, d\xi \le C \, 2^{-11j}.$$

This completes the proof.

LEMMA 4.9. *Let $m = (m_1, m_2, m_3) \in \overline{\mathbb{N}} \times \overline{\mathbb{N}} \times \overline{\mathbb{N}}$, $\xi = (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3$ and $\Gamma_{j,\ell}$ be given by (4.4), where $\ell = (\ell_1, \ell_2)$. Then*

$$\sum_{\ell_1=-2^j}^{2^j} \sum_{\ell_2=-2^j}^{2^j} \left| \frac{\partial^m}{\partial \xi^m} \Gamma_{j,\ell_1,\ell_2}(\xi) \right|^2 \leq C_m\, 2^{-2|m|j},$$

where C_m is independent of j and ξ and $|m| = m_1 + m_2 + m_3$.

Proof. Observe that $U_{j,\ell} \cap U_{j,\ell+\ell'} = \emptyset$, whenever $|\ell'_1| \geq 3$ or $|\ell'_2| \geq 3$. Since $|\ell_1|, |\ell_2| \leq 2^j$, the lemma then follows from (4.28).

LEMMA 4.10. *Let $f = g\, w_Q$, where $g \in \mathcal{E}^2(A)$ and $Q \in \mathcal{Q}_j^1$ and set*

$$T = \left(I - \frac{2^j}{(2\pi)^2}\, \Delta \right), \tag{4.39}$$

where $\Delta = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} + \frac{\partial^2}{\partial \xi_2^2}$. Then

$$\int_{\widehat{\mathbb{R}}^3} \sum_{\ell_1=-2^j}^{2^j} \sum_{\ell_2=-2^j}^{2^j} \left| T^2 \left(\hat{f}\, \Gamma_{j,\ell_1,\ell_2} \right)(\xi) \right|^2 d\xi \leq C\, 2^{-11j}.$$

Proof. Observe that, for $N \in \overline{\mathbb{N}}$,

$$\Delta^N \left(\hat{f}\, \Gamma_{j,\ell} \right) = \sum_{|\alpha|+|\beta|=2N} C_{\alpha,\beta} \left(\frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f} \right) \left(\frac{\partial^\beta}{\partial \xi^\beta}\, \Gamma_{j,\ell} \right),$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, $\beta = (\beta_1, \beta_2, \beta_3)$, and $\alpha_i, \beta_i \in \mathbb{N}$. Also notice that, by Lemma 4.9, we have that

$$\int_{\widehat{\mathbb{R}}^3} \sum_{\ell_1,\ell_2=-2^j}^{2^j} \left| \frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f}(\xi) \right|^2 \left| \frac{\partial^\beta}{\partial \xi^\beta}\, \Gamma_{j,\ell}(\xi) \right|^2 d\xi \leq C_\beta\, 2^{-2|\beta|j} \int_{U_{j,\ell}} \left| \frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f}(\xi) \right|^2 d\xi.$$

Recall that $f(x)$ is of the form $g(x)\, w(2^j x)$. It follows that $x^\alpha f(x) = 2^{-j|\alpha|}\, g(x)\, w_\alpha(2^j x)$, where $w_\alpha(x) = x^\alpha w(x)$. By Lemma 4.8, $g(x)\, w_\alpha(2^j x)$ obeys the estimate (4.38). Thus, observing that $\frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f}(\xi)$ is the Fourier transform of $(-2\pi i x)^\alpha f(x)$, we have that

$$\int_{U_{j,\ell}} \left| \frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f}(\xi) \right|^2 d\xi \leq C_\alpha\, 2^{-2j|\alpha|}\, 2^{-11j}.$$

Combining the estimates above we have that, for each α, β with $|\alpha| + |\beta| = 2N$,

$$\int_{\widehat{\mathbb{R}}^3} \sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-2^j}^{2^j} \left| \frac{\partial^\alpha}{\partial \xi^\alpha}\, \hat{f}(\xi) \right|^2 \left| \frac{\partial^\beta}{\partial \xi^\beta}\, \Gamma_{j,\ell}(\xi) \right|^2 d\xi \leq C_{\alpha,\beta}\, 2^{-11j}\, 2^{-4jN}. \tag{4.40}$$

Since $T^2 = 1 - 2\, \frac{2^j}{(2\pi)^2}\, \Delta + \frac{2^{2j}}{(2\pi)^4}\, \Delta^2$, the lemma now follows from (4.40) and Lemma 4.9.

We can now prove Theorem 3.4.

Proof of Theorem 3.4.

As in the arguments above, it is sufficient to consider the system interior shearlets in the pyramidal region $\mathcal{P}_1$.

For T given by (4.39) and $\ell = (\ell_1, \ell_2)$, a direct computation gives that

$$T\left(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-[\ell]} k}\right) = \left(1 + 2^{-2j}(k_1 - k_2\,\ell_1 - k_3\,\ell_2)^2 + k_2^2 + k_3^2\right)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-[\ell]} k}. \tag{4.41}$$

Hence,

$$T^2\left(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-[\ell]} k}\right) = \left(1 + 2^{-2j}(k_1 - k_2\,\ell_1 - k_3\,\ell_2)^2 + k_2^2 + k_3^2\right)^2\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-[\ell]} k}. \tag{4.42}$$

Fix $j \geq 0$ and let $f = f_Q$, with $Q \in \mathcal{Q}_j^1$. Then, using integration by parts as in the proof of Theorem 3.3, from (4.42) it follows that

$$\langle f, \widetilde{\psi}_\mu\rangle = |\det A_{(1)}|^{-j}\left(1 + 2^{-2j}(k_1 - k_2\,\ell_1 - k_3\,\ell_2)^2 + k_2^2 + k_3^2\right)^{-2} \int_{\widehat{\mathbb{R}}^2} T^2\left(\hat f(\xi)\,\Gamma_{j,\ell}(\xi)\right)\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-[\ell]} k}\, d\xi.$$

Let $K = (K_1, K_2, K_3) \in \mathbb{Z}^3$ and R_K be the set

$$R_K = \{(k_1, k_2, k_2) \in \mathbb{Z}^3 : k_2 = K_2, k_3 = K_3, 2^{-j}(k_1 - K_2\ell_1 - K_3\ell_2) \in [K_1, K_1 + 1]\}.$$

Observe that, for each K and each fixed ℓ, there are only $1 + 2^j$ choices for k_1 in R_K. In fact, $R_K = \{k_1 : 2^j K_1 \leq k_1 - K_2\ell_1 - K_3\ell_2 \leq 2^j(K_1 + 1)\}$. Hence the number of terms in R_K is bounded by $1 + 2^j$. Also notice that, as in the proof of Theorem 3.3, we can take advantage of the fact that, for j, ℓ fixed, the set $\{|\det A_{(1)}|^{-j/2} e^{2\pi i A_{(1)}^{-j} B_{(1)}^{[\ell]} k} : k \in \mathbb{Z}^3\}$ is an orthonormal basis for the L^2 functions supported in the set $[-\frac{1}{2}, \frac{1}{2}]^3 A_{(1)}^j B_{(1)}^{[\ell]}$. Thus, using this observation and the fact that the function $\Gamma_{j,\ell}$ is supported on the set $[-\frac{1}{2}, \frac{1}{2}]^3 A_{(1)}^j B_{(1)}^{[\ell]}$, we have that

$$\sum_{k \in R_K} |\langle f, \widetilde{\psi}_\mu\rangle|^2 \leq C\,\left(1 + K_1^2 + K_2^2 + K_3^2\right)^{-4} \int_{\widehat{\mathbb{R}}^3} \left|T^2\left(\hat f\,\Gamma_{j,\ell}\right)(\xi)\right|^2 d\xi.$$

From this inequality, using Lemma 4.10, we have that

$$\begin{aligned}\sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-2^j}^{2^j} \sum_{k \in R_K} |\langle f, \widetilde{\psi}_\mu\rangle|^2 &\leq C\,(1 + K^2)^{-4} \int_{\widehat{\mathbb{R}}^2} \sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-2^j}^{2^j} \left|T^2\left(\hat f\,\Gamma_{j,\ell}\right)(\xi)\right|^2 d\xi \\ &\leq C\,(1 + K^2)^{-4}\, 2^{-11j}. \end{aligned} \tag{4.43}$$

For any $N \in \mathbb{N}$, the Hölder inequality yields:

$$\sum_{m=1}^{N} |a_m| \leq \left(\sum_{m=1}^{N} |a_m|^2\right)^{\frac{1}{2}} N^{\frac{1}{2}}. \tag{4.44}$$

Since the cardinality of R_K is bounded by $1 + 2^j$, it follows from (4.43) and (4.44) that

$$\sum_{\ell_2=-2^j}^{2^j} \sum_{\ell_1=-2^j}^{2^j} \sum_{k \in R_K} |\langle f, \widetilde{\psi}_\mu\rangle| \leq C\,\left(2^{3j}\right)^{\frac{1}{2}} (1 + K^2)^{-2}\, 2^{-\frac{11}{2}j} \leq C\,2^{-4j}.$$

Thus, for $f = f_Q$, with $Q \in \mathcal{Q}_j^1$, we have that:

$$\sum_{\mu \in M_j} |\langle f, \widetilde{\psi}_\mu\rangle| \leq C\,2^{-4j}.$$

5. Discussion and Extensions. In this section, we collect some observations which are relevant to the results discussed in the paper. In Sec. 5.1, we extend to 3D the result of Donoho about the optimal degree of sparsity for the representation of piecewise smooth functions. This shows that, in a certain sense, N^{-1} is the optimal approximation error rate that can be achieved for functions in $\mathcal{E}^2(A)$. Next, in Sec. 5.2, we discuss the extension of the results presented in this paper to the more general setting where the surfaces of discontinuities are allowed to be piecewise smooth.

5.1. Optimal Approximation Rates. In [9], Donoho investigates the problem of finding an optimal dictionary for spaces of synthetic images that can provide a simplified model of natural images. In particular, the class of *Star-Shaped Images* $\mathcal{C}^\alpha(\mathbb{R}^2) \subset L^2(\mathbb{R}^2)$ is introduced, whose elements are defined as characteristic functions of star-shaped sets with $C^\alpha(\mathbb{R}^2)$ smooth boundaries. The approach developed in [9] considers adaptive decompositions in an overcomplete (possibly uncountable) dictionary $\Phi = \{\phi_i : i \in I\} \subset L^2(\mathbb{R}^2)$ of the form

$$f = \sum_{i \in I_f} c_i(f)\, \phi_i, \tag{5.1}$$

where $\{\phi_i : i \in I_f\}$ is a countable normalized subset of Φ, which depends on f, in general. To avoid situations which are computationally unfeasible, the selection of I_f in I is required to satisfy a *polynomial depth search* constraint. That is, the i-th term in the expansion is selected according to the selection function $\sigma(i, f)$ which obeys $\sigma(i, f) \le p(i)$, for a fixed polynomial $p(i)$. The sparsity of the expansion (5.1) is measured in terms of the quasi-norm $\|c(f)\|_{w\ell^p}$, where $c(f) = (c_i(f))$, with the *optimal degree of sparsity* being associated with the smallest p such that $\|c(f)\|_{w\ell^p}$ is bounded. The main result in [9] is that the optimal degree of sparsity for the class $\mathcal{C}^\alpha(\mathbb{R}^2)$ is $p_0 = 2/(\alpha+1)$. That is, no representation system satisfying the polynomial depth search constraint can provide approximations for $\mathcal{C}^\alpha(\mathbb{R}^2)$ with the coefficients $\|c(f)\|_{w\ell^p} < \infty$, for $p < p_0$. Hence, if we have that $\|c(f)\|_{w\ell^{p_0}} < \infty$, denoting by $|c(f)|_m$ the m-th largest entry in the coefficient sequence $(|c(f)|)$, there is a constant $C > 0$ such that

$$\sup_{f \in \mathcal{C}^\alpha} |c(f)|_m \le C\, m^{-\frac{\alpha+1}{2}}, \tag{5.2}$$

and no decay rate faster than $m^{-\frac{\alpha+1}{2}}$ is possible. As a consequence, if Φ is an optimal dictionary for $C^\alpha(\mathbb{R}^2)$ and is also a Parseval frame, then, from (5.2) it follows that

$$\|f - f_N\|^2 \le \sum_{m>N} |c(f)|_m^2 \le C \sum_{m>N} m^{-(\alpha+1)} \le C\, N^{-\alpha}.$$

If, in addition, we have that Φ is a Riesz basis, then we can conclude that

$$\|f - f_N\|^2 \approx C\, N^{-\alpha},$$

so that in this case $O(N^{-\alpha})$ is truly the optimal decay rate. If Φ is only a Parseval frame but not a Riesz basis, even though one cannot ensure that $O(N^{-\alpha})$ is truly the optimal decay rate, yet no better approximation can be achieved under the procedure described above. In this weaker sense, as it is used in [2], the rate $O(N^{-\alpha})$ is identified as the optimal approximation rate for functions in $\mathcal{C}^\alpha(\mathbb{R}^2)$. When $\alpha = 2$, this gives the optimal approximation rate which was mentioned in the introduction and which is nearly achieved by 2D shearlet and curvelet approximations (cf. eq. (1.1)).

The result about the optimal degree of sparsity in $\mathcal{C}^\alpha(\mathbb{R}^2)$ from [9] follows from an information theoretic argument which leads to determine the values p such that $\mathcal{C}^\alpha(\mathbb{R}^2)$ *contains a copy of* ℓ_0^p. By definition, a function class $\mathcal{F}$ is said to contain a copy of ℓ_0^p if $\mathcal{F}$ contains embedded orthogonal hypercubes of dimension $M(\Delta)$ and side Δ, and if, for some sequence $(\Delta_k) \to 0$, there is a constant $C > 0$ such that

$$M(\Delta_k) \ge C\, \Delta_k^{-p}, \quad k = k_0, k_0+1, \ldots ..$$

Thus, to extend the result about the optimal degree of sparsity p_0 to 3D, it is sufficient to extend Thm. 3 in [9] by proving the following result.[3]

THEOREM 5.1. *The class $\mathcal{C}^2(\mathbb{R}^3)$ contains a copy of ℓ_0^p for $p = 1$.*

Here $\mathcal{C}^2(\mathbb{R}^3) \subset L^2(\mathbb{R}^3)$ is the class Star-Shaped 3D Images, whose elements are the characteristic functions of 3-dimensional star-shaped sets with C^2 smooth boundaries. It is clear that the class $\mathcal{E}^2(A)$ of piecewise smooth function of 3 variables considered in this paper contains the class $\mathcal{C}^2(\mathbb{R}^3)$. Hence, Theorem 5.1 shows that no representation system satisfying the polynomial depth search constraint can provide approximations for $\mathcal{C}^2(\mathbb{R}^3)$ with the coefficients $\|c(f)\|_{w\ell^p} < \infty$, for $p < 1$, where $p_0 = 1$ is the optimal degree of sparsity. From that it follows that if we have that $\|c(f)\|_{w\ell^1} < \infty$, then there is a constant $C > 0$ such that

$$\sup_{f \in \mathcal{C}^2} |c(f)|_m \le C\, m^{-1},$$

and no decay rate faster than m^{-1} is possible. Using the same argument as above we conclude that, if f_N is the best N term approximation to $f \in \mathcal{C}^2(\mathbb{R}^3)$ using a Parseval frame, then

$$\|f - f_N\|^2 \le C \sum_{m>N} m^{-2} \le C\, N^{-1}.$$

This shows that, in the weaker sense described above, N^{-1} is the optimal error approximation rate, as it was indicated in the Sec. 1.

Proof of Theorem 5.1. Our proof follows very closely the proof of Thm. 3 in [9]. We will mainly emphasize the modifications needed for $D = 3$ and refer to [9] for more detail about the argument.

Let g be a smooth and nonnegative bivariate function with compact support in $[0, 2\pi] \times [0, \pi]$. For scalars A and $m(A, \delta)$ to be determined, let

$$g_{i,j,m}(t_1, t_2) = A\, m^{-2} g(mt_1 - 2\pi i, mt_2 - \pi j), \quad i, j = 0, 1, \ldots, m-1.$$

Notice that $\|g_{i,j,m}\|_{C^2} = A\, \|g\|_{C^2}$ and $\|g_{i,j,m}\|_{L^1} = A\, m^{-4} \|g\|_{L^1}$. We introduce a spherical coordinates (ρ, θ, ϕ) with origin in $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. For $\rho_0 = \frac{1}{4}$, set

$$\psi_{i,j,m} = \chi_{\{\rho \le \rho_0\}} - \chi_{\{\rho \le g_{i,j,m} + \rho_0\}}, \quad , i, j = 0, 1, \ldots, m-1.$$

Hence, we define the radius functions

$$r_\xi = \frac{1}{4} + \sum_{i,j=1}^{m} \xi_{i,j}\, g_{i,j,m}, \quad \xi_{i,j} \in \{0, 1\}$$

and the corresponding functions

$$f_\xi = \chi_{\{\rho \le \rho_0\}} + \sum_{i,j=1}^{m} \xi_{i,j}\, \psi_{i,j,m}, \quad \xi_{i,j} \in \{0, 1\}.$$

Similar to the 2D argument, the functions $\psi_{i,j,m}$ are bulges around the sphere of radius ρ_0 and have disjoint support; each f_ξ is the indicator function of the the sphere of radius ρ_0 plus some addition bulges. Using the fact that g is bounded and nonnegative, a direct calculation shows that

$$\|\psi_{i,j,m}\|_{L^2}^2 \simeq \|g_{i,j,m}\|_{L^1} = A\, m^{-4} \|g\|_{L^1},$$

[3] For simplicity, we consider only the case $\alpha = 2$, which is what is needed in this paper. A similar proof works with $\mathcal{C}^\alpha(\mathbb{R}^3)$ and yields $p = 4/(\alpha + 2)$.

and, for each radius function r_ξ,

$$\|r_\xi\|_{C^2} \le \|g_{i,j,m}\|_{C^2} = A\,\|g\|_{C^2}.$$

Hence, as in [9], the hypercube embedding is achieved whenever $A \le C/\|g\|_{C^2}$.

Now, whenever $A \le C/\|g\|_{C^2}$, the sidelength $\Delta = \|\psi_{i,j,m}\|_{L^2}$ of the hypercubes satisfies:

$$\|\psi_{i,j,m}\|_{L^2}^2 = \Delta^2 \simeq \|g_{i,j,m}\|_{L^1} = A\,m^{-4}\|g\|_{L^1} \le C\,m^{-4}\frac{\|g\|_{L^1}}{\|g\|_{C^2}}.$$

Hence, setting

$$m(\delta) = \lfloor \Big(\frac{\delta^2}{C}\frac{\|g\|_{C^2}}{\|g\|_{L^1}}\Big)^{-\frac{1}{4}} \rfloor, \quad A(\delta, C) = \delta^2 m^4/\|g\|_{L^1},$$

it follows that $A \le C/\|g\|_{C^2}$ and $\Delta \simeq \delta$, which shows that the hypercube embedding is satisfied with sidelength $\Delta \simeq \delta$ and dimension $M = m^2(\delta)$. The dimension of the hypercube obeys

$$M = m^2(\delta) \ge K\,C^{1/2}\,\delta^{-1},$$

for all $0 < \delta < \delta_0$, where δ_0 is the solution of

$$\left(\frac{\delta_0^2}{C}\frac{\|g\|_{C^2}}{\|g\|_{L^1}}\right)^{-1/2}, \quad K = \frac{1}{2}\left(\frac{\|g\|_{C^2}}{\|g\|_{L^1}}\right)^{-1/2}.$$

Since $\Delta \simeq \delta$, it follows by the observations above that there is a $c > 0$ such that $M(\Delta_k) \ge c\,\Delta_k^{-1}$, for a sequence $(\Delta_k) \to 0$.

5.2. Extensions and Modified Construction. There is an alternative construction of 3D smooth Parseval frames of shearlets which was found by the authors during the revision of the manuscript and will be briefly sketched below. Similar to Sec. 2, whose notation we adopt here, we consider three systems of shearlets associated with the pyramidal regions $\mathcal{P}_d$, $d = 1, 2, 3$. In this case, however, we consider affine-like systems of the form

$$\{\psi_{j,\ell,k}^{(d)} = |\det A_{(d)}|^{j/2}\,\psi^{(d)}(B_{(d)}^{[\ell]}\,A_{(d)}^j x - k) : j \ge 0, -2^j \le \ell_1, \ell_2 \le 2^j, k \in \mathbb{Z}^3\}, \tag{5.3}$$

where

$$\widehat{\psi}^{(d)}(\xi_1, \xi_2, \xi_3) = U(\xi_d)\,V_{(d)}(\xi_1, \xi_2, \xi_3),$$

and $U \in C_c^\infty(\mathbb{R})$ has support in $[-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and it satisfies $\sum_{j\ge 0} |U(2^{-2j}\xi_d)|^2 = 1$, for $|\xi_d| \ge \frac{1}{8}$. Notice that the functions (5.3) have exactly the same frequency support as the functions (2.7). Similar to Sec. 2, the issue is how to combine the three systems of shearlets (5.3) for $d = 1, 2, 3$ in order to obtain a Parseval frame. If we directly combine these three systems, in the Fourier domain, each ξ is covered exactly by two directional windows $V_{j,\ell,d}(\xi) = V_{(d)}(\xi A_{(d)}^{-j}\,B_{(d)}^{-[\ell]})$ and the sum of the squares of such windows is exactly one, except for the windows overlapping the boundaries of the pyramidal regions. To enforce the partition of unity and obtain a Parseval frame, we borrow the following clever idea from [35]. We redefine the directional windows as

$$\widetilde{V}_{j,\ell,d}(\xi) = \frac{V_{j,\ell,d}(\xi)}{\sqrt{\sum_{\ell',d'} V_{j,\ell',d'}^2(\xi)}}.$$

Due to the properties of V, the denominator of this expression is 1 for all indices ℓ such that $\widetilde{V}_{j,\ell,d}$ is supported away from the boundary surfaces of the pyramidal regions; for ℓ corresponding to the boundaries,

the sum at the denominator reduces to a sum of the few terms having support near the boundary regions. This construction ensures that we obtain a smooth partition of unity and that the resulting Parseval frame of shearlets is a smooth Parseval frame of $L^2(\mathbb{R}^3)$. Since frequency support and regularity conditions are the same as those of the shearlet system from Sec. 2, it is clear that the sparsity result discussed in this paper carries over to this modified Parseval frame of shearlets. It is also clear that a similar construction can be used to obtain a smooth Parseval frame of shearlets for $L^2(\mathbb{R}^2)$ as it is needed to derive the 2D sparsity result presented in [14].

Finally, it useful to make a comment about potential extensions of our result to more general objects. We have shown that the Parseval frame of 3D shearlets provides (nearly) optimally sparse approximations for C^2 regular functions of 3 variables containing discontinuities along along C^2 boundaries. This class of functions provides a simplified model for many objects typically found in applications. However, for a more realistic model one should consider the situation of piecewise smooth boundaries. Based on our previous work in the 2-dimensional case and preliminary observations, we expect that Theorem 3.2 can be extended to the situation where the surfaces of discontinuity are not simply C^2 but piecewise C^2. This extension goes beyond the scope of this paper and will be addressed elsewhere.

Acknowledgments. The authors acknowledge support from NSF grant DMS 1008900/1008907; the second author also acknowledges support from NSF grant DMS 1005799. The authors thank the anonymous reviewers for several very valuable comments and suggestions which have improved the quality of the paper.

REFERENCES

[1] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms*, SIAM Multiscale Model. Simul. **5** (2006), 861–899.

[2] 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., **57** (2004), pp. 219–266.

[3] V. Chandrasekaran, M.B. Wakin, D. Baron, and R.G. Baraniuk, *Representation and Compression of Multidimensional Piecewise Functions Using Surflets* IEEE Transactions on Information Theory **55**(1) (2009), 374–400.

[4] F. Colonna, G. Easley, K. Guo, and D. Labate, *Radon Transform Inversion using the Shearlet Representation*, Appl. Comput. Harmon. Anal. **29**(2)2 (2010), 232–250.

[5] S. Dahlke, G. Steidl, and G. Teschke, *The continuous shearlet transform in arbitrary space dimensions*, J. Fourier Anal. Appl. **16** (2010), 340–364.

[6] S. Dahlke, G. Steidl and G. Teschke, *Shearlet Coorbit Spaces: Compactly Supported Shearlets, Traces, and Embedding*, to appear in J. Fourier Anal. Appl. (2011)

[7] R. DeVore, *Nonlinear Approximation*, Acta Numerica **7** (1998), 51–150.

[8] M. N. Do and M. Vetterli, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans. Image Process. **14** (2005), 2091–2106.

[9] D. L. Donoho, *Sparse components of images and optimal atomic decomposition*, Constr. Approx. **17**(2001), 353–382.

[10] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, *Data compression and harmonic analysis*, IEEE Trans. Inform. Th. **44**(1998), 2435–2476.

[11] G. R. Easley, D. Labate, F. Colonna, *Shearlet-Based Total Variation Diffusion for Denoising*, IEEE Trans. Image Proc. **18** (2) (2009), 260–268.

[12] G. R. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal., **25** (1) (2008), 25–46.

[13] 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), pp. 189–201.

[14] K. Guo, D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, Siam J. Math. Anal., **9** (2007), pp. 298–318

[15] K. Guo, and D. Labate, *Characterization and analysis of edges using the continuous shearlet transform*, SIAM Journal on Imaging Sciences **2**(2009), 959–986.

[16] K. Guo, and D. Labate, *Analysis and Detection of Surface Discontinuities using the 3D Continuous Shearlet Transform*, Appl. Comput. Harmon. Anal. **30** (2010), 231–241.

[17] K. Guo, and D. Labate, *The Construction of Smooth Parseval Frames of Shearlets*, preprint (2011).

[18] K. Guo, W.-Q Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations*, Electron. Res. Announc. Amer. Math. Soc., **10** (2004), pp. 78–87

[19] K. Guo, W-Q. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations and their MRA properties*, Appl. Computat. Harmon. Anal. **20**(2006), pp. 231–249.

[20] G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.

[21] G. Kutyniok and W.-Q Lim, *Compactly Supported Shearlets are Optimally Sparse*, J. Approx. Theory **163** (2011), 1564-1589.

[22] G. Kutyniok, M. Shahram, and D. L. Donoho. *Development of a Digital Shearlet Transform Based on Pseudo-Polar FFT*, Wavelets XIII (San Diego, CA, 2009), 74460B-1 - 74460B-13, SPIE Proc. 7446, SPIE, Bellingham, WA, 2009.

[23] G. Kutyniok, M. Shahram, and X. Zhuang. *ShearLab: A Rational Design of a Digital Parabolic Scaling Algorithm*, preprint

[24] J. Lemvig, *Optimally Sparse Approximations of Functions in $L^2(\mathbb{R}^3)$ with C^α Singularities using Shearlet*, Oberwolfach Report **44** (2010)

[25] E. Le Pennec, and S. Mallat, *Sparse geometric image representations with bandelets*, IEEE Trans. Image Process. **14**(2005), pp. 423–438.

[26] Y. Lu, and M. N. Do, *Multidimensional Directional Filter Banks and Surfacelets* IEEE Transactions on Image Processing, **16**(4) (2007), 918–931.

[27] S. Mallat, *Geometrical grouplets* Applied and Computational Harmonic Analysis, **26**(2) (2009), 161–180.

[28] F. Natterer, and F. Wübbeling, Mathematical Methods in Image Reconstruction, SIAM Monographs on Mathematical Modeling and Computation, Philadelphia, 2001.

[29] P. S. Negi and D. Labate, *3D Discrete Shearlet Transform and Video Processing*, to appear in IEEE Trans. Image Process. (2012).

[30] V. M. Patel, G. R. Easley and D. M. Healy, Jr., *Shearlet-Based Deconvolution*, IEEE Transactions on Image Processing **18** (2009), 2673–2685.

[31] G. Peyre and S. Mallat, *Surface compression with geometric bandelets*, ACM Transactions on Graphics **24**(3) (2005), 601–608.

[32] G. Peyre and S. Mallat, *Orthogonal Bandlet Bases for Geometric Images Approximation*, Communications on Pure and Applied Mathematics **61**(9) (2008), 1173–1212.

[33] E.M. Stein, G. Weiss, Introduction to Fourier Analysis on Euclidean Spaces, Princeton University Press, Princeton, NJ (1970).

[34] S. Yi, D. Labate, G. R. Easley, and H. Krim, *A Shearlet approach to Edge Analysis and Detection*, IEEE Trans. Image Process **18**(5) (2009), 929–941.

[35] L. Ying, L. Demanet, E. Candes, *3D Discrete Curvelet Transform*, Wavelet XI, Proc. SPIE San Diego, July 2005.