

## Optimally sparse representations of cartoon-like cylindrical data

Glenn R. Easley · Kanghui Guo ·  
Demetrio Labate · Basanta R. Pahari

Received: date / Accepted: date

**Abstract** Sparse representations of multidimensional data have received a significant attention in the literature due to their applications in problems of data restoration and feature extraction. In this paper, we consider an idealized class  $\mathcal{C}^2(Z) \subset L^2(\mathbb{R}^3)$  of 3-dimensional data dominated by surface singularities that are orthogonal to the  $xy$  plane. To deal with this type of data, we introduce a new multiscale directional representation called cylindrical shearlets and prove that this new approach achieves superior approximation properties not only with respect to conventional multiscale representations but also with respect to 3-dimensional shearlets and curvelets. Specifically, the  $N$ -term approximation  $f_N^S$  obtained by selecting the  $N$  largest coefficients of the cylindrical shearlet expansion of a function  $f \in \mathcal{C}(Z)$  satisfies the asymptotic estimate

$$\|f - f_N^S\|_2^2 \leq c N^{-2} (\ln N)^3, \quad \text{as } N \rightarrow \infty.$$

This is the optimal decay rate, up the logarithmic factor, outperforming 3d wavelet and 3d shearlet approximations which only yield approximation rates of order  $N^{-1/2}$  and  $N^{-1}$  (ignoring logarithmic factors), respectively, on the same type of data.

---

Dr. Demetrio Labate  
University of Houston  
Tel.:+1 (713) 743-3492  
Fax:+1 (713) 743-3505  
E-mail: dlabate@math.uh.edu

Basanta R Pahari  
University of Houston  
E-mail: brpahari@math.uh.edu

Dr. Kanghui Guo  
Missouri State University  
E-mail: kanghuigu@missouristate.edu

Dr. Glenn Easley  
Johns Hopkins University  
E-mail: glenn.easley@jhuapl.edu

---

## 1 Introduction

Sparse multiscale representations of functions have a long and celebrated history in applied harmonic analysis [18]. Their significance goes beyond data compression as discovering a sparse representation of signals in a certain class entails a deeper understanding of the class structure and can be exploited for tasks including signal denoising, inpainting and classification [4].

Following the success of sparse wavelet approximations of piecewise regular 1d signals, several methods were later proposed to extend the same approach to the multi-dimensional setting. The main challenge concerns the more complex nature of singularities occurring in higher dimensions. In dimensions 2 and 3, data found in most applications are dominated by curvilinear edges or surface discontinuities, and conventional wavelets are inefficient at representing such structures since it takes ‘many’ wavelet coefficients to achieve a good approximation. To overcome this limitation, a more sophisticated class of multiscale systems was introduced about 15 years ago, including most notably curvelets [2] and shearlets [17]. The geometric intuition underlying their construction is that, in order to provide more efficient representations of images with edges, the elements of the analyzing system should be defined not only over a range of locations and scales, as traditional wavelets, but also at multiple orientations and with highly anisotropic shapes. It was shown that shearlets and curvelets provide optimally sparse approximations, in a precise sense, of bivariate functions that are  $C^2$  regular away from  $C^2$  edges (the, so-called, *cartoon-like images*), outperforming separable wavelet bases [2, 8, 15, 22]. The same idea was extended to the 3d setting, showing that also in this case shearlet representations provide optimally sparse approximations of functions of 3 variables that are  $C^2$  away from  $C^2$  boundaries [9, 11, 14].

However, while the cartoon-like data model adopted in the 3d setting covers a large number of problems found in applications, it is not very convenient to handle a large class of 3d data dominated by surface singularities that are perpendicularly to the  $xy$  plane. Examples of such data include, for instance, hyperspectral images (HSI) from application in remote sensing and typical movie sequences. We recall that HSI data form 3d cubes where 2 variables encode the spatial information and the 3rd one is a spectral component; in such data, discontinuities typically occur in the spatial plane corresponding to physical structures such as roads and buildings [19, 21]. Similarly, a movie sequence forms a data cube where 2 variables are associated with the image plane while the 3rd one is time; the dominant discontinuities here are edges in the image plane. Motivated by these applications, we consider in this paper a simplified solid model where discontinuities are independent of the  $z$  variables. To represent such data we introduce a new multiscale representation of functions of 3 variables, called cylindrical shearlets, especially designed to handle the geometry of data containing surface singularities perpendicular to the  $xy$  plane. This new construction follows the general philosophy of shearlets whose analyzing elements consist of well-localized waveforms defined over a range of scales, locations and orientations. However, unlike conventional 3d shearlets

where orientations are controlled by two parameters and range over the entire sphere, cylindrical shearlets only include one orientation parameter and their directional sensitivity is limited to rotations on a circle. Our main result is that cylindrical shearlets provide optimally sparse representations over a class  $\mathcal{C}^2(Z)$  of bounded and compactly supported functions on  $\mathbb{R}^3$  that are  $C^2$  regular away from surface discontinuities orthogonal to the  $xy$  plane (as defined in Sec. 3), significantly outperforming both conventional wavelets and shearlets. We include numerical experiments illustrating the computational advantages of approximating data in the class  $\mathcal{C}^2(Z)$  using cylindrical shearlets, as compared with state-of-the-art multiscale representations. Remarkably, our numerical results suggest that our approach conventional 3d shearlets not only when discontinuities are independent of the  $z$  variable but also when their dependence on  $z$  is mild.

The ideas presented in this work rely on the microlocal properties of shearlets that have been investigated in several papers [10, 12, 16]. We recall that, in the 3d setting, different versions of shearlet transform have been already proposed to deal with different singularity types [13].

The paper is organized as follows. We present the construction of the new Parseval frame of cylindrical shearlets in Sec. 2. We report the main results and their proofs in Sec. 3 and the supporting numerical experiments in Sec. 4.

## 2 Cylindrical shearlets

Cylindrical shearlets were originally introduced in [1] as a variant of 3d shearlets targeted to applications where the dominant discontinuities are perpendicular to one of the coordinate axes. Unlike classical 3-dimensional shearlets that require to partition  $\mathbb{R}^3$  into 3 pyramidal regions, we associate cylindrical shearlets to two cylindrical pyramids  $\mathcal{P}_1$  and  $\mathcal{P}_2$  in  $\mathbb{R}^3$  defined as:

$$\mathcal{P}_1 = \{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : |\frac{\xi_2}{\xi_1}| \leq 1\}, \quad \mathcal{P}_2 = \{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : |\frac{\xi_1}{\xi_2}| \leq 1\}.$$

**Definition 1** For  $d = 1, 2$ , a *pyramid-based cylindrical shearlet system* associated with the pyramid  $\mathcal{P}_d$  is a collection of functions

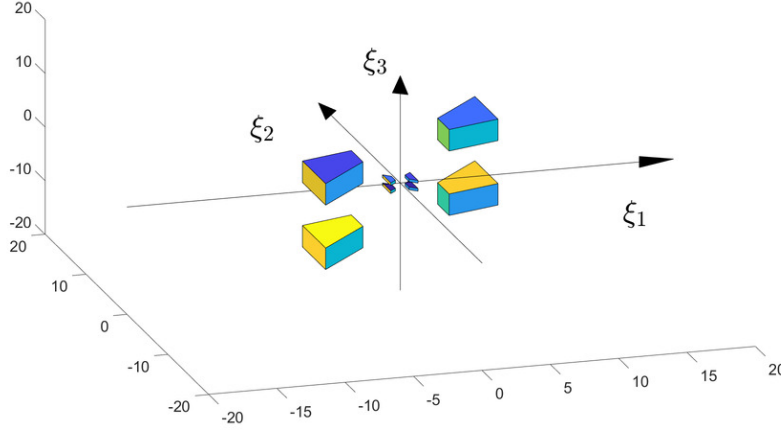
$$\{\psi_{j,\ell,k}^{(d)} : j \geq 0, |\ell| \leq 2^j, k \in \mathbb{Z}^3\}, \quad (1)$$

where the elements of the system (1) are given in the Fourier domain as

$$\hat{\psi}_{j,\ell,k}^{(d)}(\xi) = |\det A_{(d)}|^{-\frac{j}{2}} w(2^{-2j}\xi) V_{(d)}(\xi A_{(d)}^{-j} B_{(d)}^{-\ell}) e^{2\pi i \xi A_{(d)}^{-j} B_{(d)}^{-\ell} k} \quad (2)$$

and the matrices  $A_{(d)}$  and  $B_{(d)}$  are given by

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



**Fig. 1** Fourier support region  $U_{j,\ell}$  associated with a cylindrical shearlet system.

Similar to the standard shearlet construction [11, 7], we can choose the functions  $w$  and  $V$  and the elements in (1) to form a smooth Parseval frame of  $L^2(\mathbb{R}^3)$ . The construction we adopt here is different from [1] since our sparse approximation result requires the shearlet elements in (1) to be well localized.

We let  $\phi \in L^2(\mathbb{R}^3)$  be such that  $\hat{\phi} \in C_c^\infty$  with  $0 \leq \hat{\phi} \leq 1$  and

$$\hat{\phi}(\xi) = 1 \text{ if } \xi \in \left[-\frac{1}{16}, \frac{1}{16}\right]^3, \quad \hat{\phi}(\xi) = 0 \text{ if } \xi \in \mathbb{R}^3 \setminus \left[-\frac{1}{8}, \frac{1}{8}\right]^3. \quad (3)$$

Next we define

$$w(\xi) = \sqrt{\hat{\phi}^2(2^{-2}\xi) - \hat{\phi}^2(\xi)}. \quad (4)$$

It follows that

$$\hat{\phi}^2(\xi) + \sum_{j \geq 0} w^2(2^{-2j}\xi) = 1 \text{ for } \xi \in \mathbb{R}^3. \quad (5)$$

We notice that the functions  $w_j^2 = w^2(2^{-2j}\cdot)$  are supported on the sets  $C_j = [-2^{2j-1}, 2^{2j-1}]^3 \setminus [-2^{2j-4}, 2^{2j-4}]^3 \subset \mathbb{R}^3$ . In addition, we let  $v \in C^\infty(\mathbb{R})$  be such that  $\text{supp } v \subset [-1, 1]$  and

$$|v(u-1)|^2 + |v(u)|^2 + |v(u+1)|^2 = 1 \quad \text{for } |u| \leq 1. \quad (6)$$

For  $d = 1$ , since  $(\xi_1, \xi_2, \xi_3)A_{(1)}^{-j}B_{(1)}^{-\ell} = (2^{-2j}\xi_1, -2^{-2j}\ell\xi_1 + 2^{-j}\xi_2, 2^{-2j}\xi_3)$  and  $|\det A_{(1)}| = 2^5$ , an element of the system (2) can be written as

$$\hat{\psi}_{j,\ell,k}^{(1)}(\xi) = 2^{-\frac{5j}{2}} w(2^{-2j}\xi) v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k},$$

showing that the Fourier support of  $\psi_{j,\ell,k}^{(1)}$  is contained inside the region

$$U_{j,\ell} = \{\xi \in [-2^{2j-1}, 2^{2j-1}]^3 \setminus [-2^{2j-4}, 2^{2j-4}]^3 \subset \mathbb{R}^3 : \left|\frac{\xi_2}{\xi_1} - \ell 2^{-j}\right| \leq 2^{-j}\}.$$

A graphical illustration of the region  $U_{j,\ell}$  is shown in Fig. 1

Similar to conventional 3d shearlets [7], we obtain a smooth Parseval frame of cylindrical shearlets for  $L^2(\mathbb{R}^3)$  by combining the pyramid-based systems (1) with a coarse scale system. To ensure that all elements of this combined system are  $C_c^\infty$  in the frequency domain, we modify the elements of the shearlet system overlapping the boundaries of the regions  $\mathcal{P}_1$  and  $\mathcal{P}_2$  in the frequency domain. Hence, a *cylindrical shearlet system* for  $L^2(\mathbb{R}^3)$  is given by

$$\begin{aligned} & \left\{ \tilde{\psi}_{-1,k} : k \in \mathbb{Z}^3 \right\} \cup \left\{ \tilde{\psi}_{j,\ell,k,d} : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^3, d = 1, 2 \right\} \\ & \cup \left\{ \tilde{\psi}_{j,\ell,k} : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^3 \right\}, \end{aligned} \quad (7)$$

consisting of:

- the *coarse-scale cylindrical shearlets*  $\{\tilde{\psi}_{-1,k} = \phi(\cdot - k) : k \in \mathbb{Z}^3\}$ , where  $\phi$  is given by (3);
- the *interior cylindrical shearlets*  $\{\tilde{\psi}_{j,\ell,k,d} = \psi_{j,\ell,k}^{(d)} : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^3, d = 1, 2\}$ , where the functions  $\psi_{j,\ell,k}^{(d)}$ ,  $d = 1, 2$  are given by (1) ;
- the *boundary cylindrical shearlets*  $\{\tilde{\psi}_{j,\ell,k} : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^3\}$ , obtained by joining together slightly modified versions of  $\psi_{j,\ell,k}^{(1)}$  and  $\psi_{j,\ell,k}^{(2)}$ , for  $\ell = \pm 2^j$ . Namely, for  $j \geq 1, \ell = \pm 2^j, k \in \mathbb{Z}^3$ , we define

$$(\tilde{\psi}_{j,\ell,k})^\wedge(\xi) = \begin{cases} 2^{-\frac{5}{2}j - \frac{3}{2}} w(2^{-2j}\xi) v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \text{if } \xi \in \mathcal{P}_1 \\ 2^{-\frac{5}{2}j - \frac{3}{2}} w(2^{-2j}\xi) v\left(2^j \frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \text{if } \xi \in \mathcal{P}_2 \end{cases}$$

and, for  $j = 0, k \in \mathbb{Z}^3, \ell = \pm 1$ , we define

$$(\tilde{\psi}_{0,\ell,k})^\wedge(\xi) = \begin{cases} w(\xi) v\left(\frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_1 \\ w(\xi) v\left(\frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_2. \end{cases}$$

We have now the following result that we report without proof as its argument is very similar to Theorem 8 in [7].

**Theorem 1** *Let  $\phi \in L^2(\mathbb{R}^3)$ ,  $v \in C_c^\infty(\mathbb{R})$  and  $w \in C_c^\infty(\mathbb{R}^3)$  be defined as above (cf. (3), (4), (6)). The cylindrical shearlet system (7) is a Parseval frame for  $L^2(\mathbb{R}^3)$ . Furthermore, the elements of this system are  $C^\infty$  and compactly supported in the Fourier domain.*

For brevity, in the following, we denote the system (7) as

$$\{\tilde{\psi}_\mu : \mu \in M\}, \quad (8)$$

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

- $M_C = \{(j, k) : j = -1, k \in \mathbb{Z}^3\}$ ;
- $M_I = \{(j, \ell, k, d) : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^3, d = 1, 2\}$ ;

-  $M_B = \{(j, \ell, k) : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^3\}$ .

We also remark that, by direct computation, we can write the elements of the cylindrical shearlet system (1) in the space domain as

$$\psi_{j,\ell,k}^{(d)}(x) = |\det A_{(d)}|^{j/2} \psi_{j,\ell}^{(d)}\left(B_{(d)}^\ell A_{(d)}^j x - k\right),$$

where  $\hat{\psi}_{j,\ell}^{(d)}(\xi) = w(2^{2j} \xi B_{(d)}^\ell A_{(d)}^j) V_{(d)}(\xi)$ . Even though the system (1) is not generated by a finite set of generators by the action of the affine group, however the functions  $\psi_{j,\ell}^{(d)}$  depend very mildly on  $j, \ell$ . Using support and regularity properties of  $w$  and  $V$ , we have that, for any  $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in (\mathbb{N} \cup \{0\})^3$  and any  $N > 0$ , there is a constant  $C_{\gamma,N} > 0$  independent of  $j, \ell, d$  such that

$$\partial_x^\gamma \psi_{j,\ell}^{(d)}(x) \leq C_{\gamma,N} (1 + |x|)^{-N}. \quad (9)$$

### 3 Main results

Following [5], we consider  $STAR^2(Z)$ , a class of indicator functions of sets  $B$  with  $C^2$  boundaries  $\partial B$ . In polar coordinates, let  $\rho(\theta) : [0, 2\pi) \rightarrow [0, 1]^2$  be a radius function and define the set  $B = \{x \in \mathbb{R}^2 : |x| \leq \rho(\theta)\}$ . In particular, the boundary  $\partial B$  is given by the curve in  $\mathbb{R}^2$ :

$$\beta(\theta) = \begin{pmatrix} \rho(\theta) \cos(\theta) \\ \rho(\theta) \sin(\theta) \end{pmatrix}, \quad \theta \in [0, 2\pi). \quad (10)$$

Given a constant  $Z > 0$ , the class of boundaries of interest are defined by

$$\sup |\rho''(\theta)| \leq Z, \quad \rho \leq \rho_0 < 1. \quad (11)$$

We say that a set  $B \in STAR^2(Z)$  if  $B \subset [0, 1]^2$  and  $B$  is a translate of a set obeying (10) and (11). We let  $C_0^2([0, 1]^2)$  to be the collection of twice differentiable functions supported inside  $[0, 1]^2$ . Hence, the set of *cartoon-like images*  $\mathcal{E}^2(Z)$  is the collection of functions of the form  $h = h_0 \chi_B$  where  $h_0 \in C_0^2([0, 1]^2)$ ,  $B \in STAR^2(Z)$  and  $\sum_{|\alpha| \leq 2} \|D^\alpha f\|_\infty \leq 1$ . We next define the set of *cylindrical cartoon-like functions*  $\mathcal{C}^2(Z)$  of functions of the form

$$f(x, y, z) = h_0(x, y) g_0(z) + h_1(x, y) \chi_B(x, y) g_1(z). \quad (12)$$

where  $B \in STAR^2(Z)$ ,  $h_0, h_1 \in C_0^2([0, 1]^2)$ ,  $g_0, g_1 \in C_0^2([-1, 1])$ .

#### 3.1 Main theorems

Let  $\{\tilde{\psi}_\mu\}_{\mu \in M}$  be the Parseval frame of cylindrical shearlets given by (8). The *cylindrical shearlet coefficients* of a function  $f \in L^2(\mathbb{R}^3)$  are the elements of the sequence  $\{s_\mu(f) = \langle f, \tilde{\psi}_\mu \rangle : \mu \in M\}$ . We denote by  $|s(f)|_{(N)}$  the  $N$ -th largest entry in this sequence. We have the following results.

**Theorem 2** *Let  $f \in \mathcal{C}^2(Z)$  and  $\{s_\mu(f) : \mu \in M\}$  be the sequence of the corresponding cylindrical shearlet coefficients. Then, for  $N \in \mathbb{N}$  there is constant  $c$  independent of  $\mu$  and  $N$  such that*

$$\sup_{f \in \mathcal{C}^2(Z)} |s_\mu(f)|_{(N)} \leq c N^{-\frac{3}{2}} (\ln(N))^{\frac{3}{2}}.$$

Let  $f_N^S$  be the  $N$ -th term approximate of  $f \in \mathcal{C}^2(Z)$  obtained from the  $N$ -th largest coefficients of its cylindrical shearlet expansion, namely

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

where  $I_N \subset M$  is the set of indices corresponding to the  $N$ -th largest entries of the sequence  $\{|\langle f, \psi_\mu \rangle|^2 : \mu \in M\}$ . Then the approximate error satisfies

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

Therefore from Theorem 2 we obtain the following result.

**Theorem 3** *Let  $f \in \mathcal{C}^2(Z)$  and  $f_N^S$  be the  $N$ -term approximation to  $f$  defined above. Then, for  $N \in \mathbb{N}$  there is constant  $c$  independent of  $\mu$  and  $N$  such that*

$$\|f - f_N^S\|_{L^2}^2 \leq c N^{-2} (\ln(N))^3.$$

This decay rate is nearly optimal, in the sense that no other representation system can achieve an asymptotic approximation rate faster than  $N^{-2}$  for this class of functions. By contrast, separable 3d wavelets and conventional 3d shearlets only achieve approximation rates that are  $O(N^{-1/2})$  [3] and  $O(N^{-1})$  [11], respectively. The optimality argument is presented in Appendix A.

### 3.2 Arguments and constructions

The general organization of the proof follows the sparse shearlet approximation proof in [8, 11]. However, the arguments in the proof of Proposition 1 use a new idea to deal with the geometry of cylindrical shearlets.

We will use the weak- $\ell^p$  quasi-norm  $\|\cdot\|_{w\ell^p}$  which is useful to measure the sparsity of cylindrical shearlet coefficients for sequence  $s_\mu$  and is defined by

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

where  $|s_\mu|_{(N)}$  is the  $N$ -th largest entry in the sequence  $s = (s_\mu)$ . It is known that the above quantity is equivalent to

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

According with the cylindrical cartoon-like model (12), we consider functions of the form

$$f(x_1, x_2, x_3) = h_1(x_1, x_2) \chi_B(x_1, x_2) g(x_3),$$

where  $B \in STAR^2(Z)$ ,  $h_1 \in C_0^2([0, 1]^2)$  and  $g \in C_0^2([-1, 1])$ .

We remark that discontinuities only occur in the the  $x_1 x_2$  plane. Hence, to carry our our argument, we smoothly localize the function  $f$  near dyadic squares in the  $x_1 x_2$  plane as follows. For a scale parameter  $j \geq 0$  fixed, let  $\mathcal{Q}_j$  be the collections of dyadic squares of the form  $Q = [\frac{k_1}{2^j}, \frac{k_1+1}{2^j}] \times [\frac{k_2}{2^j}, \frac{k_2+1}{2^j}]$  with  $k_1, k_2 \in \mathbb{Z}^2$ . For a non-negative  $C^\infty$  function  $w$  with support in  $[-1, 1]^2$  we define the smooth partition of unity

$$\sum_{Q \in \mathcal{Q}_j} W_Q(x_1, x_2) = 1, \quad (x_1, x_2) \in \mathbb{R}^2,$$

where  $W_Q(x_1, x_2) = W(2^j x_1 - k_1, 2^j x_2 - k_2)$ .

Corresponding to each dyadic square  $Q$ , we will examine the cylindrical shearlet coefficients of the localized functions  $f_Q := f W_Q$ , that is, the terms  $\{\langle f_Q, \tilde{\psi}_\mu \rangle : \mu \in M_j\}$ , where  $\tilde{\psi}_\mu$  is an element of the set (8) and  $M_j$  denotes the collection of  $\mu \in M$  such that  $j$  is fixed. As we show below, these coefficients exhibit a different decay behaviour depending on whether  $\partial B$ , the boundary of  $B$ , intersects the support of  $W_Q$  or not. Hence, 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 dyadic squares  $Q \in \mathcal{Q}_j$  such that  $\partial B$  intersects the support of  $W_Q$ . Since each  $Q$  has side length  $2 \cdot 2^{-j}$ , then  $\mathcal{Q}_j^0$  has cardinality  $|\mathcal{Q}_j^0| \leq c_0 2^j$ , where  $c_0$  is independent of  $j$ . Since  $f$  is compactly supported in  $[0, 1]^2$ ,  $|\mathcal{Q}_j^1| \leq c_1 2^{2j}$ , where  $c_1$  is independent of  $j$ . We can now state the following propositions that are needed to prove Theorem 2.

**Proposition 1** *For  $Q \in \mathcal{Q}_j^0$  with  $j \geq 0$  fixed, the sequence of cylindrical shearlet coefficients  $\{\langle f_Q, \tilde{\psi}_\mu \rangle : \mu \in M_j\}$  satisfy the estimate*

$$\|\langle f_Q, \tilde{\psi}_\mu \rangle\|_{w\ell^{\frac{2}{3}}} \leq c 2^{-\frac{3}{2}j}$$

for some constant  $c$  independent of  $Q$  and  $j$ .

**Proposition 2** *For  $Q \in \mathcal{Q}_j^1$  with  $j \geq 0$  fixed, the cylindrical shearlet coefficients  $\{\langle f_Q, \tilde{\psi}_\mu \rangle : \mu \in M_j\}$  satisfy the estimate*

$$\|\langle f_Q, \tilde{\psi}_\mu \rangle\|_{w\ell^{\frac{2}{3}}} \leq c 2^{-3j}$$

for some constant  $c$  independent of  $Q$  and  $j$ .

Before proving Propositions 1 and 2, we show how these results are used to prove Theorem 2. First, we have the following corollary.

**Corollary 1** *Let  $f \in \mathcal{C}^2(Z)$ , and for  $j \geq 0$ , consider the sequence of cylindrical shearlet coefficients  $s_j(f) = \{\langle f, \tilde{\psi}_\mu \rangle : \mu \in M_j\}$ . Then there is a constant  $c$  independent of  $j$  such that  $\|s_j(f)\|_{w\ell^{\frac{2}{3}}} \leq c$ .*



*Proof.* Using Propositions 1 and 2 and the estimates on the cardinality of the sets  $\mathcal{Q}_j^0$  and  $\mathcal{Q}_j^1$ , we obtain

$$\begin{aligned} \|s_j(f)\|_{wl^{\frac{2}{3}}} &\leq \sum_{Q \in \mathcal{Q}_j} \|\langle f_Q, \tilde{\psi}_\mu \rangle\|_{wl^{\frac{2}{3}}} \\ &\leq \sum_{Q \in \mathcal{Q}_j^0} \|\langle f_Q, \tilde{\psi}_\mu \rangle\|_{wl^{\frac{2}{3}}} + \sum_{Q \in \mathcal{Q}_j^1} \|\langle f_Q, \tilde{\psi}_\mu \rangle\|_{wl^{\frac{2}{3}}} \\ &\leq c|\mathcal{Q}_j^0|2^{-\frac{3}{2}j} + C|\mathcal{Q}_j^1|2^{-3j} \leq c. \quad \square \end{aligned}$$

Now we can prove Theorem 2.

*Proof of Theorem 2.* By Corollary 1, we have that, for  $j \geq 0$ ,

$$R(j, \epsilon) := \#\{\mu \in M_j : |\langle f, \tilde{\psi}_\mu \rangle| > \epsilon\} \leq c\epsilon^{-\frac{2}{3}}. \quad (13)$$

Next observe that, for an interior cylindrical shearlet  $\tilde{\psi}_{j,k,\ell,d} = \psi_{j,k,\ell}^{(d)}$  given by (1), a direct calculation using (9) gives

$$\begin{aligned} |\langle f, \psi_{j,k,\ell}^{(d)} \rangle| &= \left| \int_{\mathbb{R}^3} f(x) 2^{\frac{5j}{2}} \psi_{j,\ell}^{(d)}(B_{(d)}^\ell(A_{(d)}^j x - k)) dx \right| \\ &\leq 2^{\frac{5j}{2}} \|f\|_\infty \int_{\mathbb{R}^3} |\psi_{j,\ell}^{(d)}(B_{(d)}^\ell(A_{(d)}^j x - k))| dx \\ &\leq 2^{-\frac{5j}{2}} \|f\|_\infty \int_{\mathbb{R}^3} |\psi_{j,\ell}^{(d)}(y)| dy \\ &\leq c2^{-\frac{5j}{2}}. \quad (14) \end{aligned}$$

A very similar computation on the boundary shearlets gives the same estimate.

It follows that, given any  $\epsilon > 0$ , there is scale index  $j_\epsilon > 0$  such that  $|\langle f, \tilde{\psi}_\mu \rangle| < \epsilon$  for each  $j \geq j_\epsilon$ . Thus, it follows from (14) that  $R(j, \epsilon) = 0$  for  $j > \frac{2}{5} \log_2(\epsilon^{-1}) + \log_2(c) > \frac{2}{5} \log_2(\epsilon^{-1})$ . Thus, using (13), we have that

$$\#\{\mu \in M_j : |\langle f, \tilde{\psi}_\mu \rangle| > \epsilon\} \leq \sum_{j=0}^{\frac{2}{5} \log_2(\epsilon^{-1})} R(j, \epsilon) \leq c\epsilon^{-\frac{2}{3}} \log_2(\epsilon^{-1}). \quad \square$$

### 3.3 Analysis of the edge fragments

Since the localization window  $W_Q$  acts in the  $x_1 x_2$  plane, it is convenient to analyze the localized discontinuity curve in this plane. For  $j \geq j_0$  sufficiently large, the scale  $2^{-j}$  is small enough so that over a square  $Q$  of side  $2^{-j}$ , the boundary curve  $\partial B$  can be parametrized as  $x_1 = E(x_2)$  or  $x_2 = E(x_1)$ . Without loss of generality, we may assume  $x_1 = E(x_2)$ . Also, by a translation, we may assume  $k = (0, 0)$  so that the function  $f_Q$  is localized on  $Q = [0, 2^{-j}]^2$ . Correspondingly, using the notation from Sec. 3.2, we define the *edge fragment*:

$$h_Q(x_1, x_2) = W(2^j(x_1, x_2))h_1(x_1, x_2)\chi_{[x_1 > E(x_2)]}(x_1, x_2), \quad (15)$$

where  $h_1 \in C_0^2([0, 1]^2)$  and  $x_1, x_2 \in [0, 2^{-j}]$ .

Next we derive an estimate for the elements of the Parseval frame of cylindrical shearlets against  $f_Q(x_1, x_2, x_3) = h_Q(x_1, x_2)g(x_3)$ , where  $h_Q$  is given by (15) and  $g \in C_0^2([0, 1])$ . We will only examine cylindrical shearlet coefficients associated with interior shearlets since boundary shearlets can be analyzed very similarly. In addition, since the behavior of cylindrical shearlets in the pyramidal regions  $\mathcal{P}_1$  and  $\mathcal{P}_2$  is very similar, we only consider the cylindrical shearlets in  $\mathcal{P}_1$ . For  $\xi = (\xi_1, \xi_2, \xi_3) \in \mathcal{P}_1$ ,  $j \geq 0$  and  $|\ell| \leq 2^j$ , letting

$$\Gamma_{j,\ell}(\xi) = w(2^{-2j}\xi) v(2^j \frac{\xi_2}{\xi_1} - \ell) \quad (16)$$

we can write the interior cylindrical shearlets (2) associated with  $\mathcal{P}_1$  as

$$\hat{\psi}_{j,k,\ell}^{(1)} = 2^{-\frac{5j}{2}} \Gamma_{j,\ell}(\xi) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}.$$

We need the following lemmata, where the first one is Corollary 2.4 in [8] and the second one is a simple extension of Lemma 2.5 in [8].

**Lemma 1** *Let  $h_Q$  be the edge fragment (15),  $j \geq 1$ , and  $\theta \in [0, 2\pi]$ . For  $m = (m_1, m_2) \in (\{0\} \cup \mathbb{N}) \times (\{0\} \cup \mathbb{N})$ , we have*

$$\begin{aligned} & \int_{2^{2j-4}}^{2^{2j+2}} \left| \frac{\partial^{m_1}}{\partial \xi_1^{m_1}} \left[ \frac{\partial^{m_2}}{\partial \xi_2^{m_2}} \hat{h}_Q(\lambda \cos \theta, \lambda \sin \theta) \right] \right|^2 d\lambda \\ & \leq c_m 2^{-2j(m_1+m_2)} \left( 2^{-(4+2m_1)j} (1 + 2^j |\sin \theta|)^{-5} + 2^{-10j} \right), \end{aligned}$$

where the constant  $c_m$  is independent of  $j, \theta$ .

**Lemma 2** *Let  $\Gamma$  be given by (16),  $j \geq 1$ ,  $|\ell| \leq 2^j$ . For  $m = (m_1, m_2, m_3) \in (\{0\} \cup \mathbb{N}) \times (\{0\} \cup \mathbb{N}) \times (\{0\} \cup \mathbb{N})$ , we have*

$$\left| \frac{\partial^{m_1}}{\partial \xi_1^{m_1}} \frac{\partial^{m_2}}{\partial \xi_2^{m_2}} \frac{\partial^{m_3}}{\partial \xi_3^{m_3}} [\Gamma_{j,\ell}] \right| \leq c_m 2^{-j(2m_1+m_2+2m_3)} (1 + |\ell|)^{m_1}$$

where the constant  $c_m$  is independent of  $j, \ell$ .

Since  $g \in C_0^2([0, 1])$ , Lemmata 1 and 2 imply the following key estimate.

**Lemma 3** *Let  $h_Q$  be the edge fragment (15),  $\Gamma$  be given by (16),  $g \in C_0^2([0, 1])$ ,  $j \geq 0$ ,  $|\ell| \leq 2^j$ , and  $L$  be the differential operator:*

$$L = \left( I - \left( \frac{2^{2j}}{2\pi(1+|\ell|)} \right)^2 \frac{\partial^2}{\partial \xi_1^2} \right) \left( I - \left( \frac{2^j}{2\pi} \right)^2 \frac{\partial^2}{\partial \xi_2^2} \right) \left( I - \left( \frac{2^{2j}}{2\pi} \right)^2 \frac{\partial^2}{\partial \xi_3^2} \right). \quad (17)$$

Then there is constant  $c$  independent of  $j, \ell$  such that

$$\int_{\hat{\mathbb{R}}^3} |L(\hat{h}_Q(\xi_1, \xi_2) \hat{g}(\xi_3) \Gamma_{j,\ell}(\xi))|^2 d\xi \leq c 2^{-3j} (1 + |\ell|)^{-5}.$$

### 3.4 Proof of Proposition 1

With the notation above, we write  $f_Q(x_1, x_2, x_3) = h_Q(x_1, x_2)g(x_3)$  and

$$\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle = |\det A_{(1)}|^{-\frac{j}{2}} \int_{\mathbb{R}^3} \hat{h}_Q(\xi_1, \xi_2) \hat{g}(\xi_3) \Gamma_{j,\ell}(\xi_1, \xi_2, \xi_3) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\xi,$$

where  $\xi = (\xi_1, \xi_2, \xi_3)$ ,  $\Gamma_{j,\ell}$  is given by (16) and  $A_{(1)}, B_{(1)}$  are defined in (1). Observing that

$$\begin{aligned} 2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k &= 2\pi i \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 \end{pmatrix} \begin{pmatrix} 2^{-2j} & 0 & 0 \\ 0 & 2^{-j} & 0 \\ 0 & 0 & 2^{-2j} \end{pmatrix} \begin{pmatrix} 1 - \ell & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} \\ &= 2\pi i ((k_1 - \ell k_2) 2^{-2j} \xi_1 + 2^{-j} k_2 \xi_2 + 2^{-2j} k_2 \xi_3), \end{aligned}$$

a direct computation gives that

$$\begin{aligned} \left( \frac{\partial}{\partial \xi_1} (2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k) \right)^2 &= -(2\pi)^2 2^{-4j} (k_1 - k_2 \ell)^2 \\ &= \begin{cases} -(2\pi)^2 \ell^2 2^{-4j} \left(\frac{k_1}{\ell} - k_2\right)^2 & \ell \neq 0 \\ -(2\pi)^2 2^{-4j} k_1^2 & \ell = 0 \end{cases} \\ \left( \frac{\partial}{\partial \xi_2} (2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k) \right)^2 &= -(2\pi)^2 2^{-2j} k_2^2 \\ \left( \frac{\partial}{\partial \xi_3} (2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k) \right)^2 &= -(2\pi)^2 2^{-4j} k_3^2 \end{aligned}$$

Let  $L$  be the second order partial differential operator in (17). Hence:

$$\begin{aligned} &L(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}) \\ &= \begin{cases} \left(1 + \left(\frac{\ell}{(1+|\ell|)}\right)^2 \left(\frac{k_1}{\ell} - k_2\right)^2\right) (1 + k_2^2)(1 + k_3^2) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} & \ell \neq 0 \\ (1 + k_1^2)(1 + k_2^2)(1 + k_3^2) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} & \ell = 0 \end{cases} \quad (18) \end{aligned}$$

Integrating by parts we get

$$\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle = |\det A_{(1)}|^{-\frac{j}{2}} \int_{\mathbb{R}^3} L(\hat{f}_Q(\xi_1, \xi_2, \xi_3) \Gamma_{j,\ell}(\xi)) L^{-1}(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}) d\xi.$$

**Case 1:  $\ell \neq 0$ .** By (18) we have

$$L^{-1}(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}) = G(k, \ell)^{-1} e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k},$$

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

$$\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle = |\det A_{(1)}|^{-\frac{j}{2}} G(k, \ell)^{-1} \int_{\mathbb{R}^3} L(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi)) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\xi$$

or, equivalently,

$$\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle G(k, \ell) = |\det A_{(1)}|^{-\frac{j}{2}} \int_{\mathbb{R}^3} L(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi)) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\xi.$$

For  $K = (K_1, K_2, K_3) \in \mathbb{Z}^3$  define  $R_K = \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : \frac{k_1}{\ell} \in [K_1, K_1 + 1], k_2 = K_2, k_3 = K_3\}$ . For  $j, \ell$  fixed, the set  $\{|\det A_{(1)}|^{-\frac{j}{2}} e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} : k \in \mathbb{Z}^3\}$  is an orthonormal basis of the  $L^2$  functions supported on the set  $[-\frac{1}{2}, \frac{1}{2}]^3 A_{(1)}^j B_{(1)}^\ell$ . Hence, observing that  $\Gamma_{j,\ell}(\xi)$  is supported on this set, by Plancherel theorem we have

$$\sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle G(k, \ell)|^2 \leq \int_{\mathbb{R}^3} |L(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi))|^2 d\xi.$$

By the definition of  $R_K$  there is a constant  $c$  independent of  $j, \ell$  such that

$$\sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^2 \leq c L_K^{-2} \int_{\mathbb{R}^3} |L(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi))|^2 d\xi,$$

where  $L_k = (1 + (K_1 - K_2)^2)(1 + K_2^2)(1 + K_3^2)$ . By Lemma 3, it follows that

$$\sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^2 \leq c L_K^{-2} 2^{-3j} (1 + |\ell|)^{-5}. \quad (19)$$

Letting  $N_{j,\ell,K}(\epsilon) = \#\{k \in R_K : |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle| > \epsilon\}$ , we have that  $N_{j,\ell,K}(\epsilon) \leq c(1 + |\ell|)$  and, by (19),  $N_{j,\ell,K}(\epsilon) \leq c L_K^{-2} 2^{-3j} \epsilon^{-2} (1 + |\ell|)^{-5}$ . Thus

$$N_{j,\ell,K}(\epsilon) \leq c \min((1 + |\ell|), L_K^{-2} 2^{-3j} \epsilon^{-2} (1 + |\ell|)^{-5}). \quad (20)$$

Using (20), we can show that

$$\sum_{\ell=-2^j}^{2^j} N_{j,\ell,K}(\epsilon) \leq c L_K^{-\frac{2}{3}} 2^{-j} \epsilon^{-\frac{2}{3}}. \quad (21)$$

In fact, let  $\ell^*$  be defined by  $(\ell^* + 1) = L_K^{-2} 2^{-3j} \epsilon^{-2} (1 + \ell^*)^{-5}$ . So we have that  $(\ell^* + 1)^2 = L_K^{-\frac{2}{3}} 2^{-j} \epsilon^{-\frac{2}{3}}$ . By direct calculation

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

which proves (21). Since  $\sum_{k \in \mathbb{Z}^3} L_K^{-\frac{2}{3}} < \infty$ , using (21), we then have

$$\begin{aligned} \#\{\mu \in M_j : |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle| > \epsilon\} &\leq \sum_{K \in \mathbb{Z}^3} \sum_{\ell=-2^j}^{2^j} N_{j,\ell,K}(\epsilon) \\ &\leq c 2^{-j} \epsilon^{-\frac{2}{3}} \sum_{k \in \mathbb{Z}^3} L_K^{-\frac{2}{3}} \\ &\leq c 2^{-j} \epsilon^{-\frac{2}{3}}. \end{aligned}$$

Hence  $\|\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle\|_{w\ell^{\frac{2}{3}}} \leq c 2^{-\frac{3}{2}j}$  and this complete the proof in the case  $\ell \neq 0$ .

**Case 2:**  $\ell = 0$ . In this case, using the operator  $L$  given by (18), we have

$$L^{-1}(e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}) = L_k^{-1} e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}$$

where  $L_k = (1 + k_1^2)(1 + k_2^2)(1 + k_3^2)$ . Then it is clear that  $\sum_{k \in \mathbb{Z}^3} L_k^{-\frac{2}{3}} < \infty$ . An application of integration by parts similar to the one above gives that

$$\langle f_Q, \psi_{j,0,k}^{(1)} \rangle = |\det A_{(1)}|^{-\frac{j}{2}} L_k^{-1} \int_{\mathbb{R}^3} L(\hat{f}(\xi) \Gamma_{j,\ell}(\xi)) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\xi.$$

It follows that there is a constant  $c$  independent of  $j, \ell$  such that

$$\sum_{k \in \mathbb{Z}^3} |\langle f_Q, \psi_{j,0,k}^{(1)} \rangle|^2 L_k^2 = |\det A_{(1)}|^{-j} \int_{\mathbb{R}^3} |L(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi))|^2 d\xi \leq c 2^{-5j}.$$

In particular, for each  $k \in \mathbb{Z}^3$ , we have  $|\langle f_Q, \psi_{j,0,k}^{(1)} \rangle| \leq c L_k^{-1} 2^{-\frac{5}{2}j}$  and, hence,

$$\sum_{k \in \mathbb{Z}^3} |\langle f_Q, \psi_{j,0,k}^{(1)} \rangle| \leq \sum_{k \in \mathbb{Z}^3} c L_k^{-1} 2^{-\frac{5}{2}j} \leq c 2^{-\frac{5}{2}j} \sum_{k \in \mathbb{Z}^3} L_k^{-1} \leq c 2^{-\frac{5}{2}j} \leq c 2^{-\frac{3}{2}j}.$$

This completes the proof when  $j$  is large ( $j \geq j_0$ , as in the definition of the edge fragment). For  $j < j_0$ , the simpler argument is very similar to [11, Sec. 4.6].

### 3.5 Proof of Proposition 2

Similar to the proof of Proposition 1, using the notation from Sec. 3.2, we let  $\tilde{h}_Q(x_1, x_2) = h_1(x_1, x_2) W_Q(x_1, x_2)$ , where now  $Q \in \mathcal{Q}_j^1$ . With this notation, we write the localized function as  $f_Q(x_1, x_2, x_3) = \tilde{h}_Q(x_1, x_2) g(x_3)$ .

To carry out this proof, we need the following lemma ([8], Lemma 2.8).

**Lemma 4** *Let  $\tilde{h}_Q$  be given as above where  $Q \in \mathcal{Q}_j^1$ , let  $\Gamma_{j,\ell}$  be given by (16) and set  $T = (1 - \frac{2^j}{(2\pi)^2} \Delta)$ , where  $\Delta = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2}$ . Then*

$$\int_{\mathbb{R}^2} \sum_{|\ell| \leq 2^j} |T^2(\hat{\tilde{h}}_Q \Gamma_{j,\ell})(\xi)| d\xi \leq c 2^{-10j},$$

where the constant  $c$  is independent of  $j, \ell$ .

From Lemma 4 we directly derive the following estimate.

**Lemma 5** *Let  $f_Q$  be given as above where  $Q \in \mathcal{Q}_j^1$ ,  $\Gamma_{j,\ell}$  be given by (16),  $T$  be as in Lemma 4 and set  $\Lambda = I - (\frac{2^{2j}}{2\pi})^2 \frac{\partial^2}{\partial \xi_3^2}$ . Then*

$$\sum_{|\ell| \leq 2^j} \int_{\mathbb{R}^3} |T^2 \Lambda^2(\hat{f}_Q \Gamma_{j,\ell})(\xi)| d\xi \leq c 2^{-10j},$$

where the constant  $c$  is independent of  $j, \ell$ .

As in the proof of Proposition 1, we only examine the interior cylindrical shearlet (1) in the pyramidal region  $\mathcal{P}_1$ . A direct calculation shows that

$$(T^2 \Lambda^2)^{-1} e^{2\pi i \xi A^{-j} B^{-\ell} k} = ((1 + 2^{-2j}(k_1 - k_2 \ell))^2 + k_2^2)^{-2} (1 + k_3^2)^{-1} e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}.$$

For a fixed  $j \geq 0$ , using integration by parts we have

$$\begin{aligned} \langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle &= |\det A_{(1)}|^{-\frac{j}{2}} ((1 + 2^{-2j}(k_1 - k_2 \ell))^2 + k_2^2)^{-2} (1 + k_3^2)^{-1} \times \\ &\quad \int_{\mathbb{R}^3} \left( T^2 \Lambda^2(\hat{f}_Q(\xi) \Gamma_{j,\ell})(\xi) \right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\xi. \end{aligned}$$

Let  $K = (K_1, K_2, K_3) \in \mathbb{Z}^3$  and define the set  $R_K = \{(k_1, k_2, k_3) \in \mathbb{Z}^3, \text{ such that } k_2 = K_2, k_3 = K_3 \text{ and } k_1 \text{ satisfies } 2^{-j}(k_1 - K_2 \ell) \in [K_1, K_1 + 1]\}$ . For each fixed  $K$  and  $\ell$ , there are about  $2^{2j}$  choices for  $k_1$  in  $R_K$  (since  $|\ell| \leq 2^j$ ). Hence the cardinality of  $R_K$  is bounded by  $2^{2j}$ . Similar to Proposition 1, for fixed  $j, \ell$ , the set  $\{|\det A_{(1)}|^{-\frac{j}{2}} e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k} : k \in \mathbb{Z}^3\}$  is an orthonormal basis for the  $L^2$  functions with support on  $[-\frac{1}{2}, \frac{1}{2}]^3 A_{(1)}^j B_{(1)}^\ell$ . Plancherel theorem implies that there is a constant  $c$  independent of  $j, \ell$  such that

$$\sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^2 \leq c (1 + K_1^2 + K_2^2)^{-4} (1 + K_3^2)^{-2} \int_{\mathbb{R}^3} |T^2 \Lambda^2(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi))|^2 d\xi.$$

From this inequality, using Lemma 5, we have

$$\begin{aligned} \sum_{\ell=-2^j}^{2^j} \sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^2 &\leq c (1 + |K|^2)^{-4} \int_{\mathbb{R}^3} \sum_{\ell=-2^j}^{2^j} |T^2 \Lambda^2(\hat{f}_Q(\xi) \Gamma_{j,\ell}(\xi))|^2 d\xi \\ &\leq c (1 + K_1^2 + K_2^2)^{-4} (1 + K_3^2)^{-4} 2^{-10j}. \end{aligned} \quad (22)$$

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

$$\sum_{m=1}^N |a_m|^{\frac{2}{3}} \leq \left( \sum_{m=1}^N |a_m|^2 \right)^{\frac{1}{3}} N^{\frac{2}{3}} \quad (23)$$

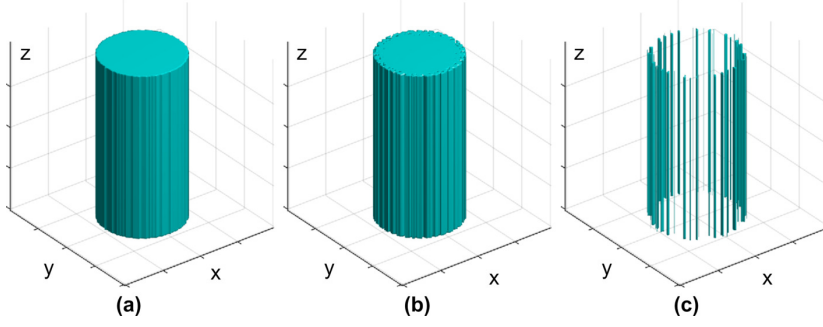
Since the cardinality of  $R_K$  is  $2^{2j}$ , it follows from (22) and (23) that

$$\sum_{\ell=-2^j}^{2^j} \sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^{\frac{2}{3}} \leq c (2^{2j})^{\frac{2}{3}} (1 + K_1^2 + K_2^2)^{-\frac{4}{3}} (1 + K_3^2)^{-\frac{2}{3}} 2^{-\frac{10}{3}j}.$$

It follows that

$$\|\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle\|_{w\ell^{\frac{2}{3}}}^{\frac{2}{3}} \leq \|\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle\|_{\ell^{\frac{2}{3}}}^{\frac{2}{3}} = \sum_{K \in \mathbb{Z}^3} \sum_{\ell=-2^j}^{2^j} \sum_{k \in R_K} |\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle|^{\frac{2}{3}} \leq c2^{-2j}.$$

Thus  $\|\langle f_Q, \psi_{j,k,\ell}^{(1)} \rangle\|_{w\ell^{\frac{2}{3}}} \leq c2^{-3j}$  and this finishes the proof of Proposition 2.



**Fig. 2** Reconstruction of a digitized cartoon-like cylindrical solid. (a) Original solid (size  $128 \times 128 \times 128$ ), (b) corresponding reconstruction using cylindrical shearlets and (c) the difference of the reconstruction from the original solid. The reconstructed solid was generated by taking the  $N$  largest coefficients in magnitude with  $N = 4192256$ . The Relative Mean Square Error of the reconstruction is 0.020.

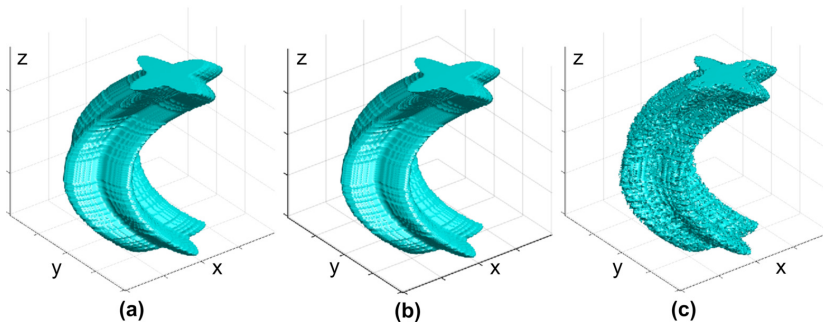
#### 4 Numerical implementation of cylindrical shearlet

In this section, we illustrate the numerical advantages of using cylindrical shearlets rather than conventional 3d shearlets when processing data whose geometry is consistent with the model of cylindrical cartoon-like functions. For that, we developed a numerical implementation of the *cylindrical shearlet transform*  $f \mapsto \langle f, \tilde{\psi}_\mu \rangle$ , where  $\{\tilde{\psi}_\mu : \mu \in M\}$  is the system (8).

Our new numerical implementation adapts the algorithmic pipeline in [6, 20] which was originally introduced for conventional shearlets and includes the following steps.

- (1) Application of the 3d Laplacian pyramid transform to decompose the input image  $f$  into a low-pass  $f_d$  and a high-pass component  $f_a$ .
- (2) Computation of the Discrete Fourier Transform  $\hat{f}_a$  of the high-pass component of the image on a pseudo-polar grid, obtaining  $P\hat{f}_a$ .
- (3) Application of band pass filters to  $P\hat{f}_a$  in order to achieve the partition into appropriate directional subbands.

The main difference with respect to the 3d discrete shearlet transform in [20] is the selection of the directional subbands since the new transform applies



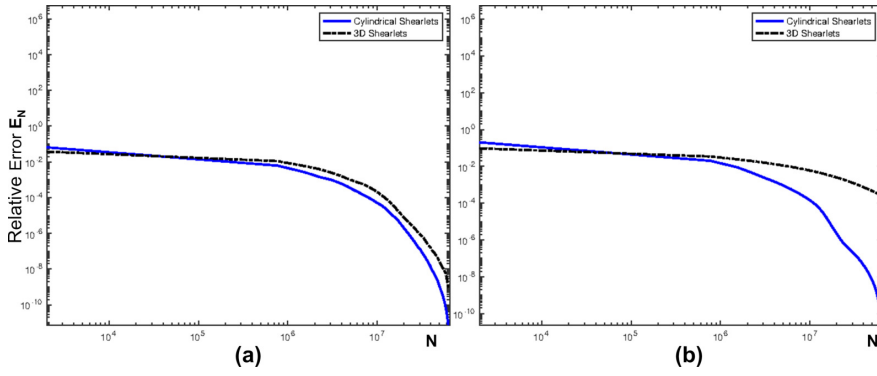
**Fig. 3** Reconstruction of a digitized solid. (a) Original solid (size  $128 \times 128 \times 128$ ) and corresponding reconstructions using (b) cylindrical shearlets and (c) conventional 3d shearlets. The reconstructed solid was generated by taking the  $N$  largest coefficients in magnitude with  $N = 1638400$ . The Relative Mean Square Error of the reconstructions in panels (b) and (c) are 0.0032 and 0.058, respectively.

directional filtering with respect to one angular rotation only (as illustrated in Fig. 1) unlike the former one where directional filtering is associated with two cone regions and two orientations.

To test our code, we first considered a discrete cylindrical region with vertical  $z$  axis as shown in Fig. 2. The figure displays the  $N$ -term reconstruction of the digitized solid using our discrete cylindrical shearlet transform and the corresponding error. To further illustrate the approximation properties of the cylindrical shearlet transform, we also considered a more complex discrete solid generated by taking a star-shape region in the horizontal plane and moving it along a spiral path, as shown in Fig. 3. Even though the surface boundary of this solid is not independent of  $z$  and, hence, the solid does not satisfy the definition of a cylindrical cartoon-like function, yet the dependence on  $z$  is mild (the solid can be locally approximated as a cylindrical cartoon-like function). Fig. 3 shows the  $N$ -term reconstruction of this digitized solid using our discrete cylindrical shearlet transform and compares it to the corresponding  $N$ -term reconstruction obtained using conventional 3d discrete shearlets for which we used the implementation from [20]. The quality of reconstruction using cylindrical shearlets is clearly superior to conventional 3d shearlets.

To illustrate the approximation properties of the discrete cylindrical shearlet transform, we plotted in Fig. 4 the relative approximation error  $\|f - f_N\|_2^2 / \|f\|_2^2$  where  $f$  is the solid in Fig. 2 or Fig. 3 and  $f_N$  is the  $N$ -term approximation of  $f$  obtained from the  $N$  largest expansion coefficients in magnitude. Fig. 4 shows that, as compared with a similar  $N$ -term approximation using conventional 3d shearlets, the decay rate of the error observed for cylindrical shearlets is significantly faster; this behavior is consistently with our theoretical prediction.





**Fig. 4** Reconstruction relative error  $E_N = \frac{\|f - f_N\|_2^2}{\|f\|_2^2}$  of (a) the solid  $f$  in Fig. 2 and (b) the solid  $f$  in Fig. 3, where  $f_N$  is the  $N$ -term approximation of  $f$  obtained using discrete cylindrical shearlets (solid line) or conventional 3d shearlets (dashed line). The plots are displayed with logarithmic axes.

## A Optimal Approximation Rates

Our analysis of the optimal sparsity rate adapts the method in [5] where, for functions  $f$  in a function class  $\mathcal{F}$ , one considers adaptive representations in an overcomplete dictionary  $\Phi = \{\phi_i : i \in I\} \subset L^2(\mathbb{R}^2)$  of the form

$$f = \sum_{i \in I_f} c_i(f) \phi_i, \quad (24)$$

and the selection of  $I_f$  in  $I$  is required to satisfy a *polynomial depth search* constraint to avoid situations which are computationally unfeasible. The sparsity of the expansion (24) 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 defined as the smallest  $p$  such that  $\|c(f)\|_{w\ell^p}$  is bounded. Hence, denoting by  $|c(f)|_m$  the  $m$ -th largest entry in the coefficient sequence  $(|c(f)|)$ , there is a constant  $C > 0$  such that

$$\sup_{f \in \mathcal{F}} |c(f)|_m \leq C m^{-\frac{1}{p}}, \quad (25)$$

and no decay rate faster than  $m^{-\frac{1}{p}}$  is possible. It follows that, if  $\Phi$  is also a Parseval frame, then, from (25) we have

$$\|f - f_N\|^2 \leq \sum_{m > N} |c(f)|_m^2 \leq C \sum_{m > N} m^{-2/p} \leq C N^{-2/p+1}$$

and  $O(N^{-2/p+1})$  is the optimal decay rate as no better approximation can be achieved under the procedure described above.

The argument in [5] to determine the optimal  $p$  for which  $\|c(f)\|_{w\ell^p}$  is bounded for  $f \in \mathcal{F}$  requires to assess the value  $p$  such that  $\mathcal{F}$  contains a copy of  $\ell^p$ . Recall that  $\mathcal{F}$  contains a copy of  $\ell^p$  if it contains embedded orthogonal hypercubes of dimension  $M(\Delta)$  and side  $\Delta$  such that, for some sequence  $(\Delta_k) \rightarrow 0$ , there is a constant  $C > 0$  such that

$$M(\Delta_k) \geq C \Delta_k^{-p}, \quad k = k_0, k_0 + 1, \dots$$

**Theorem 4** *The class  $\mathcal{C}^2(Z) \subset L^2(\mathbb{R}^3)$  contains a copy of  $\ell^p$  for  $p = \frac{2}{3}$ .*

Theorem 4 implies that no representation system satisfying polynomial depth search constraint can provide approximations for  $C^2(Z)$  with the coefficients  $\|c(f)\|_{w\ell^p} < \infty$ , for  $p < \frac{2}{3}$ , i.e.,  $p = \frac{2}{3}$  is the optimal value. It follows that, if  $\|c(f)\|_{w\ell^{\frac{2}{3}}} < \infty$ , then there is a constant  $C > 0$  such that

$$\sup_{f \in C^2} |c(f)|_m \leq C m^{-\frac{3}{2}},$$

and no decay rate faster than  $m^{-\frac{3}{2}}$  is possible. As seen above, the last inequality implies that, if  $f_N$  is the best  $N$  term approximation to  $f \in C^2(Z)$  using a Parseval frame, then

$$\|f - f_N\|^2 \leq C \sum_{m > N} m^{-3} \leq C N^{-2}.$$

Hence Theorem 3 shows that the approximation rate of cylindrical shearlets is nearly optimal. Note the rate  $O(N^{-2})$  is also the optimal approximation rate for functions in the class of 2d cartoon like images [5].

**Proof of Theorem 4.** Since our proof follows closely the proof of Thm. 3 in [5], we will mainly emphasize the modifications needed for our case.

Let  $g$  be a smooth and nonnegative function with compact support in  $[0, 2\pi]$ . For scalars  $A$  and  $m(A, \delta)$  to be determined, let  $g_{i,m}(t) = A m^{-2} g(mt - 2\pi i)$  for  $i = 0, 1, \dots, m-1$ . Notice that  $\|g_{i,m}\|_{C^2} = A \|g\|_{C^2}$  and  $\|g_{i,m}\|_{L^1} = A m^{-3} \|g\|_{L^1}$ . We introduce polar coordinates  $(\rho, \theta)$  with origin in  $(\frac{1}{2}, \frac{1}{2})$ . For  $\rho_0 = \frac{1}{4}$ , using cylindrical coordinates in  $\mathbb{R}^3$ , we set

$$\psi_{i,m}(\rho, \theta, z) = \left( \chi_{\{\rho \leq \rho_0\}}(\rho, \theta) - \chi_{\{\rho \leq g_{i,m} + \rho_0\}}(\rho, \theta) \right) w(z), \quad i = 0, 1, \dots, m-1,$$

where  $w$  is a  $C^2$  function with compact support. Hence, we define the radius functions  $r_\xi = \frac{1}{4} + \sum_{i=1}^m \xi_i g_{i,m}$ , where  $\xi_i \in \{0, 1\}$  and the corresponding functions

$$f_\xi(\rho, \theta, z) = \chi_{\{\rho \leq \rho_0\}}(\rho, \theta) w(z) + \sum_{i=1}^m \xi_i \psi_{i,m}(\rho, \theta, z), \quad \xi_i \in \{0, 1\}.$$

The functions  $\psi_{i,m}$  are bulges around a cylinder of radius  $\rho_0$  and have disjoint support; each  $f_\xi$  is the indicator function of the cylinder of radius  $\rho_0$  plus some addition bulges. Since  $g$  is bounded and nonnegative, a direct calculation shows that

$$\|\psi_{i,m}\|_{L^2}^2 \simeq \|g_{i,m}\|_{L^1} = A m^{-3} \|g\|_{L^1},$$

and, for each radius function  $r_\xi$ ,  $\|r_\xi\|_{C^2} \leq \|g_{i,m}\|_{C^2} = A \|g\|_{C^2}$ . Hence, the hypercube embedding is achieved whenever  $A \leq Z/\|g\|_{C^2}$ .

Whenever  $A \leq Z/\|g\|_{C^2}$ , the sidelength  $\Delta = \|\psi_{i,m}\|_{L^2}$  of the hypercubes satisfies:

$$\|\psi_{i,m}\|_{L^2}^2 = \Delta^2 \simeq \|g_{i,m}\|_{L^1} = A m^{-3} \|g\|_{L^1} \leq Z m^{-3} \frac{\|g\|_{L^1}}{\|g\|_{C^2}}.$$

Hence, setting

$$m(\delta) = \lfloor \left( \frac{\delta^2 \|g\|_{C^2}}{Z \|g\|_{L^1}} \right)^{-\frac{1}{3}} \rfloor, \quad A(\delta, Z) = \delta^2 m^3 / \|g\|_{L^1},$$

it follows that  $A \leq Z/\|g\|_{C^2}$  and  $\Delta \simeq \delta$ , which shows that the hypercube embedding is satisfied with sidelength  $\Delta \simeq \delta$  and the dimension of the hypercube obeys  $m(\delta) \geq K Z^{\frac{1}{3}} \delta^{-\frac{2}{3}}$ , for all  $0 < \delta < \delta_0$ , where  $\delta_0$  is the solution of

$$2 = \left( \frac{\delta_0^2 \|g\|_{C^2}}{Z \|g\|_{L^1}} \right)^{-1/3}, \quad K = \frac{1}{2} \left( \frac{\|g\|_{C^2}}{\|g\|_{L^1}} \right)^{-1/3}.$$

## References

1. Bernhard G. Bodmann, Demetrio Labate, and Basanta R. Pahari. Smooth projections and the construction of smooth parseval frames of shearlets. *Advances in Computational Mathematics*, 45(5):3241–3264, Dec 2019.
2. Emmanuel J Candès and David L Donoho. New tight frames of curvelets and optimal representations of objects with piecewise c2 singularities. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 57(2):219–266, 2004.
3. Ronald A DeVore. Nonlinear approximation. *Acta numerica*, 7:51–150, 1998.
4. David L. Donoho, Martin Vetterli, Ronald A. DeVore, and Ingrid Daubechies. Data compression and harmonic analysis. *IEEE Transactions on Information Theory*, 44(6):2435–2476, 1998.
5. David Leigh Donoho. Sparse components of images and optimal atomic decompositions. *Constructive Approximation*, 17(3):353–382, 2001.
6. Glenn Easley, Demetrio Labate, and Wang-Q Lim. Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1):25–46, 2008.
7. K. Guo and D. Labate. The construction of smooth parseval frames of shearlets. *Mathematical Modelling of Natural Phenomena*, 8(1):82–105, 2013.
8. Kanghui Guo and Demetrio Labate. Optimally sparse multidimensional representation using shearlets. *SIAM J. Math. Analysis*, 39:298–318, 2007.
9. Kanghui Guo and Demetrio Labate. Optimally sparse shearlet approximations of 3d data. *Independent Component Analyses, Wavelets, Neural Networks, Biosystems, and Nanoengineering IX*, 8058, 2011.
10. Kanghui Guo and Demetrio Labate. Characterization of piecewise-smooth surfaces using the 3d continuous shearlet transform. *Journal of Fourier Analysis and Applications*, 18(3):488–516, 2012.
11. Kanghui Guo and Demetrio Labate. Optimally sparse representations of 3d data with c2 surface singularities using parseval frames of shearlets. *SIAM J. Math. Analysis*, 44:851–886, 2012.
12. Kanghui Guo and Demetrio Labate. Detection of singularities by discrete multiscale directional representations. *The Journal of Geometric Analysis*, 28(3):2102–2128, 2018.
13. Robert Houska and Demetrio Labate. Detection of boundary curves on the piecewise smooth boundary surface of three dimensional solids. *Applied and Computational Harmonic Analysis*, 40(1):137–171, 2016.
14. Gitta Kutyniok, Jakob Lemvig, and Wang-Q Lim. Optimally sparse approximations of 3d functions by compactly supported shearlet frames. *SIAM Journal on Mathematical Analysis*, 44(4):2962–3017, 2012.
15. Gitta Kutyniok and Wang-Q Lim. Compactly supported shearlets are optimally sparse. *Journal of Approximation Theory*, 163(11):1564–1589, 2011.
16. Gitta Kutyniok and Philipp Petersen. Classification of edges using compactly supported shearlets. *Applied and Computational Harmonic Analysis*, 42(2):245–293, 2017.
17. Demetrio Labate, Wang-Q Lim, Gitta Kutyniok, and Guido Weiss. Sparse multidimensional representation using shearlets. In *Wavelets XI*, volume 5914, page 59140U. International Society for Optics and Photonics, 2005.
18. Stephane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, Inc., USA, 3rd edition, 2008.
19. Dimitris G Manolakis, Ronald B Lockwood, and Thomas W Cooley. *Hyperspectral imaging remote sensing: physics, sensors, and algorithms*. Cambridge University Press, 2016.
20. Pooran Singh Negi and Demetrio Labate. 3-d discrete shearlet transform and video processing. *IEEE transactions on Image Processing*, 21(6):2944–2954, 2012.
21. Saurabh Prasad, Demetrio Labate, Minshan Cui, and Yuhang Zhang. Morphologically decoupled structured sparsity for rotation-invariant hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 55(8):4355–4366, 2017.
22. Daniel Vera. Democracy of shearlet frames with applications. *Journal of Approximation Theory*, 213:23–49, 2017.