Noname manuscript No.
(will be inserted by the editor)

Optimal Recovery of 3D X-Ray Tomographic Data via Shearlet Decomposition

Kanghui Guo · Demetrio Labate

the date of receipt and acceptance should be inserted later

Abstract This paper introduces a new decomposition of the 3D X-ray transform based on the shearlet representation, a multiscale directional representation which is optimally efficient in handling 3D data containing edge singularities. Using this decomposition, we derive a highly effective reconstruction algorithm yielding a near-optimal rate of convergence in estimating piecewise smooth objects from 3D X-ray tomographic data which are corrupted by white Gaussian noise. This algorithm is achieved by applying a thresholding scheme on the 3D shearlet transform coefficients of the noisy data which, for a given noise level ε, can be tuned so that the estimator attains the essentially optimal mean square error rate $O(\log(\varepsilon^{-1})\varepsilon^{2/3})$, as $\varepsilon \to 0$. This is the first published result to achieve this type of error estimate, outperforming methods based on Wavelet-Vaguelettes decomposition and on SVD, which can only achieve MSE rates of $O(\varepsilon^{1/2})$ and $O(\varepsilon^{1/3})$, respectively.

Keywords computed tomography · inverse problems · Radon transform · shearlets · X-ray tomography · X-ray transform · wavelets

Mathematics Subject Classification (2000) 42C15 · 42C40

1 Introduction

The 3D X-ray transform is the mathematical model underlying computed tomography (CT) and similar computational methods used in medical imaging (diagnostic radiology) and in industrial nondestructive testing (quality control) to determine the structural properties of 3-dimensional objects from their projected information. In transmission tomography, for example, a solid body is scanned by a narrowly

Kanghui Guo
Department of Mathematics, Missouri State University, Springfield,Missouri 65804, USA E-mail: KanghuiGuo@MissouriState.edu

Demetrio Labate
Department of Mathematics, University of Houston, Houston, TX 77204, USA E-mail: dlabate@math.uh.edu

focused X-ray beam whose intensity loss is recorded by a detector. Hence, letting $f(x)$ be the X-ray attenuation coefficient of the body at the point x and denoting by I_0 the initial intensity of the X-ray beam, the fractional decrease of the intensity I received at the detector can be modelled as

$$\frac{I}{I_0} = \exp\left\{ - \int_L f(y)\, dy \right\},$$

where the line integral is define along the line L of propagation of the X-ray beam. By taking the natural logarithm, we obtain the *projection*

$$-\ln\left(\frac{I}{I_0}\right) = \int_L f(y)\, dy.$$

It is customarily to parametrize the lines of integration as $L = L(\Theta, x)$, where $\Theta \in S^2$ and $x \in \mathbb{R}^3$ correspond to the orientation and point of intersection of the line L, respectively. This yields the following classical formulation of the *X-ray transform* P of f:

$$P(\Theta, x)f = \int_{\mathbb{R}} f(t\Theta + x)\, dt,$$

which is also called the projection of f onto $\Theta^\perp$. Notice that this definition is different from the 3-dimensional Radon transform, which maps a function on $\mathbb{R}^3$ into the sets of its integrals over planes in $\mathbb{R}^3$, rather than into the set of its line integrals. However, the two transforms are equivalent in dimension $D = 2$ [21].

For both the Radon transform and the X-ray transform, the problem of interest is the reconstruction of f from its projected information and this calls formally for the inversion of the transforms. This problem has a formal solution and, in the case of the 2D Radon transform, an inversion formula was proposed by J. Radon already in 1917 [24]. However, the inversion of the Radon and X-ray transforms is an ill-posed problem whose computation is very sensitive to small perturbations in the data. Indeed, in practical situations, the projected information is typically known on a discrete set only, with a limited accuracy, and is corrupted by noise, so that the inversion process requires an appropriate regularization to accurately recover the unknown function f without blowing up the noise during reconstruction.

Starting from the rediscovery of the Radon transform in the 60's aimed at its application to computed tomography for medical imaging [4,5], several methods have been introduced to regularize the inverse problem associated with the Radon and X-ray transforms, including Fourier methods, backprojection and singular value decomposition [21]. The main drawback of all these methods is that they yield reconstructions where the high frequency features of the data are smoothed away, with the result that the reconstructed images appeared to be blurred versions of the original ones. While a number of heuristic solutions have been introduced to deal with the phenomenon of blurring, only the approaches recently introduced by Candès and Donoho in [2] and by the authors in [3,8] offer the capability to reconstruction 2D images from their noisy 2D Radon projections with minimal loss of high-frequency information and with a precise assessment of the method performance. These new methods exploit the properties of curvelets and shearlets, a new generation of multiscale representation systems which are especially designed to handle anisotropic information with optimal efficiency.

Despite these valuable contributions, no similar results are currently known in the 3D setting, which is indeed the case of major interest in applications. The major difficulties in extending these improved reconstruction methods to the case of noisy 3D X-ray data come from: (i) the need to extend the analysis of the optimal approximation properties of curvelets/shearlets to the 3D setting; (ii) the need to establish the optimal theoretical MSE rate in the 3D setting for the type of data under consideration. Both challenges are addressed in this paper, leading to the a new highly efficient approach for the regularized reconstruction of noisy 3D X-ray data. Specifically, our method employs the framework of the multidimensional shearlet representation introduced by the authors in [12,14] to obtain a new decomposition of the 3D X-ray transform. Taking advantage of the special ability of 3D shearlets to sparsely represent piecewise smooth data, we derive a highly effective algorithm for the reconstruction from noisy X-ray data whose error rate is near-optimal. This algorithm is based on a thresholding scheme on the noisy shearlet coefficients associated with the decomposition of the 3D X-ray transform. For a given noise level ε, the proposed thresholding scheme can be tuned so that the estimator will attain the essentially optimal mean square error (MSE) $O(\log(\varepsilon^{-1})\varepsilon^{2/3})$, as $\varepsilon \to 0$. Our result is the 3D analogue of a similar 2D estimate derived by the authors in [3] and by Candès and Donoho in [2], and is the first published result to yield an essentially optimal MSE rate for the recovery of noisy 3D x-ray data.

1.1 Background and Motivation

It is useful to briefly recall some background on the Wavelet-Vaguelette decomposition (WVD), introduced by Donoho [6]. This method applies a collection of functionals called *vaguelettes* to simultaneously invert an operator and compute the wavelet coefficients of the noise-corrupted data. Hence the unknown function is estimated by first applying a nonlinear shrinkage to the wavelet coefficients and then inverting the wavelet transform. One major benefit of the Wavelet-Vaguelette decomposition is that it allows one to select a representation that well approximate the space of solutions and to exploit the estimating capabilities of this representation. This is in contrast to the Singular Value Decomposition (SVD) that uses basis functions depending solely on the operator whose inversion one attempts to compute. Since its appearance, the WVD strategy has received lot of attention and was applied to a number of inverse problems including the inversion of the Radon transform (e.g., [20,19,22]). However, as it was observed in [2], even though the WVD strategy outperforms regularization methods based on SVD and other traditional methods, it falls short from being optimal in terms of estimation capabilities, in general.

Indeed, let us consider the problem of recovering a piecewise smooth two-dimensional image f from the noisy Radon data

$$Y = Rf + \varepsilon W,$$

where εW is white Gaussian noise, and ε measures the noise level. Then an inversion based on the WVD approach yields a Mean Squared Error (MSE) that is bounded, within a logarithmic factor, by $O(\varepsilon^{2/3})$ as $\varepsilon \to 0$. This is better than the MSE rate of $O(\varepsilon^{1/2})$, as $\varepsilon \to 0$, which is achieved when the inversion techniques are SVD-based [2]. However, if one replaces the wavelet system used in the WVD with the curvelet or the shearlet systems, which are more efficient than wavelets in dealing with the class of bivariate piecewise smooth functions, then it is proved in [2, 3] that it possible to recover the function f with an MSE rate

$$O(\log(\varepsilon^{-1})\varepsilon^{4/5}) \quad \text{as } \varepsilon \to 0.$$

This rate is essentially optimal for this class of functions.

In this paper, we consider the analogous 3D problem and investigate the task of recovering a piecewise smooth 3D image f from the noisy X-ray data

$$Y = Pf + \varepsilon W,$$

where εW is white Gaussian noise, and ε measures the noise level. In this setting, a regularized inversion of the 3D X-ray transform based on the WVD approach yields a Mean Squared Error (MSE) that is bounded, within a logarithmic factor, by $O(\varepsilon^{1/2})$ as $\varepsilon \to 0$. By contrast, we prove that it possible to improve the reconstruction performance by using a 3D systems of shearlets to decompose the 3D X-ray transform. In this case, the function f is recovered with an MSE rate

$$O(\log(\varepsilon^{-1})\varepsilon^{2/3}) \quad \text{as } \varepsilon \to 0.$$

Similar to the 2D case, the method that we propose adapts the basic ideas of the WVD framework. In order to establish the new MSE estimation rate, our argument relies critically on results recently derived from the authors showing that the 3D shearlet representation exhibits essentially optimal approximation properties in the class of piecewise functions of 3 variables [12, 13]. The optimality of our MSE estimation rate is also proved in this paper using a new argument extending to the 3D setting a similar 2D result from [2].

1.2 Paper Organization

The paper is organized as follows. Section 2 provides the background on the 3D shearlet representation. Section 3 develops a new decomposition of the 3D X-ray transform based on the shearlet representation. This decomposition is the analogue of our 2D decomposition from [3]. Section 4 contains the main original contributions and main results of the paper: the new shearlet-based algorithm for the recovery of 3D images from X-ray data which are corrupted by additive Gaussian noise; the analysis of the method's performance, showing that the method is nearly optimal in the class of piecewise smooth functions of three variables.

2 The Shearlet Representation

The shearlet representation, originally derived from the framework of wavelets with composite dilations [15, 16], 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 their ability to deal with directionality and anisotropy, shearlets capture the geometric content of multivariate functions and data much more efficiently than using wavelets or other traditional methods. These properties have made shearlets very successful in applications such as image and video denoising [9,10,23], deconvolution [25], edge analsysis and detection [11,26] and sparse decompositions [17].

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. The construction given below is similar to the so-called digital curvelets in [1].

To define such systems, let b be a univariate function such that $\hat{b} \in C^\infty$, $0 \le \hat{b} \le 1$, $\hat{b} = 1$ on $[-\frac{1}{16}, \frac{1}{16}]$ and $\hat{\phi} = 0$ outside the interval $[-\frac{1}{8}, \frac{1}{8}]$. That is, b 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 ϕ by

$$\widehat{\phi}(\xi) = \widehat{\phi}(\xi_1, \xi_2, \xi_3) = \hat{b}(\xi_1)\, \hat{b}(\xi_2)\, \hat{b}(\xi_3) \tag{1}$$

and $W(\xi) = \sqrt{|\widehat{\phi}(2^{-2}\xi)|^2 - |\widehat{\phi}(\xi)|^2}$. It follows that

$$|\widehat{\phi}(\xi)|^2 + \sum_{j \ge 0} |W(2^{-2j}\xi)|^2 = 1 \text{ for } \xi \in \mathbb{R}^3.$$

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

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

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

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

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

$$|v(u-1)|^2 + |v(u)|^2 + |v(u+1)|^2 = 1 \qquad \text{for } |u| \le 1. \tag{4}$$

In addition, we will assume that $v(0) = 1$ and that $v^{(n)}(0) = 0$ for all $n \ge 1$. It was shown in [13] that there are several examples of functions satisfying these properties.

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

$$\{\psi^{(d)}_{j,\ell,k} : j \ge 0, -2^j \le \ell_1, \ell_2 \le 2^j, k \in \mathbb{Z}^3\}, \tag{5}$$

where

$$\hat{\psi}^{(d)}_{j,\ell,k}(\xi) = |\det A_{(d)}|^{-j/2}\, W(2^{-2j}\xi)\, V_{(d)}(\xi A^{-j}_{(d)} B^{[-\ell]}_{(d)})\, e^{2\pi i \xi A^{-j}_{(d)} B^{[-\ell]}_{(d)} k}, \tag{6}$$

$V_{(1)}(\xi_1,\xi_2,\xi_3) = v(\frac{\xi_2}{\xi_1})v(\frac{\xi_3}{\xi_1})$, $V_{(2)}(\xi_1,\xi_2,\xi_3) = v(\frac{\xi_1}{\xi_2})v(\frac{\xi_3}{\xi_2})$, $V_{(3)}(\xi_1,\xi_2,\xi_3) = v(\frac{\xi_1}{\xi_3})v(\frac{\xi_2}{\xi_3})$, 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 *shearing 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 (5) 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

$$\hat{\psi}^{(1)}_{j,\ell_1,\ell_2,k}(\xi) = 2^{-2j}\, W(2^{-2j}\xi)\, v\Big(2^j \tfrac{\xi_2}{\xi_1} - \ell_1\Big)\, v\Big(2^j \tfrac{\xi_3}{\xi_1} - \ell_2\Big)\, e^{2\pi i \xi A^{-j}_{(1)} B^{[-\ell_1,-\ell_2]}_{(1)} k}, \quad (7)$$

showing that their supports are contained inside the trapezoidal regions

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

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

A simple computation shows that the elements of the shearlets systems (6) can be written in space domain as

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

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

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

The functions $\psi^{(d)}_{j,\ell}$ depend very little on j, ℓ. Indeed, thanks to support and regularity conditions on W and $V_{(d)}$, one can show [14] that for each $\gamma \in \mathbb{N}^3$ and each $N \ge 0$ there is a constant $C_{\gamma,N,d} > 0$, independent of j, ℓ, such that,

$$\left|\partial^\gamma_x \psi^{(d)}_{j,\ell}(x)\right| \le C_{\gamma,N,d}\,(1 + |x|)^{-N}. \quad (8)$$

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 from 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. To ensure the regularity and decay of the system, the elements of the shearlet systems overlapping the boundaries of

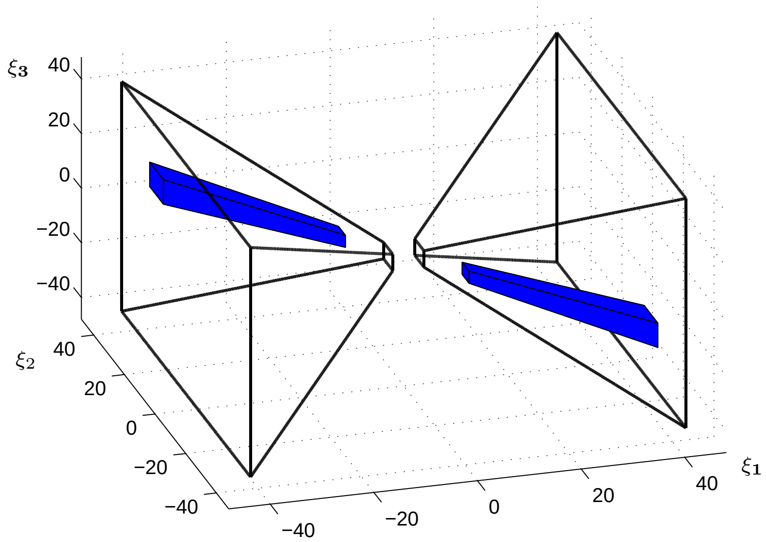

Fig. 1 *Frequency support of a representative shearlet function $\psi^{(1)}_{j,\ell,k}$, inside the pyramidal region $\mathcal{P}_1$. The orientation of the support region is controlled by $\ell = (\ell_1, \ell_2)$; its shape is becoming more elongated as j increases ($j = 4$ in this plot)*

the pyramidal regions $\mathcal{P}_d$ in the Fourier domain are modified. More precisely, we define the *3D shearlet systems for* $L^2(\mathbb{R}^3)$ as the collections

$$\Big\{\phi_k : k \in \mathbb{Z}^3\Big\} \bigcup \Big\{\widetilde{\psi}_{j,\ell,k,d} : j \geq 0, |\ell_1| < 2^j, |\ell_2| \leq 2^j, k \in \mathbb{Z}^3, d = 1,2,3\Big\}$$
$$\bigcup \Big\{\widetilde{\psi}_{j,\ell,k} : j \geq 0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3\Big\} \tag{9}$$

consisting of:

- the *coarse-scale shearlets* $\{\phi_k = \phi(\cdot - k) : k \in \mathbb{Z}^3\}$, where ϕ is given by (1);
- the *interior shearlets* $\{\widetilde{\psi}_{j,\ell,k,d} = \psi^{(d)}_{j,\ell,k} : j \geq 0, |\ell_1||\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$, where $\psi^{(d)}_{j,\ell,k}$ are given by (6);
- the *boundary shearlets* $\{\widetilde{\psi}_{j,\ell,k,d} : j \geq 0, |\ell_1| < 2^j, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$ and $\{\widetilde{\psi}_{j,\ell,k} : j \geq 0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3\}$, 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 $\ell_1, \ell_2 = \pm 2^j$. We refer to [**?**,14] for detail. Here it suffices to observe that the boundary shearlets are both compactly supported and smooth in the frequency domain.

For brevity, in the following it will be convenient to denote the system of shearlets (9) using the compact notation:

$$\{s_\mu, \ \mu \in M\}, \tag{10}$$

where $M = \mathbb{Z}^3 \cup M_I \cup M_B$, $s_\mu = \phi_\mu$ if $\mu \in \mathbb{Z}^3$ and $s_\mu = \widetilde{\psi}_\mu$ if $\mu \in M_I \cup M_B$, and M_I, M_B are the indices associated with the interior shearlets and the boundary shearlets, respectively, given by

- $M_I = \{\mu = (j, \ell_1, \ell_2, k, d) : j \geq 0, |\ell_1|\&|\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ (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 [14].

Theorem 1 *The 3D shearlet system* (10) *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.*

3 3D X-Ray Transform Inversion via Shearlet Representation

In this section, we derive a formula for the decomposition of the 3D X-ray transform based on the shearlet representation. This construction is similar to the 2D construction in [3].

3.1 Companion Representation

In order to derive our new decomposition, it is useful to introduce the following companion representation of the 3D shearlet system.

Definition 1 For a rational number α and $f \in C^\infty(\mathbb{R}^3)$, the Riesz potential I^α in $\mathbb{R}^n$ is defined by $\widehat{I^\alpha f}(\xi) = |\xi|^{-\alpha} \hat{f}(\xi), \;\; \alpha < n$.

Definition 2 For $\{s_\mu : \mu \in M\}$ given by (10), the companion shearlet representation is the set $\{s_\mu^+ = 2^{-j} I^{-\frac{1}{2}} s_\mu : \mu \in M\}$. In particular, we use the notation $\widetilde{\psi}_\mu^+ = 2^{-j} I^{-\frac{1}{2}} \widetilde{\psi}_\mu$, for $\mu \in M_I \cup M_B$, and $\phi_\mu^+ = I^{-\frac{1}{2}} \phi_\mu$, for $\mu \in \mathbb{Z}^3$.

It is easy to verify that the functions $\{s_\mu^+ : \mu \in M\}$ are smooth and compactly supported in the frequency domain. In addition, since each element of the fine-scale shearlet system $\widetilde{\psi}_\mu$, $\mu \in M_I \cup M_B$, has frequency support inside the compact region $C_j = [-2^{2j-1}, 2^{2j-1}]^3 \setminus [-2^{2j-4}, 2^{2j-4}]^3$, it follows that, in this region, $2^{-j}|\xi|^{1/2} \simeq 1$ and, thus, $\|\widetilde{\psi}_\mu^+\| \simeq \|\widetilde{\psi}_\mu\|$. This is the key observation which is used in the following result.

Theorem 2 *The system* $\{\widetilde{\psi}_\mu^+\}_{\mu \in M_I \cup M_B}$ *is a frame for* $L^2(\mathbb{R}^3 \setminus [-\frac{1}{8}, \frac{1}{8}]^2)^\vee$. *That is, there are positive constants* A *and* B, *with* $0 < A \leq B$, *such that*

$$A\,\|f\|_2 \leq \left\| \sum_\mu \langle f, \widetilde{\psi}_\mu^+ \rangle \right\|_{\ell^2} \leq B \|f\|_2,$$

for all functions f *such that* $\operatorname{supp} \hat{f} \subset \mathbb{R}^3 \setminus [-\frac{1}{8}, \frac{1}{8}]^3$.

Notice that the larger system $\{s_\mu^+ : \mu \in M\}$ is not a frame for $L^2(\mathbb{R}^3)$ since the lower frame bound condition is not satisfied. To enlarge $\{\widetilde{\psi}_\mu^+\}_{\mu \in M}$ and make it into a frame for the whole space $L^2(\mathbb{R}^3)$, one can include a "coarse scale" system of the form $\{\varphi(x - k) : k \in \mathbb{Z}^3\}$ similar to what is done for the original shearlet system.

Proof. To show that the system $\{\widetilde{\psi}_\mu^+\}_{\mu\in M_I\cup M_B}$ is a frame, we introduce the following notation. For $j\geq 0$, let M_j be the set of indices $\mu\in M_I\cup M_B$ such that j is fixed. Explicitly, $M_j=\{\mu=(j,\ell_1,\ell_2,k,d):|\ell_1|<2^j,|\ell_2|\leq 2^j,k\in\mathbb{Z}^3,d=1,2,3\}\cup\{\mu=(j,\ell_1,\ell_2,k):|\ell_1|,|\ell_2|=2^j,k\in\mathbb{Z}^3\}$. Recall that, for each $j\geq 0$, a function $\widetilde{\psi}_\mu$, $\mu\in M_j$, has frequency support contained in the set C_j which is given by (2). Also recall that $\bigcup_{j\geq 0}C_j=\mathbb{R}^3\setminus[-\frac{1}{8},\frac{1}{8}]^3$. For brevity, we adopt the notation $\mathcal{C}=\mathbb{R}^3\setminus[-\frac{1}{8},\frac{1}{8}]^3$.

Let $\hat{w}_j$ be a smooth window function which is supported in $C_{j-1}\cup C_j\cup C_{j+1}$ and is equal 1 on the set C_j. It follows that $\hat{w}_j\,W_j=W_j$ and that $w_j*g=g$ for each function g such that $\hat{g}$ is supported on the set C_j. Hence, using the fact that $\operatorname{supp}(\widetilde{\psi}_\mu)^\wedge\subset C_j$ when $\mu\in M_j$, it follows that for each $f\in L^2(\mathcal{C})^\vee$:

$$\begin{aligned}\sum_{\mu\in M_j}|\langle f,\widetilde{\psi}_\mu^+\rangle|^2&=\sum_{\mu\in M_j}|\langle f,2^{-j}\,I^{-\frac12}\widetilde{\psi}_\mu\rangle|^2\\&=\sum_{\mu\in M_j}|\langle w_j*f,2^{-j}\,I^{-\frac12}\widetilde{\psi}_\mu\rangle|^2\\&=\sum_{\mu\in M_j}|\langle 2^{-j}\,I^{-\frac12}(w_j*f),\widetilde{\psi}_\mu\rangle|^2\\&=\sum_{\mu\in M_j}|\langle h_j,\widetilde{\psi}_\mu\rangle|^2,\end{aligned}$$

where $h_j=2^{-j}\,I^{-\frac12}(w_j*f)$. Notice that the smoothness assumption on w_j guarantees that h_j is well defined. Furthermore $h_j\in L^2(\mathcal{C})^\vee$.

Using the properties of the shearlet system (in particular, (**??**)), a calculation similar to Theorem 3.2 in [14] gives that, for any $h\in L^2(\mathcal{C})^\vee$ we have the following equalities:

$$\begin{aligned}&\sum_{\mu\in M_j}|\langle h,\widetilde{\psi}_\mu\rangle|^2\\&=\sum_{d=1}^{3}\sum_{|\ell_1|\leq 2^j}\sum_{|\ell_2|<2^j}\sum_{k\in\mathbb{Z}^3}|\langle h,\widetilde{\psi}_{j,\ell_1,\ell_2,k,d}\rangle|^2+\sum_{j\geq 0}\sum_{\ell_1,\ell_2=\pm 2^j}\sum_{k\in\mathbb{Z}^3}|\langle h,\widetilde{\psi}_{j,\ell_1,\ell_2,k}\rangle|^2\\&=\int_{\mathcal{C}}|\hat{h}(\xi)|^2\,|W(2^{-2j}\xi)|^2\left(\sum_{|\ell_1|,|\ell_2|\leq 2^j}|v(2^j\frac{\xi_2}{\xi_1}-\ell_1)|^2|v(2^j\frac{\xi_3}{\xi_1}-\ell_2)|^2\chi_{\mathcal{P}_1}(\xi)\right.\\&+\sum_{|\ell_1||\ell_2|\leq 2^j}|v(2^j\frac{\xi_1}{\xi_2}-\ell_1)|^2|v(2^j\frac{\xi_3}{\xi_2}-\ell_2)|^2\chi_{\mathcal{P}_2}(\xi)\\&\left.+\sum_{|\ell_1||\ell_2|\leq 2^j}|v(2^j\frac{\xi_1}{\xi_3}-\ell_1)|^2|v(2^j\frac{\xi_2}{\xi_3}-\ell_2)|^2\chi_{\mathcal{P}_3}(\xi)\right)d\xi\\&=\int_{\mathcal{C}}|\hat{h}(\xi)|^2\,|W(2^{-2j}\xi)|^2\,d\xi,\end{aligned}\tag{11}$$

where $W_j=W(2^{-2j}\cdot)$ is supported inside the set $C_j\subset\mathcal{C}$.

Using (11) with $h = h_j$ it follows that:

$$\sum_{\mu \in M_j} |\langle h_j, \widetilde{\psi}_\mu \rangle|^2 = \int_{\mathcal{C}} |\hat{h}_j(\xi)|^2 \, |W(2^{-2j}\xi)|^2 \, d\xi$$
$$= 2^{-2j} \int_{C_j} |\xi| \, |\hat{f}(\xi)|^2 \, |W(2^{-2j}\xi)|^2 \, d\xi.$$

Notice that, for $\xi \in C_j$, the function $2^{-2j}|\xi|$ is bounded above and below by positive constants, independently of j. Hence, from the expression above, we have that

$$\sum_{\mu \in M_j} |\langle h_j, \widetilde{\psi}_\mu \rangle|^2 \simeq \int_{C_j} |\hat{f}(\xi)|^2 \, |W(2^{-2j}\xi)|^2 \, d\xi$$

Adding up over all $j \geq 0$ and using (3) we conclude that

$$\sum_{\mu \in M} |\langle f, \widetilde{\psi}_\mu^+ \rangle|^2 = \sum_{j \geq 0} \sum_{\mu \in M_j} |\langle f, \widetilde{\psi}_\mu^+ \rangle|^2$$
$$\simeq \sum_{j \geq 0} \int_{C_j} |\hat{f}(\xi)|^2 \, |W(2^{-2j}\xi)|^2 \, d\xi$$
$$= \int_{\mathcal{C}} |\hat{f}(\xi)|^2 d\xi$$

for all $f \in L^2(\mathcal{C})^\vee$. □

3.2 Shearlet decomposition of the 3D X-ray transform

We first recall some basic properties of the X-ray transform, which will be useful in the construction of our decomposition based on the shearlet representation (cf. [21] for additional details). For $\Theta \in S^2$ and $x \in \mathbb{R}^3$, then the 3D *X-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 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 $\mathcal{T} = \{(\Theta, x) : \Theta \in S^2, x \in \Theta^\perp\}$.

The adjoint operator P^* of P is acting on functions on $\mathcal{T}$ and is defined by

$$(P^* g)(x) = \int_{S^2} g(\Theta, E_\Theta x) d\Theta,$$

where $E_\Theta x = x - (x \cdot \Theta)\Theta$ is the orthogonal projection of x on $\Theta^\perp$.

For F, G on $\mathcal{T}$, we define the inner product of F, $G \in L^2(\mathcal{T})$ by

$$[F, G] = \int_{S^2} \int_{\Theta^\perp} F(\Theta, x) \, \overline{G(\Theta, x)} dx \, d\Theta,$$

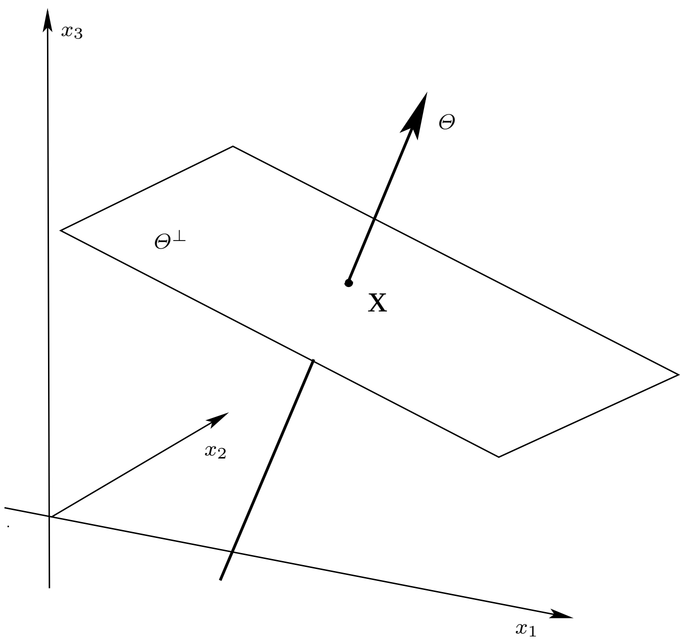

Fig. 2 *The 3D-ray transform is defined by integration over the lines through the point x with direction Θ.*

while as usual for f, g on $\mathbb{R}^3$, define the inner product of f and g by

$$\langle f, g \rangle = \int_{\mathbb{R}^3} f(x)\, \overline{g(x)} dx.$$

In this setting, we have

$$[Pf, G] = \langle f, P^* G \rangle.$$

It is useful to recall the *Fourier Slice Theorem* which establishes that following relationship between the 3D X-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 η.

For a function F on $\mathcal{T}$, we define the operator I_2^{α}, $\alpha \in \mathbb{R}$, on F by

$$\mathcal{F}_2[I_2^{\alpha} F](\Theta, \eta) = |\eta|^{-\alpha} \mathcal{F}_2[F](\Theta, \eta), \;\; \eta \in \Theta^{\perp}.$$

It follows that, for each fixed Θ, we have:

$$\mathcal{F}_2[I_2^{-\frac{1}{2}} Pf](\Theta, \eta) = |\eta|^{\frac{1}{2}} \mathcal{F}_2[Pf](\Theta, \eta) = |\eta|^{\frac{1}{2}} \hat{f}(\eta), \;\; \eta \in \Theta^{\perp}.$$

Since, for each fixed Θ,

$$\mathcal{F}_2[P(I^{-\frac{1}{2}} f)](\Theta, \eta) = \widehat{(I^{-\frac{1}{2}} f)}(\eta) = |\eta|^{\frac{1}{2}} \hat{f}(\eta), \;\; \eta \in \Theta^{\perp},$$

it follows that $I_2^{-\frac{1}{2}} Pf = PI^{-\frac{1}{2}} f$.

We will need the following formula for the 3D X-ray transform which, up to a simple change of constant (needed to make it consistent with our definition of Fourier transform), is a special case of Theorem 2.14 in [21]. For $f \in \mathcal{S}(\mathbb{R}^3)$, we have

$$f = 2^{-1} I^{-\frac{1}{2}} P^* \, I_2^{-\frac{1}{2}} Pf.$$

Let $\mathcal{P} = I_2^{-\frac{1}{2}} P$. Since the operator $I^{-\frac{1}{2}}$ is self-adjoint, the formula above yields the following *X-ray isometry condition*

$$[\mathcal{P}f, \mathcal{P}g] = 2 \, \langle f, g \rangle.$$

Using the companion shearlet representations $\{s_\mu^+ : \mu \in M\}$ from Section 3.1, we define the system $\{U_\mu : \mu \in M\}$ by the formula

$$U_\mu = \mathcal{P} s_\mu^+, \quad \mu \in M. \tag{12}$$

Using the X-ray isometry and Theorem 2, one can show that $\{U_\mu : \mu \in M\}$ is a frame sequence (that is, a frame for its span), although it is not a frame since $\{s_\mu^+ : \mu \in M\}$ is not a frame for $L^2(\mathbb{R}^3)$.

We are now ready to introduce a decomposition formula for the 3D X-ray transform based on the 3D shearlet representation. This result is in the spirit of the Wavelet-Vaguelette Decomposition [6] and extends to the 3D setting the 2D decomposition of the Radon transform from [2,3].

Theorem 3 *Let $\{s_\mu : \mu \in M\}$ be the Parseval frame of shearlets given by* (10) *and $\{U_\mu : \mu \in M\}$ be the system defined by (12). For all $f \in L^2(\mathbb{R}^3)$ the following reproducing formula holds:*

$$f = 2^{-1} \sum_\mu 2^j [Pf, U_\mu] \, s_\mu.$$

Proof. A direct computation show that:

$$\begin{aligned}
\langle f, s_\mu \rangle &= 2^{-1} [\mathcal{P}f, \mathcal{P} \, s_\mu] \\
&= 2^{-1} [I_2^{-\frac{1}{2}} \; Pf, I_2^{-\frac{1}{2}} \; Pf \; P \, s_\mu] \\
&= 2^{-1} [Pf, I_2^{-\frac{1}{2}} I_2^{-\frac{1}{2}} \; P \, s_\mu] \\
&= 2^{-1} [Pf, I_2^{-\frac{1}{2}} \; PI^{-\frac{1}{2}} \; s_\mu] \\
&= 2^{-1} [Pf, \mathcal{P}(2^j s_\mu^+)] \\
&= 2^{-1} 2^j \, [Pf, U_\mu]. \quad \square
\end{aligned}$$

4 Optimal Inversion of Noisy 3D X-Ray Data

Let us consider the classical problem of recovering an unknown function f from its noisy X-ray projections. More precisely, we assume that the observed 3D X-ray transform of f is corrupted by white Gaussian noise as:

$$Y = Pf + \varepsilon \, W, \tag{13}$$

where W is a Wiener sheet and ε is measuring the noise level. This means that each measurement $[Y, U_\mu]$ of the observed data is normally distributed with mean $[Pf, U_\mu]$ and variance $\varepsilon^2 \|U_\mu\|^2_{L^2(\mathcal{T})}$.

While the white noise model may not describe precisely the types of noise typically found in practical applications, the asymptotic theory derived from this assumption in practice has been found to lead to very acceptable results. This framework allows one to derive a theoretical assessment of the performance of the method which would be extremely complicated to handle otherwise.

In order to obtain an upper bound on the risk of the estimator, it is necessary to specify the type of functions we are considering. Following [13], we fix a constant $A > 0$ and denote by $\mathcal{M}(A)$ the class 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 $\Sigma_\alpha = \{(v,\ E_\alpha(v)),\quad v \in V_\alpha \subset \mathbb{R}^2\}$, such that $\|E_\alpha\|_{C^2(V_\alpha)} \le A$ for all α. Also, let $C^2_c([0,1]^3)$ 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 $f_0, f_1 \in C^2_c([0,1]^3)$, $B \in \mathcal{M}(A)$ and $\|f\|_{C^2} = \sum_{|\alpha|\le 2} \|D^\alpha f\|_\infty \le 1$.

Projecting the data (13) onto the frame $\{U_\mu :\ \mu \in M\}$, and rescaling, we obtain

$$\begin{aligned} y_\mu &:= 2^j\, [Y, U_\mu] \\ &= 2^j\, [Pf, U_\mu] + \varepsilon\, 2^j\, [W, U_\mu] \\ &= 2\langle f, s_\mu\rangle + \varepsilon\, 2^j\, n_\mu, \end{aligned} \tag{14}$$

where n_μ is a (non-i.i.d.) Gaussian noise with zero mean and variance $\sigma_\mu = \|\psi^+_\mu\|_2$. We observe that there are positive numbers α_1, α_2 such that $\alpha_1 \le \sigma_\mu \le \alpha_2$ for all $\mu \in M$. In order to estimate f, we need to estimate the shearlet coefficients $\langle f, \psi_\mu\rangle$, $\mu \in M$, from the data y_μ. To accomplish this, we will devise a thresholding rule, to be applied to $\{y_\mu : \mu \in M\}$, which exploits the sparsity properties of the 3D shearlet representation.

4.1 Shearlet-based decomposition of the noisy data

For the application of the shearlet representation to the estimation problem, we need to modify the shearlet system by rescaling the coarse scale system and changing the range of scales for which the directional fine-scale system is defined. Hence, for a fixed $j_0 \in N$ (which will be chosen as a function of the noise level), we let

- $M_C = \{\mu = (j,k) : j = j_0 - 1, k \in \mathbb{Z}^3\}$ (coarse-scale shearlets, replacing $\Phi(x-k)$ by $2^{\frac{3}{2}j_0}\Phi(2^{j_0}x - k)$)
- $M_I = \{\mu = (j, \ell_1, \ell_2, k, d) :\ j \ge j_0, |\ell_1| \& |\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1,2,3\}$ (interior shearlets)
- $M_B = \{\mu = (j, \ell_1, \ell_2, k, d) :\ j \ge j_0, |\ell_1| < 2^j, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3, d = 1,2,3\} \cup \{\mu = (j, \ell_1, \ell_2, k) :\ j \ge j_0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3\}$ (boundary shearlets)

so that the modified system is still a Parseval frame in $\mathbb{R}^3$.

Similar to the original shearlet system given in Theorem 1, the modified shearlet system is made of coarse and fine scale systems, with the coarse scale system now associated with the coarse scale j_0. Proceeding as above, we introduce the index set $\mathcal{M}^0 = N \cup M^0$, where $N = \mathbb{Z}^3$,

$$M^0 = \{\mu = (j, \ell, k, d) : j \geq j_0, -2^j \leq \ell_1,\ \ell_2 \leq 2^j, k \in \mathbb{Z}^3\}, d = 1, 2, 3;$$

we shall denote the new shearlet system using the compact notation $\{s_\mu : \mu \in \mathcal{M}^0\}$, where $s_\mu = \psi_\mu = \psi^{(d)}_{j,\ell,k}$ if $\mu \in M^0$ and $s_\mu = 2^{\frac{3}{2}j_0}\Phi(2^{j_0}x - \mu)$ if $\mu \in N$. For ψ_μ, $\mu \in M^0$, it is understood that the boundary elements are modified as in Theorem 1.

In our reconstruction method, the selection of the scale j_0 will depend on the noise level ε. Namely, we set $j_0 = \frac{2}{15}\log_2(\varepsilon^{-1})$. We also introduce the scale index $j_1 = \frac{1}{3}\log_2(\varepsilon^{-1})$ (so that $2^{j_0} = \varepsilon^{-\frac{2}{15}}$ and $2^{j_1} = \varepsilon^{-\frac{1}{3}}$).

Hence, depending on the noise level ε, for a given a function $f \in \mathcal{E}^2(A)$ we define the *set of significant indices* associated with the shearlet representation of f as the subset of $\mathcal{M}^0$ given by $\mathcal{N}(\varepsilon) = M_1(\varepsilon) \cup N_0(\varepsilon)$, where

$$N_0(\varepsilon) = \{\mu = k \in \mathbb{Z}^3 : \ |k| \leq 2^{2j+1}\}, \text{ and}$$

$$M_1(\varepsilon) = \{\mu = (j, \ell, k, d) : j_0 < j \leq j_1, |k| \leq 2^{2j+1}, d = 1, 2, 3\}.$$

The *significant coefficients* in the shearlet representation of f are the elements $\langle f, s_\mu \rangle$ for which μ belongs to $\mathcal{N}(\varepsilon)$.

We obtain the following result which is proved in the Appendix. Notice that the proof of this result relies on the nearly optimal approximation properties of the 3D shearlets (and does not follow from the 2D proof).

Theorem 4 *For $f \in \mathcal{E}^2(A)$, let ε denote the noise level, and $\mathcal{N}(\varepsilon)$ be the set of significant indices associated with the shearlet representation of f where f is represented as*

$$f = \sum_{\mu \in \mathcal{M}^0} \langle f, s_\mu \rangle \, s_\mu.$$

Then there exist positive constants C', C'', and C''' such that the following properties hold:

1. *The neglected shearlet coefficients $\{\langle f, s_\mu \rangle : \mu \notin \mathcal{N}(\varepsilon)\}$ satisfy:*

$$\sup_{f \in \mathcal{E}^2(A)} \sum_{\mu \notin \mathcal{N}(\varepsilon)} |\langle f, s_\mu \rangle|^2 \leq C' \varepsilon^{2/3}.$$

2. *The risk proxy satisfies:*

$$\sup_{f \in \mathcal{E}^2(A)} \sum_{\mu \in \mathcal{N}(\varepsilon)} \min(|\langle f, s_\mu \rangle|^2, 2^{2j}\varepsilon^2) \leq C'' \varepsilon^{2/3}.$$

3. *The cardinality of $\mathcal{N}(\varepsilon)$ obeys:*

$$\#\mathcal{N}(\varepsilon) \leq C''' \varepsilon^{-\frac{8}{3}}.$$

4.2 3D X-ray data recovery via shearlet thresholding

To estimate f from the noisy observations (14), we will apply the soft thresholding function $T_s(y,t) = \mathrm{sgn}(y)(|y|-t)_+$. The analysis of the estimation error follows the general framework of the wavelet shrinkage developed in [7]. Letting $\#\mathcal{N}(\varepsilon)$ be the number of significant coefficients of the shearlet representation of f, we estimate the function f by

$$\tilde{f} = \sum_{\mu \in \mathcal{M}^0} \tilde{c}_\mu \, s_\mu, \tag{15}$$

where the coefficients are obtained by the rule

$$\tilde{c}_\mu = \begin{cases} T_s(y_\mu, \varepsilon\sqrt{2\log(\#\mathcal{N}(\varepsilon))}2^j \sigma_\mu), & \mu \in \mathcal{N}(\varepsilon), \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

and $\sigma_\mu = \|s_\mu^+\|_2^2$. Notice that the terms σ_μ, $\mu \in \mathcal{M}^0$, are uniformly bounded.

Our main theorem can now be established. This result is the first published result of this type and, as will be shown in the next section, is essentially optimal for functions in $\mathcal{E}^2(A)$. It is the the 3D analogue of the 2D results obtain using curvelets and shearlets in [2, Thm.6] and in [3, Thm.4.2], respectively.

Theorem 5 *Let $f \in \mathcal{E}^2(A)$ be the solution of the problem $Y = Rf + \varepsilon W$ and let $\tilde{f}$ be the approximation to f given by the formulas (15) and (16). Then there is a constant $C > 0$ such that*

$$\sup_{\mathcal{E}^2(A)} E\|\tilde{f} - f\|_2^2 \le C \, \log(\varepsilon^{-1}) \, \varepsilon^{\frac{2}{3}}, \qquad \text{as } \varepsilon \to 0,$$

where E is the expectation operator.

Proof. For $\mu \in \mathcal{M}^0$, let $c_\mu = \langle f, s_\mu \rangle$ and $\tilde{c}_\mu$ be given by (16). By the Parseval frame property of the shearlet system $\{s_\mu : \mu \in \mathcal{M}^0\}$, it follows that

$$\| \sum_{\mu \in \mathcal{M}^0} c_\mu \, s_\mu \|_2^2 \le \sum_{\mu \in \mathcal{M}^0} |c_\mu|^2,$$

and that

$$E\|\tilde{f} - f\|_2^2 \le E \left(\sum_{\mu \in \mathcal{M}^0} |\tilde{c}_\mu - c_\mu|^2 \right). \tag{17}$$

On the other hand, by the oracle inequality [7] we have

$$E\left(\sum_{\mu \in \mathcal{N}(\varepsilon)} |\tilde{c}_\mu - c_\mu|^2 \right) \le L(\varepsilon) \left(\varepsilon^2 \sum_{\mu \in \mathcal{N}(\varepsilon)} \left(\frac{2^{2j}\sigma_\mu^2}{\#\mathcal{N}(\varepsilon)} + \min(c_\mu^2, \varepsilon^2 2^{2j} \sigma_\mu^2) \right) \right) \tag{18}$$

where $L(\varepsilon) = (1 + 2\log(\#\mathcal{N}(\varepsilon))$. Now observe that, by Theorem 4.1, there exist positive constants C_1, C_2, C_3, C_4 such that

$$\sum_{\mu\in\mathcal{N}(\varepsilon)} \min(c_\mu^2, \varepsilon^2 2^{2j}\sigma_\mu^2) \le C_1\, \varepsilon^{\frac{2}{3}},$$

$$\sum_{\mu\notin\mathcal{N}(\varepsilon)} c_\mu^2 \le C_2\, \varepsilon^{\frac{2}{3}},$$

$$\log(\#\mathcal{N}(\varepsilon)) \le C_3 \, \log(\varepsilon^{-1}).$$

$$\varepsilon^2 \sum_{\mu\in\mathcal{N}(\varepsilon)} 2^{2j} \frac{\sigma_\mu^2}{\#\mathcal{N}(\varepsilon)} \le \alpha_2^2 \varepsilon^2 \sum_{\mu\in\mathcal{N}(\varepsilon)} 2^{2j} \le C_4\, \varepsilon^2\, 2^{2j_1} = C_4 \varepsilon^{\frac{4}{3}}.$$

Thus, using these observations and equations (17) and (18), we deduce that there is a constant $C > 0$ such that

$$E\|\tilde f - f\|_2^2 \le E\left(\sum_{\mu\in\mathcal{N}(\varepsilon)} |\tilde c_\mu - c_\mu|^2\right) + \sum_{\mu\in\mathcal{N}(\varepsilon)^c} c_\mu^2 \le C\, \log(\varepsilon^{-1})\, \varepsilon^{\frac{2}{3}}. \quad \square$$

As also observed in [2], Theorem 5 remains valid if the soft thresholding operator $T_s(y,t)$ is replaced by the hard thresholding operator $T_h(y,t) = y\,\chi_{\{|y|\ge t\}}$. In fact, also using hard thresholding one can obtain estimates similar to (18).

4.3 Analysis of the MSE estimation rate

We will prove that estimate in Theorem 5 is nearly optimal with respect to the rate of convergence; that is, up to the log-like factor, no estimator can achieve a better rate uniformly over $\mathcal{E}^2(A)$. This is the content of the following result which extends a similar 2-dimensional statement from [2].

Theorem 6 *Let $f \in \mathcal{E}^2(A)$ and consider the minimax mean square error*

$$\mathcal{M}(\varepsilon, \mathcal{E}^2(A)) = \inf_{\tilde f} \sup_{\mathcal{E}^2(A)} E\|\tilde f - f\|_2^2.$$

This satisfies

$$\mathcal{M}(\varepsilon, \mathcal{E}^2(A)) \ge C\varepsilon^{\frac{2}{3}} (\log(\varepsilon^{-1}))^{-\frac{2}{3}}, \quad \varepsilon \to 0,$$

for some $C \in \mathbb{R}^+$.

Notice that the proof of Theorem 6 does not follow from the proof of the 2D case in [2] even though the general architecture of the proof is similar. In particular, the crucial Lemma 1, estimating the L^2 norm of the X-ray transform for the indicator function of ellipsoids in $\mathbb{R}^3$, is much more involved than the corresponding 2D case.

The proof of Theorem 6 requires some construction. To begin with, let α be a smooth and nonnegative bivariate function with compact support in $[0, 2\pi] \times [0, \pi]$ with $\|\alpha\|_{C^2} = 1$. For $m \ge 1$, let

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

Notice that $\|\alpha_{i,j,m}\|_{C^2} = \|\alpha\|_{C^2}$ and $\|\alpha_{i,j,m}\|_{L^1} = m^{-4}\|\alpha\|_{L^1}$. 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

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

That is, similar to the 2D argument, the functions $\psi_{i,j,m}$ are characteristic functions of bulges around the sphere of radius ρ_0; notice that they have disjoint supports. Also, we define the radius functions

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

and the corresponding functions

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

each f_ξ is the indicator function of a part of the the sphere of radius ρ_0 plus some addition bulges. Using the fact that α is bounded and nonnegative, a direct calculation shows that

$$\|\psi_{i,j,m}\|^2_{L^2(\mathbb{R}^3)} \simeq \|\alpha_{i,j,m}\|_{L^1([0,2\pi]\times[0,\pi])} = m^{-4}\|\alpha\|_{L^1([0,2\pi]\times[0,\pi])},$$

and, for each radius function r_ξ,

$$\|r_\xi\|_{C^2} \leq \|\alpha_{i,j,m}\|_{C^2} = \|\alpha\|_{C^2}.$$

Let

$$\mathcal{H}_m = \{h = f_0 + \sum_{i,j} \xi_{i,j}\psi_{i,j,m},\ \xi_{i,j} \in \{0, 1\}\}.$$

It is clear that $\mathcal{H}_m \subset \mathcal{E}^2(A)$ and for $f \in \mathcal{H}_m$, the estimator $\hat{f}$ can be chosen from the same set $\mathcal{H}_m$. Set

$$\mathcal{M}(\varepsilon, \mathcal{H}_m) = \inf_{\tilde{f}} \sup_{\mathcal{H}_m} E\|\tilde{f} - f\|_2^2.$$

It follows that

$$\mathcal{M}(\varepsilon, \mathcal{E}^2(A)) \geq \mathcal{M}(\varepsilon, \mathcal{H}_m).$$

We need the following critical lemma.

Lemma 1 *For $m \geq 2$, let g_m be the indicator function of the ellipsoid in $\mathbb{R}^3$ with the center at the origin and the length of the axes $a \simeq m^{-2}, b \simeq m^{-1}$ and $c \simeq m^{-1}$ and let $(Pg_m)(\Theta, u)$ be the X-ray transform of g_m. Then there is an absolute constant C such that*

$$\|Pg_m(\Theta, u)\|^2_{L^2(\Theta,u)} \leq C\, m^{-6} (\ln m)^2.$$

Proof. Let $g_0(x)$ be the indicator function of the unit ball in $\mathbb{R}^3$ and the $g_1(x)$ be the indicator function of the ellipsoid with the center at the origin and the length of the axes a, b and c such that $g_1(x) = g_0(\frac{x_1}{a}, \frac{x_2}{b}, \frac{x_3}{c})$.

A simple calculation shows that $Pg_0(\Theta, u) = 2\left(\sqrt{1-|u|^2}\right)^+$ for any $\Theta \in S^2$ and $u \in \Theta^\perp$, where $\left(\sqrt{1-|u|^2}\right)^+ = \sqrt{1-|u|^2}$ for $|u| \le 1$ and $\left(\sqrt{1-|u|^2}\right)^+ = 0$ for $|u| > 1$. In spherical coordinates, write $\Theta = \cos\theta\sin\phi\,\mathbf{i} + \sin\theta\sin\phi\,\mathbf{j} + \cos\phi\,\mathbf{k}$. We choose $\mathbf{L_1} = \cos\theta\cos\phi\,\mathbf{i} + \sin\theta\cos\phi\,\mathbf{j} - \sin\phi\,\mathbf{k}$, $\mathbf{L_2} = -\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j} + 0\,\mathbf{k}$ as an orthonormal basis in $\Theta^\perp$ so that $u \in \Theta^\perp$ can be written as $u = u_1\mathbf{L_1} + u_2\mathbf{L_2}$. We have:

$$
\begin{aligned}
Pg_1(\Theta,u) &= \int_{-\infty}^{\infty} g_1(t\Theta + u_1\mathbf{L_1} + u_2\mathbf{L_2})dt \\
&= \int_{-\infty}^{\infty} g_0\Big(t(\frac{1}{a}\cos\theta\sin\phi, \frac{1}{b}\sin\theta\sin\phi, \frac{1}{c}\cos\phi) \\
&+\ u_1(\frac{1}{a}\cos\theta\cos\phi, \frac{1}{b}\sin\theta\cos\phi, -\frac{1}{c}\sin\phi) + u_2(-\frac{1}{a}\sin\theta, \frac{1}{b}\cos\theta, 0)\Big)\,dt \\
&= \frac{abc}{\sigma}\int_{-\infty}^{\infty} g_0\Big(t\Theta_0 + u_1(\frac{1}{a}\cos\theta\cos\phi, \frac{1}{b}\sin\theta\cos\phi, -\frac{1}{c}\sin\phi) \\
&+\ u_2(-\frac{1}{a}\sin\theta, \frac{1}{b}\cos\theta, 0)\Big)\,dt,
\end{aligned}
$$

where $\sigma = \sigma(\theta,\phi,a,b,c) = (b^2c^2\cos^2\theta\sin^2\phi + a^2c^2\sin^2\theta\sin^2\phi + a^2b^2\cos^2\phi)^{\frac{1}{2}}$ and $\Theta_0 = \frac{abc}{\sigma}(\frac{1}{a}\cos\theta\sin\phi, \frac{1}{b}\sin\theta\sin\phi, \frac{1}{c}\cos\phi) \in S^2$.

Let $\mathbf{L_3} = \frac{1}{a}\cos\theta\cos\phi\,\mathbf{i} + \frac{1}{b}\sin\theta\cos\phi\,\mathbf{j} - \frac{1}{c}\sin\phi\,\mathbf{k}$, $\mathbf{L_4} = -\frac{1}{a}\sin\theta\,\mathbf{i} + \frac{1}{b}\cos\theta\,\mathbf{j} + 0\,\mathbf{k}$. We can choose $\mathbf{L_5} = \frac{1}{cb}\cos\theta\cos\phi\;\mathbf{i} + \frac{1}{ca}\sin\theta\cos\phi\;\mathbf{j} - \frac{1}{ab}\sin\phi\;\mathbf{k}$, $\mathbf{L_6} = -(\frac{1}{ab^2}\sin\theta\sin^2\phi + \frac{1}{ac^2}\sin\theta\cos^2\phi)\,\mathbf{i} + (\frac{1}{a^2b}\cos\theta\sin^2\phi + \frac{1}{bc^2}\cos\theta\cos^2\phi)\,\mathbf{j} + (\frac{1}{a^2c} - \frac{1}{b^2c})\cos\theta\sin\theta\cos\phi\sin\phi\,\mathbf{k}$ to be an orthogonal basis in $\Theta_0^\perp$. Letting $\mathbf{L_{5_0}} = \frac{1}{\|\mathbf{L_5}\|}\mathbf{L_5}$, $\mathbf{L_{6_0}} = \frac{1}{\|\mathbf{L_6}\|}\mathbf{L_6}$, we can write $\mathbf{L_3}$, $\mathbf{L_4}$ as

$$
\mathbf{L_3} = \langle\mathbf{L_3},\Theta_0\rangle\Theta_0 + \langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle\mathbf{L_{5_0}} + \langle\mathbf{L_3},\mathbf{L_{6_0}}\rangle\mathbf{L_{6_0}},
$$

$$
\mathbf{L_4} = \langle\mathbf{L_4},\Theta_0\rangle\Theta_0 + \langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle\mathbf{L_{5_0}} + \langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle\mathbf{L_{6_0}}.
$$

It follows that

$$
\begin{aligned}
Pg_1(\Theta,u) &= \frac{abc}{\sigma}\int_{-\infty}^{\infty} g_0\big((t + u_1\langle\mathbf{L_3},\Theta_0\rangle + u_2\langle\mathbf{L_4},\Theta_0\rangle)\Theta_0 + (u_1\langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle + \\
&+\ u_2\langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle)\mathbf{L_{5_0}} + (u_1\langle\mathbf{L_3},\mathbf{L_{6_0}}\rangle + u_2\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle)\mathbf{L_{6_0}}\big)dt \\
&= \frac{abc}{\sigma}\int_{-\infty}^{\infty} g_0\big(t\Theta_0 + (u_1\langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle + u_2\langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle)\mathbf{L_{5_0}} + \\
&+\ (u_1\langle\mathbf{L_3},\mathbf{L_{6_0}}\rangle + u_2\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle)\mathbf{L_{6_0}}\big)dt \\
&= \frac{2abc}{\sigma}\left(\sqrt{1 - (u_1\langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle + u_2\langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle)^2 - (u_1\langle\mathbf{L_3},\mathbf{L_{6_0}}\rangle + u_2\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle)^2}\right)^+.
\end{aligned}
$$

Let $D = \langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle - \langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle\langle\mathbf{L_3},\mathbf{L_{6_0}}\rangle$. It is easy to check that $\langle\mathbf{L_4},\mathbf{L_{5_0}}\rangle = 0$, $\langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle = \frac{1}{abc\|\mathbf{L_5}\|}$ and $\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle = \frac{1}{a^2b^2c^2\|\mathbf{L_6}\|}\sigma_1^2$, where $\sigma_1^2 = c^2\sin^2\phi + (b^2\sin^2\theta + a^2\cos^2\theta)\cos^2\phi$. It follows that $D = \langle\mathbf{L_3},\mathbf{L_{5_0}}\rangle\langle\mathbf{L_4},\mathbf{L_{6_0}}\rangle = \frac{\sigma_1^2}{a^3b^3c^3\|\mathbf{L_5}\|\|\mathbf{L_6}\|} =$

$\frac{\sigma_1^2}{\sigma_2\sigma_3}$, where $\sigma_2 = (a^2\cos^2\theta\cos^2\phi + b^2\sin^2\theta\cos^2\phi + c^2\sin^2\phi)^{\frac{1}{2}}$, $\sigma_3 = \big((ac^2\sin\theta\sin^2\phi + ab^2\sin\theta\cos^2\phi)^2 + (bc^2\cos\theta\sin^2\phi + a^2b\cos\theta\cos^2\phi)^2 + (b^2c - a^2c)^2\cos^2\theta\sin^2\theta\cos^2\phi\sin^2\phi\big)^{\frac{1}{2}}$. Hence we have that

$$\begin{aligned}\|Pg_1(\Theta,u)\|_2^2 &= 4(abc)^2\int_0^{\pi}\int_0^{2\pi}\frac{1}{\sigma^2 D}\int_{R^2}\left(\left(\sqrt{1-|u|^2}\right)^+\right)^2 du\,d\theta\,d\phi\\ &= 2(abc)^2\,\pi\int_0^{\pi}\int_0^{2\pi}\frac{\sigma_2\sigma_3}{\sigma^2\sigma_1^2}\,d\theta\,d\phi\end{aligned}$$

For $m \geq 2$, let $a \simeq m^{-2}$, $b \simeq m^{-1}$, $a \simeq m^{-1}$, where $a \simeq m^{-2}$ means that there are positive constants c_1, c_2 such that $c_1 m^{-2} \leq a \leq c_2 m^{-2}$. It follows that

$$\begin{aligned}\sigma^2 &\simeq m^{-4}(\cos^2\theta\sin^2\phi + m^{-2}(\sin^2\theta\sin^2\phi + \cos^2\phi)\\ &\simeq m^{-4}(\cos^2\theta\sin^2\phi + m^{-2});\\ \sigma_1^2 &\simeq m^{-2}(\sin^2\phi + (\sin^2\theta + m^{-2}\cos^2\theta)\cos^2\phi)\\ &\simeq m^{-2}(1-(1-m^{-2})\cos^2\theta\cos^2\phi);\\ \sigma_2 &\simeq m^{-1}(m^{-2}\cos^2\theta\cos^2\phi + \sin^2\theta\cos^2\phi + \sin^2\phi)^{\frac{1}{2}} \simeq \sigma_1;\\ \sigma_3 &\simeq m^{-3}\big((m^{-2}\sin^2\theta + (m^{-2}+\sin^2\phi)^2\cos^2\theta + \cos^2\theta\sin^2\theta\cos^2\phi\sin^2\phi\big)^{\frac{1}{2}}.\end{aligned}$$

Using this observation, we have that

$$\begin{aligned}&m^6\,\|Pg_m(\Theta,u)\|_2^2\\ &\simeq \int_0^{\pi}\!\!\int_0^{2\pi}\frac{\big(m^{-2}\sin^2\theta + (m^{-2}+\sin^2\phi)^2\cos^2\theta + \cos^2\theta\sin^2\theta\cos^2\phi\sin^2\phi\big)^{\frac{1}{2}}}{(\cos^2\theta\sin^2\phi + m^{-2})(1-(1-m^{-2})\cos^2\theta\cos^2\phi)^{\frac{1}{2}}}d\theta d\phi\end{aligned}$$

Since the two factors of the denominator in the above integral have different zeros, in order to estimate $\|Pg_1(\Theta,u)\|_2^2$ we can consider the two factors $\cos^2\theta\sin^2\phi + m^{-2}$ and $(1-(1-m^{-2})\cos^2\theta\cos^2\phi)^{\frac{1}{2}}$ separately. In the following, we only examine the first factor (setting the second factor $= 1$), which has higher order of zeros than the second one has. The discussion for the second factor is similar. Also we may assume that $(\theta,\phi) \in [0,\frac{\pi}{2}] \times [0,\frac{\pi}{2}]$. Hence, we then can write

$$\begin{aligned}&m^6\,\|Pg_m(\Theta,u)\|_2^2\\ &\simeq \int_0^{\frac{\pi}{2}}\!\!\int_0^{\frac{\pi}{2}}\frac{\big(m^{-2}\sin^2\theta + (m^{-2}+\sin^2\phi)^2\cos^2\theta + \cos^2\theta\sin^2\theta\cos^2\phi\sin^2\phi\big)^{\frac{1}{2}}}{(\cos^2\theta\sin^2\phi + m^{-2})}d\theta d\phi\\ &\simeq \int_0^{\frac{\pi}{2}}\!\!\int_0^{\frac{\pi}{2}}\frac{\big(m^{-1}\sin\theta + (m^{-2}+\sin^2\phi)\cos\theta + \cos\theta\sin\theta\cos\phi\sin\phi}{(\cos^2\theta\sin^2\phi + m^{-2})}d\theta d\phi.\end{aligned}$$

Since the behavior of this expression is determined by the zeros of $\cos\theta$ and $\sin\phi$, we will decompose the above double integral $\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}$ as

$$\int_0^{\frac{1}{m}}\int_{\frac{\pi}{2}-\frac{1}{m}}^{\frac{\pi}{2}} + \int_{\frac{1}{m}}^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}-\frac{1}{m}} + \int_0^{\frac{1}{m}}\int_0^{\frac{\pi}{2}-\frac{1}{m}} + \int_{\frac{1}{m}}^{\frac{\pi}{2}}\int_{\frac{\pi}{2}-\frac{1}{m}}^{\frac{\pi}{2}}.$$

Noticing that $(\cos^2\theta\sin^2\phi + m^{-2}) \geq \max\{\cos^2\theta\sin^2\phi, m^{-2}, m^{-1}\cos\theta\sin\phi\}$ on $[0,\frac{\pi}{2}] \times [0,\frac{\pi}{2}]$, that $\cos\theta \simeq \frac{\pi}{2} - \theta$ for θ near $\frac{\pi}{2}$, and that $\sin\phi \simeq \phi$ for ϕ near 0, we have:

$$\int_0^{\frac{1}{m}}\int_{\frac{\pi}{2}-\frac{1}{m}}^{\frac{\pi}{2}}\frac{m^{-1}\sin\theta + (m^{-2}+\sin^2\phi)\cos\theta + \cos\theta\sin\theta\cos\phi\sin\phi}{\cos^2\theta\sin^2\phi + m^{-2}}d\theta d\phi$$

$$\leq \int_0^{\frac{1}{m}} \int_{\frac{\pi}{2}-\frac{1}{m}}^{\frac{\pi}{2}} (m+1+2m^2) d\theta d\phi \leq C;$$

$$\int_{\frac{1}{m}}^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}-\frac{1}{m}} \frac{(m^{-1}\sin\theta + (m^{-2}+\sin^2\phi)\cos\theta + \cos\theta\sin\theta\cos\phi\sin\phi}{(\cos^2\theta\sin^2\phi + m^{-2})} d\theta d\phi$$

$$\leq \int_{\frac{1}{m}}^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}-\frac{1}{m}} (\frac{1}{\cos\theta\sin\phi} + 1 + \frac{1}{\cos\theta} + \frac{1}{\cos\theta\sin\phi})) d\theta d\phi$$

$$\leq C\,(\ln m)^2;$$

$$\int_0^{\frac{1}{m}} \int_0^{\frac{\pi}{2}-\frac{1}{m}} \frac{m^{-1}\sin\theta + (m^{-2}+\sin^2\phi)\cos\theta + \cos\theta\sin\theta\cos\phi\sin\phi}{\cos^2\theta\sin^2\phi + m^{-2}} d\theta d\phi$$

$$\leq \int_0^{\frac{1}{m}} \int_0^{\frac{\pi}{2}-\frac{1}{m}} (m+1+m+m) d\theta d\phi \leq C.$$

Finally, the integral $\int_{\frac{1}{m}}^{\frac{\pi}{2}} \int_{\frac{\pi}{2}-\frac{1}{m}}^{\frac{\pi}{2}}$ can be estimates exactly like the last term. Hence it follows that

$$\|Pg_m(\Theta,u)\|_2^2 \leq Cm^{-6}\,(\ln m)^2,$$

and this completes the proof of Lemma 1. □

We also need the following lower bound on the estimator from [2, Lemma 8.3].

Lemma 2 *For $N \geq 1$, let $\xi \in \{0,1\}^N$ and $X = N(\xi, V)$ be a multivariate Gaussian vector. Assume that V is invertible such that $\tau_i^2 = Var(X_l | X_k, k \neq l) = \frac{1}{(V^{-1})_{ll}} \geq 1$ for all $1 \leq l \leq N$. Then there is an absolute constant B such that*

$$\inf_{\hat{\xi}} \sup_{\xi \in \{0,1\}^N} E\|\hat{\xi} - \xi\|_2^2 \geq B \cdot N.$$

We can now prove Theorem 6. This argument follows the general idea from [2].

Proof of Theorem 6. From the inequality $\mathcal{M}(\varepsilon, \mathcal{E}^2(A)) \geq \mathcal{M}(\varepsilon, \mathcal{H}_m)$ we see that, in order to estimate f from the noisy data $Y = Pf + \epsilon W$, it is enough to consider $f \in \mathcal{H}_m$. Furthermore, via the L^2 projection onto the smallest affine subspace containing $\mathcal{H}_m$, we can restrict the estimator $\hat{f}$ to be in this subspace. That is, we can write $\hat{f} = f_0 + \sum_{1\leq i,j\leq m} \hat{\xi}_{i,j}\,\psi_{i,j,m}$, where $\hat{\xi} = (\hat{\xi}_{i,j}) \in \mathbb{R}^N$, $N = m^2$. Since $f \in \mathcal{H}_m$, we can write $f = f_0 + \sum_{1\leq i,j\leq m} \xi_{i,j}\,\psi_{i,j,m}$, where $\xi = (\xi_{i,j}) \in \{0,1\}^N$, $N = m^2$. From the orthogonality of the functions $\psi_{i,j,m}$ and the fact that $\|\psi_{i,j,m}\|_2^2 \simeq m^{-4}$, we have

$$\|\hat{f} - f\|_{L^2} \simeq m^{-4} \|\hat{\xi} - \xi\|_{\ell^2}. \tag{19}$$

Hence, it remains to control the term $\|\hat{\xi} - \xi\|_{\ell^2}$.

Let $g_{i,j} = P\psi_{i,j,m}$ be the X-ray transform of the hypercube $\mathcal{H}_m$ generators $\psi_{i,j,m}$. We notice that, even though the functions $g_{i,j}$ are no longer orthogonal, they are still linearly independent. Let V_m denote the affine space $\{Pf_0 + \sum_{i,j} \theta_{i,j} g_{i,j},\ 1 \leq i,j \leq m\}$ for arbitrary choices $(\theta_{i,j})$. For each (i,j), let $Y_{i,j} = \langle Y, g_{i,j}\rangle - \langle Pf_0, g_{i,j}\rangle$. Similar to the 2D argument in [2, Sec.8], the vector $Y = (Y_{i,j})$ is a sufficient statistic for the ξ's and we may restrict our attention to estimators that are functions of Y alone. In matrix notation, we have $Y \approx N(G\xi, \epsilon G)$, where G is the Gram matrix of the functions $g_{i,j}$, i.e., $G_{i,j;i',j'} = \langle g_{i,j}, g_{i',j'}\rangle$. Since the functions $g_{i,j}$ are linearly independent, it follows that G is invertible. Define $X = G^{-1}Y$. Since Y is a sufficient statistic for the ξ's, so is X and we may restrict our attention to estimators that are functions of X alone. We have that $X \approx N(\xi, \epsilon G^{-1})$. Now let $V = \epsilon G^{-1}$ so that $(V^{-1})_{i,j;i,j} = \epsilon^{-2}\|g_{i,j}\|^2$. By choosing a sufficiently large m (this may depend on ϵ) such that $\|g_{i,j}\|^2 \leq \epsilon^2$ for all $1 \leq i, j \leq m$, then one can apply Lemma 2 for X and for $N = m^2$ to get

$$\inf_{\hat{\xi}} \sup_{\xi \in \{0,1\}^N} E\|\hat{\xi} - \xi\|_2^2 \geq Bm^2.$$

From the above inequality, using (19) we have that $\mathcal{M}(\varepsilon, \mathcal{H}_m) \geq Bm^{-2}$ and hence

$$\mathcal{M}(\varepsilon, \mathcal{E}^2(A)) \geq Bm^{-2}. \tag{20}$$

For any $\psi_{i,j,m}$, there are ellipsoid D and D_1, both having the three axes of side length $\simeq m^{-2}, m^{-1}, m^{-1}$, such that $1_D \leq \psi_{i,j,m} \leq 1_{D_1}$. We observe that, by the definition of $\psi_{i,j,m}$, the ratio of axis lengths for the inscribing and circumscribing ellipsoids are bounded independently of i, j, and m. Since the L^2 norm of Pf is invariant with respect to both rotation and translation of the variables of f in $\mathbb{R}^3$, we may assume that, for any i, j, the function $\psi_{i,j,m}$ is the indicator function of the ellipsoid in $\mathbb{R}^3$ with center at the origin and axis side lengths $\simeq m^{-2}, m^{-1}, m^{-1}$.

By Lemma 1, we have that there is a constant $C > 0$ such that $\|g_{i,j}\|_2^2 \leq Cm^{-6}(\ln m)^2$ for all i, j. Given $\epsilon > 0$, one can choose a large m with $\epsilon \simeq m^{-3} \ln m$ so that $\|g_{i,j}\|_2^2 \leq \epsilon^2$ for all $1 \leq i, j \leq m$, which is the condition for us to apply Lemma 2.

Finally from $\epsilon \simeq m^{-3} \ln m$, we have $m^{-2} \simeq \epsilon^{\frac{2}{3}}(\ln m)^{-\frac{2}{3}}$ and $\ln m \simeq \ln \epsilon^{-1}$, which implies $m^{-2} \simeq \epsilon^{\frac{2}{3}}(\ln \epsilon^{-1})^{-\frac{2}{3}}$. The proof is completed by replacing m^{-2} in (20) with $\epsilon^{\frac{2}{3}}(\ln \epsilon^{-1})^{-\frac{2}{3}}$. □

5 Appendix: The Proof of Theorem 4.1

Let $\{s_\mu : \mu \in \mathcal{M}^0\}$ be the modified shearlet system introduced in Section 4.1. Recall that $\mathcal{M}^0 = N \cup M^0$, where $N = \mathbb{Z}^3$, $M^0 = \{\mu = (j, \ell_1, \ell_2, k, d) : j \geq j_0, -2^j \leq \ell_1,\ \ell_2 \leq 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$, and the system is made of the coarse scale system $\{s_\mu = 2^{j_0}\varphi(2^{j_0}x - \mu) : \mu \in N\}$ and the fine scale system $\{s_\mu = \psi_\mu = \psi^{(d)}_{j,\ell_1,\ell_2,k} : \mu \in M^0\}$.

In order to prove Theorem 4.1, we need the following lemma which provides an estimate for the size of the shearlet coefficients at a fixed scale j (where $j \geq j_0$). For such j fixed, recall from Section 3 that $M_j = \{(j, \ell, k, d) : -2^j \leq \ell_1, \ell_2 \leq 2^j, k \in \mathbb{Z}^2, d = 1, 2, 3\}$. Due to the symmetry of ℓ_1 and ℓ_2 in M_j, there is no loss of generality in assuming that $|\ell_1| \leq |\ell_2|$.

Lemma 3 *Let $f \in \mathcal{E}^2(A)$ and $j \geq j_0$. Then there is a positive constant C such that*

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

Proof. It is useful to introduce a smooth localization of the function f near dyadic squares. Let $\mathcal{Q}_j$ be the collection of dyadic squares of the form

$Q = [\frac{\nu_1}{2^j}, \frac{\nu_1+1}{2^j}] \times [\frac{\nu_2}{2^j}, \frac{\nu_2+1}{2^j}] \times [\frac{\nu_3}{2^j}, \frac{\nu_3+1}{2^j}]$, with $\nu_1, \nu_2, \nu_3 \in \mathbb{Z}$. For a nonnegative C^∞ function w with support in $[-1,1]^3$, we define a smooth partition of unity

$$\sum_{Q \in \mathcal{Q}_j} w_Q^2(x) = 1, \quad x \in \mathbb{R}^3,$$

where, for each dyadic square $Q \in \mathcal{Q}_j$, $w_Q(x) = w(2^j x_1 - \nu_1, 2^j x_2 - \nu_2, 2^j x_3 - \nu_3)$.

Given $f \in \mathcal{E}^2(A)$, the coefficients $\{\langle f, \psi_\mu \rangle\}$ will exhibit a very different behavior depending on whether the edge surface of f intersects the support of w_Q or not. We split $\mathcal{Q}_j$ into the disjoint sets $\mathcal{Q}_j^0$ and $\mathcal{Q}_j^1$ that indicate whether the collection of dyadic squares Q intersects an edge surface or not. Since each dyadic square Q has side-length $2 \cdot 2^{-j}$ and f has compact support in $[0,1]^3$, there are $O(2^{2j})$ dyadic cubes in $Q \in \mathcal{Q}_j^0$ intersecting the edge surface and $O(2^{3j})$ dyadic cubes in $Q \in \mathcal{Q}_j^1$ not intersecting the edge surface.

For each such cube $Q \in \mathcal{Q}_j^0$, as a corollary of Theorem 4.1 in [13], we have

$$\sum_{k \in \mathbb{Z}^3} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, 2^{-4j}\, (1 + |\ell_2|)^{-5}.$$

More precisely, in the proof of Theorem 3.3 in [13], it is shown that

$$\sum_{k \in R_K} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, L_K^{-2}\, 2^{-3j}\, (1 + |\ell|)^{-5},$$

where $K \in \mathbb{Z}^3$, $\cup_{K \in \mathbb{Z}^3} R_K = \mathbb{Z}^3$ and $\sum_{K \in \mathbb{Z}^3} L_K^{-2} < \infty$. Hence

$$\sum_{\mu \in M_j} |\langle f_Q, \psi_\mu \rangle|^2 = \sum_{|\ell_1| \le |\ell_2| \le 2^j} \sum_{k \in \mathbb{Z}^3} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, 2^{-4j}.$$

Adding up over all cubes in $\mathcal{Q}_j^0$, we have that

$$\sum_{\mu \in M_j} \sum_{Q \in \mathcal{Q}_j^0} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, 2^{-2j}. \tag{21}$$

Similarly for $Q \in \mathcal{Q}_j^1$, from the proof of Theorem 3.4 in [13], one can obtain that

$$\sum_{\mu \in M_j} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, 2^{-11j}.$$

Hence adding up over all $Q \in \mathcal{Q}_j^1$, it follows that

$$\sum_{\mu \in M_j} \sum_{Q \in \mathcal{Q}_j^1} |\langle f_Q, \psi_\mu \rangle|^2 \le C\, 2^{-8j}. \tag{22}$$

The proof is completed by combining the estimates (21) and (22). □

We now proceed with the proof of Theorem 4.1.

Proof of Theorem 4.1.1(1): We need to establish

$$\sup_{f \in \mathcal{E}^2(A)} \sum_{\mu \notin \mathcal{N}(\varepsilon)} |\langle f, s_\mu \rangle|^2 \le C' \varepsilon^{2/3}.$$

(I) We start by examining the situation at fine scales, for $j \geq j_1(\varepsilon) = \frac{1}{3}\log_2(\varepsilon^{-1})$, so that $2^{-j} \leq \varepsilon^{\frac{1}{3}}$. Notice that, for these values of the index j, $s_\mu = \psi_\mu$.

By Lemma 3, for each $f \in \mathcal{E}^2(A)$ and each $j \geq j_0$, we have that

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

Hence

$$\sum_{j > j_1} \sum_{\mu \in M_j} |\langle f, \psi_\mu \rangle|^2 \leq C \sum_{j > j_1} 2^{-2j} \leq C\, \varepsilon^{\frac{2}{3}}.$$

(II) Let $Q_0 = [0,1]^3$ and $\operatorname{supp} f \subset Q_0$. We will show that the terms $\langle f, s_\mu \rangle$ decay very rapidly for locations k away from Q_0.

We will start by examining the decay of a fine-scale term $\langle f, s_\mu \rangle$, for $\mu = (j, \ell, k, d) \in M^0$ and $d = 1$. Let $E_{j\ell} = B_{(1)}^{[\ell]} A_{(1)}^j$, where $A_{(1)}$ and $B_{(1)}^{[\ell]}$ are given in Sec. 2.

By the assumptions on ψ, it follows that, for each $m \in \mathbb{N}$, there is a constant $C_m > 0$ such that

$$|\psi(x)| \leq C_m\, (1 + \|x\|)^{-m}. \tag{23}$$

It follows that

$$|\psi_{j,\ell,k}^{(1)}(x)| \leq C_m\, 2^{2j}\, (1 + \|E_{j\ell}\, x - k\|)^{-m}.$$

We will use two simple facts. The first one is that $\|E_{j\ell}\, x\| \leq \|E_{j\ell}\| \|x\| = 2^{2j} \|x\| \leq \sqrt{3}\, 2^{2j}$ for $x \in Q_0$ and the second one is that for $a > 0,\ 0 \leq b \leq c \leq a$, we have $a - b \geq a - c$. It follows that, for $|k| \geq 2^{2j+1}$, we have

$$\begin{aligned}
|\langle f, \psi_{j,\ell,k}^{(1)} \rangle| &\leq \|f\|_\infty \int_Q |\psi_{j,\ell,k}^{(1)}(x)|\, dx \\
&\leq C_m\, 2^{2j} \int_{Q_0} \Big(1 + \|E_{j\ell} x - k\|\Big)^{-m} dx \\
&\leq C_m\, 2^{2j} \int_{Q_0} \Big(1 + |k| - \|E_{j\ell}\, x\|\Big)^{-m} dx \\
&\leq C_m\, 2^{2j} \int_{Q_0} (|k| - 2^{2j}\|x\|)^{-m} dx \\
&\leq C_m\, 2^{2j}\, (|k| - \sqrt{3}\, 2^{2j})^{-m}.
\end{aligned} \tag{24}$$

Thus:

$$\begin{aligned}
\sum_{|\ell_1| \leq |\ell_2| \leq 2^j} \sum_{|k| \geq 2^{2j+1}} |\langle f, \psi_{j,\ell,k}^{(1)} \rangle|^2 &\leq C_m \sum_{|\ell_1| \leq |\ell_2| \leq 2^j} 2^{6j} \sum_{|k| \geq 2^{2j+1}} (|k| - \sqrt{3}\, 2^{2j})^{-2m} \\
&= C_m\, 2^{6j} \sum_{|k| \geq 2^{2j+1}} (|k| - \sqrt{3}\, 2^{2j})^{-2m} \\
&\leq C_m\, 2^{6j}\, 2^{-2j(2m-2)} \\
&= C_m\, 2^{10j}\, 2^{-4jm}.
\end{aligned}$$

Now we can add up all contributions for $j \geq j_0$. Since we can choose m arbitrarily large, for an appropriate choice of the constant C, we have:

$$\sum_{j\geq j_0} \sum_{|\ell_1|\leq|\ell_2|\leq 2^j} \sum_{|k|\geq 2^{2j+1}} |\langle f, \psi^{(1)}_{j,\ell,k}\rangle|^2 \leq C_m \sum_{j\geq j_0} 2^{10j}\, 2^{-4jm} \leq C\, 2^{-5j_0} \leq C\varepsilon^{\frac{2}{3}}.$$

The analysis in the case where $\mu \in M^0$ and $d = 2,3$ is essentially the same as the one given above. For the coarse case terms, notice first that φ satisfies the same decay behavior as (23) for ψ. Hence, letting $\varphi_{j_0,k}(x) = 2^{\frac{3}{2}j_0}\varphi(2^{j_0}x - k)$ and proceeding as in (24), we have that

$$|\langle f, \varphi_{j_0,k}\rangle| \leq \|f\|_\infty \int_Q |\varphi_{j_0,k}(x)|\, dx \leq C_m\, 2^{\frac{3}{2}j_0}\, (|k| - \sqrt{3}\, 2^{j_0})^{-m}.$$

Now we can proceed as above by summing over $|k| \geq 2^{2j_0+1}$ and using the fact that m can be chosen arbitrarily, to conclude that

$$\sum_{|k|\geq 2^{2j_0+1}} |\langle f, \varphi_{j_0,k}\rangle|^2 \leq C\varepsilon^{\frac{2}{3}}.$$

Combining the estimates from parts (I) and (II) and of the proof, we finally have that

$$\sum_{\mu\notin\mathcal{N}(\varepsilon)} |\langle f, s_\mu\rangle|^2 \leq C\,\varepsilon^{\frac{2}{3}}. \quad \square$$

Proof of Theorem 4.1.1(2): For $\mu \in \mathcal{M}^0$, we use the notation $c_\mu = \langle f, s_\mu\rangle$ and we define the set

$$R(j,\varepsilon) = \{\mu \in M_j : |c_\mu| > \varepsilon\},$$

to denote the set of "large" shearlet coefficients, at a fixed scale j.

By Corollary 3.5 in [13] (which is valid both for coarse and fine scale shearlets), there is a constant $C > 0$ such that, as $\varepsilon \to 0$,

$$\# R(j,\varepsilon) \leq C\,\varepsilon^{-1}.$$

It follows by a simple rescaling argument that

$$\# R(j, 2^j\varepsilon) \leq C\, 2^{-j}\varepsilon^{-1}.$$

Since $\widehat{\psi} \in C_0^\infty(\mathbb{R}^3)$, for $\mu = (j,\ell,k,d) \in M^0$ and $d = 1$, we have that

$$\begin{aligned} |c_\mu| = |\langle f, \psi_\mu\rangle| &= \left| \int_{\mathbb{R}^3} f(x) 2^{2j} \psi(B_1^\ell A_1^j x - k) dx \right| \\ &\leq 2^{-2j} \|f\|_\infty \int_{\mathbb{R}^3} |\psi(z)| dz = C' 2^{-2j}. \end{aligned}$$

Thus, we can assume that $R(j, 2^j\varepsilon) = \emptyset$ when $2^j > \varepsilon^{-1/3}$ (that is, $j > j_1(\varepsilon) = -\frac{1}{3}\log_2(\varepsilon^{-1})$). Similarly, $R(j,\varepsilon) = \emptyset$ when $2^j > \varepsilon^{-1/2}$ (that is, $j > j_2(\varepsilon) = \frac{1}{2}\log_2(\varepsilon^{-1})$). For $\mu \in M^0$ and $d = 2,\ 3$, we get exactly the same estimates.

For the risk proxy, notice that

$$\sum_{\{\mu\in\mathcal{N}(\varepsilon)\}} \min(c_\mu^2, 2^{2j}\varepsilon^2) = S_1(\varepsilon) + S_2(\varepsilon),$$

where

$$S_1(\varepsilon) = \sum_{\{\mu\in\mathcal{N}(\varepsilon)\colon |c_\mu|\geq 2^j\,\varepsilon\}} \min(c_\mu^2, 2^{2j}\varepsilon^2)$$

$$S_2(\varepsilon) = \sum_{\{\mu\in\mathcal{N}(\varepsilon)\colon |c_\mu|< 2^j\,\varepsilon\}} \min(c_\mu^2, 2^{2j}\varepsilon^2).$$

Hence, using the observations above, we have:

$$\begin{aligned}
S_1(\varepsilon) &= \sum_{\{\mu\in\mathcal{N}(\varepsilon)\colon |c_\mu|\geq 2^j\,\varepsilon\}} 2^{2j}\,\varepsilon^2 \\
&\leq \sum_{j\leq j_1}\ \sum_{\{\mu\in M_j\colon |c_\mu|\geq 2^j\,\varepsilon\}} 2^{2j}\,\varepsilon^2 \\
&\leq C \sum_{j\leq j_1} (2^{-j}\varepsilon^{-1})\, 2^{2j}\,\varepsilon^2 \\
&= C \sum_{j\leq j_1} 2^j\,\varepsilon \\
&\leq C\, 2^{j_0}\varepsilon \leq C\,\varepsilon^{\frac{2}{3}}.
\end{aligned}$$

For S_2, we have

$$\begin{aligned}
S_2(\varepsilon) &= \sum_{\{\mu\in\mathcal{N}(\varepsilon)\colon |c_\mu|< 2^j\,\varepsilon\}} |c_\mu|^2 \\
&= \sum_{j_0\leq j\leq j_1}\sum_{n=0}^{\infty}\ \sum_{\{2^{j-n-1}\varepsilon\leq |c_\mu|<2^{j-n}\varepsilon\}} |c_\mu|^2 \\
&\leq C \sum_{j_0\leq j\leq j_1}\sum_{n=0}^{\infty} 2^{-(j-n-1)}\varepsilon^{-1} 2^{2(j-n)}\,\varepsilon^2 \\
&= C \sum_{j_0\leq j\leq j_1}\sum_{n=0}^{\infty} 2^{-n+1}\, 2^j\,\varepsilon \\
&\leq C \sum_{j_0\leq j\leq j_1} 2^j\,\varepsilon \\
&\leq C\,\varepsilon^{\frac{2}{3}}. \quad \square
\end{aligned}$$

Proof of Theorem 4.1.1(3): For each fixed scale $j_0 \leq j \leq j_1$, the number of indices μ in $\mathcal{N}(\varepsilon)\cap M_j$ is of the order $O(2^{8j})$. In fact, $\mathcal{N}(\varepsilon)\cap M_j \subset \{(j,\ell,k,d) : |k| \leq 2^{2j+1}, |\ell_1| \leq 2^j, |\ell_2| \leq 2^j\}$ and this set contains $O(2^{6j})$ terms for the k variable and $O(2j)$ terms for the ℓ_1, ℓ_2 variables. Hence, adding up the contributions corresponding to the various scales, we obtain:

$$\#\mathcal{N}(\varepsilon) \leq C \sum_{j\leq j_1} 2^{8j} \leq C\varepsilon^{-\frac{8}{3}}. \quad \square$$

Acknowledgements The authors thank E. Candès for useful discussions. The authors acknowledge partial support from NSF grants 1008900/1008907 and DMS 1005799.

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. Candeś, and D. L. Donoho, Recovering edges in ill-posed inverse problems: optimality of curvelet frames, Annals Stat., **30**(3) (2002), 784–842.
3. F. Colonna, G.R. Easley, K. Guo, D. Labate, Radon transform inversion using the shearlet representation, Appl. Comput. Harmon. Anal, **29**(2) (2010), 232-250.
4. A. M. Cormack, Representation of a function by its line integrals, with some radiological applications, J. Appl. Phys. **34** (1963), 2722–2727.
5. A. M. Cormack, Representation of a function by its line integrals, with some radiological applications II, J. Appl. Phys., **35** (1964), 195–207.
6. D. L. Donoho, Nonlinear solution of linear inverse problem by wavelet-vaguelettedecomposition, Appl. Comput. Harmon. Anal., **2** (1995), pp. 101–126.
7. D. L. Donoho, I. M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika, **81** (1994), 425–455.
8. G. R. Easley, F. Colonna, D. Labate, Improved Radon Based Imaging using the Shearlet Transform, Proc. SPIE, Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks II, **7343** (2009), Orlando.
9. G. R. Easley, D. Labate, F. Colonna, Shearlet-Based Total Variation Diffusion for Denoising, IEEE Trans. Image Proc. **18** (2) (2009), 260–268.
10. 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.
11. K. Guo, and D. Labate, Characterization and analysis of edges using the continuous shearlet transform, SIAM J. Imag. Sci. **2** (2009), 959–986.
12. K. Guo, and D. Labate, Optimally Sparse 3D Approximations using Shearlet Representations, Electronic Research Announcements in Mathematical Sciences, **17** (2010), 126–138.
13. K. Guo, D. Labate, Optimally Sparse Representations of 3D Data with C^2 Surface Singularities Using Parseval Frames of Shearlets, to appear in SIAM J. Math. Anal. (2012)
14. K. Guo, and D. Labate, The Construction of Smooth Parseval Frames of Shearlets, preprint (2011).
15. K. Guo, W. Lim, D. Labate, G. Weiss, E. Wilson, Wavelets with composite dilations, ERA Amer. Math. Soc., **10** (2004), 78–87.
16. K. Guo, W. Lim, D. Labate, G. Weiss, E. Wilson, Wavelets with composite dilations and their MRA properties, Appl. Computat. Harmon. Anal., **20** (2006) 231–249.
17. B. Han, G. Kutyniok, and Z. Shen, Adaptive Multiresolution Analysis Structures and Shearlet Systems SIAM J. Numer. Anal. 49 (2011), 1921-1946.
18. S. Helgason, *The Radon Transfrom*, Birkhäuser, Boston, 1980.
19. E. D. Kolaczyk, A wavelet shrinkage approach to tomographic image reconstruction, J. Amer. Statist. Assoc., **91** (1996), 1079–1090.
20. N.Y. Lee, and B. J. Lucier, Wavelet Methods for Inverting the Radon Transform with Noisy Data, IEEE Trans. Image Proc., **10** (1) (2001), 79–94.
21. F. Natterer, and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM Monographs on Mathematical Modeling and Computation, Philadelphia, 2001.
22. R. Neelamani, R. Nowak, and R. Baraniuk, Model-based inverse halftoning with Wavelet-Vaguelette Deconvolution, in Proc. IEEE Int. Conf. Image Processing ICIP 2000, (Vancouver, Canada), pp. 973-976, Sept. 2000.
23. P. S. Negi and D. Labate, 3D Discrete Shearlet Transform and Video Processing, to appear in IEEE Trans. Image Process. (2012).
24. J. Radon, Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, Berichte Sächsische Akademie der Wissenschaften, Leipzig, Math.–Phys. Kl., **69**, 262–267, reprinted in [18] , 177–192, 1917.
25. V.M. Patel, G. Easley, D. M. Healy, Shearlet-based deconvolution IEEE Trans. Image Process. **18**(12) (2009), 2673-2685
26. 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.