Acta Appl Math
https://doi.org/10.1007/s10440-018-0159-0

Inpainting via High-dimensional Universal Shearlet Systems

Z. Amiri[1,2] · R.A. Kamyabi-Gol[1,2]

Received: 18 October 2016 / Accepted: 15 January 2018

Abstract Thresholding and compressed sensing in combination with both wavelet and shearlet transforms have been very successful in inpainting tasks. Recent results have demonstrated that shearlets outperform wavelets in the problem of image inpainting. In this paper, we provide a general framework for universal shearlet systems in high dimensions. This theoretical framework is used to analyze the recovery of missing data via ℓ^1 minimization in an abstract model situation. In addition, we set up a particular model inspired by seismic data and a box mask to model missing data. Finally, the results of numerical experiments comparing various inpainting methods are presented.

Keywords Inpainting · ℓ^1 Minimization · Compressed sensing · Cluster coherence · Shearlets

Mathematics Subject Classification (2000) 42C40 · 42C15 · 65J22 · 65T60

1 Introduction

Reconstructing missing data is a popular challenge in both analog and digital fields. Also known as inpainting, this activity is the process of filling in a missing region or for making undetectable modifications to images, modifying the corrupted regions which are not consistent with the original images. Applications of inpainting range from restoring of missing blocks in video data to removal of occlusions such as text from images and repairing of scratched photos [1, 2, 13, 14].

✉ R.A. Kamyabi-Gol
kamyabi@um.ac.ir

Z. Amiri
za_am10@um.ac.ir

1 Department of Mathematics, Ferdowsi University of Mashhad, P.O. Box 1159-91775, Mashhad, Iran

2 Center of Excellence in Analysis on Algebraic Structures (CEAAS), Ferdowsi University of Mashhad, P.O. Box 1159-91775, Mashhad, Iran

Published online: 25 January 2018

Due to the vast interest on this topic, there exist several excellent works on inpainting via compressed sensing which is a fundamental method to recover sparsified data by ℓ^1 minimization [7, 16]. Previous works have focused on the concept of clustered sparsity which have led to theoretical bounds and results. In this setting, directional representation systems such as shearlets have been shown to outperform not only wavelets, but also other directional systems [5, 6, 18, 23]. In addition, the superiority of shearlets over wavelets for a basic thresholding algorithm and geometric separation was shown in [12, 16].

In [9], Genzel and Kutyniok introduced the more flexible universal shearlet systems, which are associated with an arbitrary scaling sequence. The performance for inpainting of this novel construction shearlet system in two dimensions was also analyzed. In this paper, we extend the framework of asymptotic analysis of inpainting by Genzel and Kutyniok to the higher-dimensional setting by generalizing the concept of universal shearlet systems. Using a method based on the original construction of Guo and Labate in [11], it is possible to provide a general framework for a signal in high dimensions, covering various challenges from the fields of image and video inpainting. A practical implementation of such a generalization in the three dimensions is applied to video inpainting in [18] which benefits from an optimally sparse approximation of three-dimensional data with C^2 surface singularities.

We continue this line of research by assuming that the missing area is known and has a different structure. Although the general strategy of our work is the same as in [9], the technical details are dependent on the chosen model for this area. The recovery errors of inpainting in Theorems 3 and 4 are the main results of this paper and formally justify the success of inpainting via universal shearlet frames. In this regard, we suggest conditions to relate the degree of anisotropic scaling to the admissible gap size. By considering the gap size as asymptotically smaller than the length of the corresponding shearlet elements, perfect inpainting results are achieved. These achievements are based on a recent work by Genzel and Kutyniok [9].

Finally, we consider inpainting as a task to assess the performance of different transformations in different types of images in the discrete domain. Our experiments show that shearlet transforms in general achieve a better trade off between computational efforts and reconstruction quality.

This review is organized as follows. Section 2 provides background of inpainting via a combination of applied harmonic analysis and compressed sensing. Section 3 contains the results of D-dimensional universal shearlet systems as sparsifying systems and the analysis of the performance for inpainting of this class of systems. In Sect. 4, we present a new model and investigate a recovery problem when using ℓ^1-analysis minimization algorithm for reconstruction. Section 5 is devoted to the results of numerical experiments, comparing various approaches.

2 Abstract Model and Inpainting via ℓ^1 Minimization

We start by analyzing the abstract Hilbert space, which is considered later on. Let x^0 be a signal in separable Hilbert Space $\mathcal{H}$. We assume that $\mathcal{H}$ can be decomposed into a direct sum of two closed subspaces, namely, a subspace $\mathcal{H}_M$ which is associated with the missing part of x^0 and a subspace $\mathcal{H}_K$ which is related to the known part of the signal. Hence,

$$\mathcal{H} = \mathcal{H}_K \oplus \mathcal{H}_M = P_K\mathcal{H} \oplus P_M\mathcal{H},$$

where P_M and P_K denote the orthogonal projections onto those subspaces, respectively. Note that, we will try to find the missing part $P_M x^0$, so the problem of data recovery can

Algorithm 1: Inpainting via ℓ^1 minimization

Input:
– Incomplete signal $P_K x_0 \in \mathcal{H}_K$.
– Parseval frame $\Phi = (\phi_i)_{i\in I}$.
Compute:
(ℓ^1-INP) $\quad x^\star = \text{argmin}_{x\in\mathcal{H}} \|T_\Phi x\|_{\ell^1(I)} \quad$ subject to $P_K x^0 = P_K x$
where T_Φ is analysis operator respect Φ ($T_\Phi : \mathcal{H} \to \ell^2(I)$, $x \to (\langle x, \phi_i\rangle)_{i\in I}$).
Output:
recovered signal $x^\star \in \mathcal{H}$.

be formulated as follows: Given a corrupt signal $P_K x^0$, recover the missing part $P_M x^0$. Depending on the dimension of the given model, we consider $\mathcal{H} = L^2(\mathbb{R}^D)$, $D \in \mathbb{N}$. If the measurable subset $\mathfrak{M} \subseteq \mathbb{R}^D$ is the missing area of the image, we may set $\mathcal{H}_M = L^2(\mathfrak{M})$. We recall that a sequence $\Phi = (\phi_i)_{i\in I}$ in a separable Hilbert space $\mathcal{H}$ is a frame if $A\|x\|^2 \le \|\langle x, \phi_i\rangle\|_{\ell^2(I)} \le B\|x\|^2$ for all $x \in \mathcal{H}$. If $A = B = 1$, it is called a Parseval frame. We assume that x^0 can be represented by a certain Parseval frame $\Phi = (\phi_i)_{i\in I}$ for $\mathcal{H}$, which can be selected non-adaptively or adaptively.

Now, we present the methods for recovering a signal which will be useful in the sequel. In fact, one of the fundamental methodologies for sparse recovery is ℓ^1 minimization [3, 8], which recovers the original signal by the recovery Algorithm 1 [15].

Since Parseval frames are not bases in general, there are many solutions such as c where $x = T_\Phi^\star c$, only the specific solution $T_\Phi x$ produces the desired numerical stabilities. The assumption of the sparsity signal x^0 by Φ provides a good recovery which is expected to occur.

Another reconstruction method from compressed sensing to achieve recovery is one-step thresholding which was introduced in [17] and is adapted from [16]. Note that, this minimization problem can also be regarded as a relaxation of the co-sparsity problem

$$x^\star = \text{argmin}_{x\in\mathcal{H}} \|T_\Phi x\|_{\ell^0(I)}, \quad \text{subject to } P_K x^0 = P_K x,$$

where $\|T_\Phi x\|_{\ell^0(I)} = \sharp\{\langle x, \phi_i\rangle \mid \langle x, \phi_i\rangle \neq 0\}$. Theoretical results associated to co-sparsity may be found in [21, 22].

In order to analyze the optimization problem using the inpainting algorithm, we need to introduce two important notions, δ-clustered sparsity and cluster coherence. These notions were applied to study the geometric separation problem and sparsity [4].

Definition 1 ([16]) Fix $\delta > 0$. A signal $x \in \mathcal{H}$ is called δ-clustered sparse in a Parseval frame Φ (with respect to $\Lambda \subseteq I$) if

$$\|1_{\Lambda^c} T_\Phi x\|_{\ell^1} \le \delta. \tag{1}$$

In this case, Λ is said to be δ-cluster for x in Φ.

The δ-clustered sparsity elucidates that coefficients outside of Λ are small. In fact, the cluster sparsity depends on the chosen set of indices Λ, enlarging Λ leads to smaller δ in (1).

Cluster coherence was introduced in [4] to investigate the missing part of signal x^0 on $\mathcal{H}_M$ and is defined as follows:

Definition 2 ([16]) Let $\Lambda \subseteq I$. The cluster coherence $\mu_c(\Lambda, P_M\Phi)$ of Parseval frame Φ with respect to $\mathcal{H}_M$ and Λ is defined by

$$\mu_c(\Lambda, P_M\Phi) = \max_{j\in J} \sum_{i\in\Lambda} \big|\langle P_M\phi_i, P_M\phi_j\rangle\big|,$$

where $P_M\Phi = (P_M\phi_i)_{i\in I}$.

By presenting these analysis tools (cluster sparsity and cluster coherence) we are able to state a main theorem that estimates the recovery error of the algorithm on the Hilbert space

$$\mathcal{H}_{1,\Phi} = \big\{x \in \mathcal{H} \,\big|\, \|x\|_{1,\Phi} = \|T_\Phi x\|_{\ell^1}\big\},$$

where Φ is a Parseval frame for Hilbert space $\mathcal{H}$. Details of the following theorem can be found in [9, 16].

Theorem 1 *Fix $\delta > 0$. Let $\Lambda \subseteq I$ be δ-cluster for x^0 in Parseval frame Φ and $\mu_c(\Lambda, P_M\Phi) < \frac{1}{2}$. If $x^0 \in \mathcal{H}_{1,\Phi}$, then*

$$\big\|x^\star - x^0\big\|_{1,\Phi} \le \frac{2\delta}{1-\mu_c(\Lambda, P_M\Phi)}.$$

A heuristic explanation of the theorem is as follows: If Φ sparsifies x^0 then there is a small set of analysis coefficients which contain most of the information of x^0 and those elements of Φ which capture that information do not fall too much into the hole of missing data. Indeed, we would like to select a cluster Λ such that x^0 becomes δ-cluster for small δ and cluster coherence μ_c shall not exceed the bound of $\frac{1}{2}$. While, both clustered sparsity and cluster coherence depend on the chosen set of indices Λ. Therefore, Theorem 1 provides a suitable cluster.

3 Inpainting via Universal Shearlet Systems

3.1 The Construction of D-dimensional Universal Shearlet Systems

In order to clarify the significance of universal shearlet systems, let us recall the main idea of classical shearlet systems in dimension $D = 2$ [10, 19]. For generator $\psi \in L^2(\mathbb{R}^2)$, a system of shearlets is defined by:

$$\big\{\psi_{j,l,k} = 2^{\frac{3j}{2}}\psi\big(S^l A^j[.] - k\big): \quad j \in \mathbb{Z}, l \in \mathbb{Z}, k \in \mathbb{Z}^2\big\},$$

where

$$A = \begin{pmatrix} 2^2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

denote the parabolic scaling matrix and shearing matrix, respectively. Using an appropriate cone-adapted construction and an appropriate generator, one can prove the optimally sparse approximation of cartoon-like functions, as shown in [10]. The tiling of the frequency domain of such shearlet systems is illustrated in Fig. 1. Finally, universal shearlet systems were introduced with associated scaling matrix

$$A^j_{\alpha_j} = \begin{pmatrix} 2^2 & 0 \\ 0 & 2^{\alpha_j} \end{pmatrix},$$

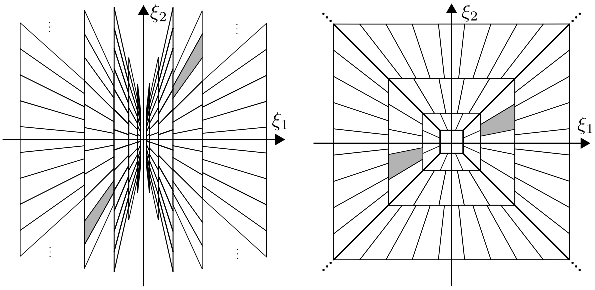


Fig. 1 (**A**) Frequency tiling of a classical shearlet system. (**B**) Frequency tiling of a cone-adapted shearlet system

where $(\alpha_j)_j \subseteq (-\infty, 2)$ to produce more flexibility in each scale [9]. The construction of such systems might be interesting for sparse-frequency localization and smoothness. The construction of universal shearlets can be generalized to higher dimensions. The following structures are based on the construction of smooth Parseval shearlet frames in [11] and 2-D universal shearlets in [9].

To start the construction of a D-dimensional universal shearlet, let us use the compact notation $\langle |x| \rangle = (1 + |x|^2)^{\frac{1}{2}}$ and recall that the Schwartz functions (the rapidly decreasing functions),

$$\mathcal{S}(\mathbb{R}^D) = \left\{ f \in C^\infty(\mathbb{R}^D) \,\middle|\, \forall K, N \in \mathbb{N}_0 : \sup_{x \in \mathbb{R}^D} \langle |x| \rangle^{-N} \sum_{|\alpha| \le K} |D^\alpha f(x)| < \infty \right\}.$$

As in the 2D case, let $\phi \in \mathcal{S}(\mathbb{R})$ such that $0 \le \hat{\phi} \le 1$, $\hat{\phi}(\xi) = 1$ on $\xi \in [-\frac{1}{16}, \frac{1}{16}]$ and $\operatorname{supp} \hat{\phi} \subset [-\frac{1}{8}, \frac{1}{8}]$.

For $\xi = (\xi_1, \dots, \xi_D) \in \mathbb{R}^D$, $j \in \mathbb{N}_0$, a smooth low pass function $\Phi(\xi)$ and corona scaling functions are defined by

$$\begin{aligned} \widehat{\Phi}(\xi) &= \hat{\phi}(\xi_1)\hat{\phi}(\xi_2)\dots\hat{\phi}(\xi_n), \\ W(\xi) &= \sqrt{\hat{\Phi}^2(2^{-2}\xi) - \hat{\Phi}^2(\xi)}, \\ W_j(\xi) &= W(2^{-2j}\xi). \end{aligned} \tag{2}$$

The functions satisfy this,

$$\widehat{\Phi}^2(\xi) + \sum_{j \ge 0} W_j^2(\xi) = 1, \quad \xi \in \mathbb{R}^D. \tag{3}$$

Note that, by the definition of ϕ, the sequence of functions W_j are compactly supported in

$$\mathscr{K} j = [-2^{2j-1}, 2^{2j-1}]^D \setminus (-2^{2j-4} 2^{2j-4})^D, \quad j \in \mathbb{N}_0. \tag{4}$$

Next, we consider function $\upsilon \in C^\infty(\mathbb{R})$ which satisfies $\operatorname{supp} \upsilon \subset [-1, 1]$ and

$$|\upsilon(u-1)|^2 + |\upsilon(u)|^2 + |\upsilon(u+1)|^2 = 1, \quad u \in [-1, 1], \tag{5}$$

moreover, we will presume that

$$\upsilon(0) = 1 \quad \text{and} \quad \upsilon^n(0) = 0, \quad n \ge 1. \tag{6}$$

A function with these properties was shown in [10]. In order to define the D-dimension of universal shearlet systems, we need to introduce the scaling matrices,

$$A_{\alpha,(1)} = \begin{pmatrix} 2^2 & 0 & \dots & 0 \\ 0 & 2^\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 2^\alpha \end{pmatrix}, \qquad A_{\alpha,(2)} = \begin{pmatrix} 2^\alpha & 0 & 0 & \dots & 0 \\ 0 & 2^2 & 0 & \dots & 0 \\ 0 & 0 & 2^\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2^\alpha \end{pmatrix}, \dots,$$

$$A_{\alpha,(D)} = \begin{pmatrix} 2^\alpha & 0 & \dots & 0 \\ 0 & 2^\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 2^2 \end{pmatrix},$$

where $\alpha \in (-\infty, 2)$ is the scaling parameter and for $l = (l_1, \dots, l_{D-1})$, the shear matrices

$$S^l_{(1)} = \begin{pmatrix} 1 & l_1 & \dots & l_{D-1} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \qquad S^l_{(2)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ l_1 & 1 & \dots & l_{D-1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \dots,$$

$$S^l_{(D)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ l_1 & l_2 & \dots & 1 \end{pmatrix}.$$

In order to define the universal shearlet systems in D-dimension, we need to combine a set of coarse scaling functions, a set of interior shearlets and boundary shearlets. In particular, the coarse scaling functions are defined

$$\psi_{-1,k}(x) = \Phi(x-k), \quad x \in \mathbb{R}^D, k \in \mathbb{Z}^D.$$

The interior shearlets are the elements

$$\big\{\psi^{\alpha,(d)}_{j,l,k} : \quad j \in \mathbb{N}_0, l \in \mathbb{Z}, |l_i| < 2^{(2-\alpha)j}, \\ k \in \mathbb{Z}^D, d = 1, \dots, D, i = 1, \dots, D-1\big\},$$

where

$$\widehat{\psi}^{\alpha,(d)}_{j,l,k}(\xi) = |\det A_{\alpha,(d)}|^{-\frac{j}{2}} W\big(2^{-2j}\xi\big) V_{(d)}\big(\xi A^{-j}_{\alpha,(d)} S^{-l}_{(d)}\big) e^{-2\pi i \xi A^{-j}_{\alpha,(d)} S^{-l}_{(d)} k},$$

and

$$V_{(d)}(\xi_1, \dots, \xi_D) = \prod_{\substack{m=1,\dots,D \\ m \neq d}} \upsilon\left(\frac{\xi_m}{\xi_d}\right).$$

Note that, the index d is associated with the D-dimensional pyramid

$$P_d = \left\{\xi \in \mathbb{R}^D : \left|\frac{\xi_1}{\xi_d}\right| \le 1, \dots, \left|\frac{\xi_D}{\xi_d}\right| < 1\right\},$$

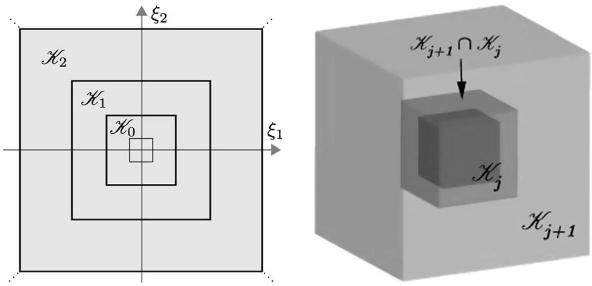


Fig. 2 Decomposition of the frequency plane by corona functions W_j and Φ in $D = 2, 3$. Note that the corona shapes $\mathcal{K}_j$ could slightly overlap

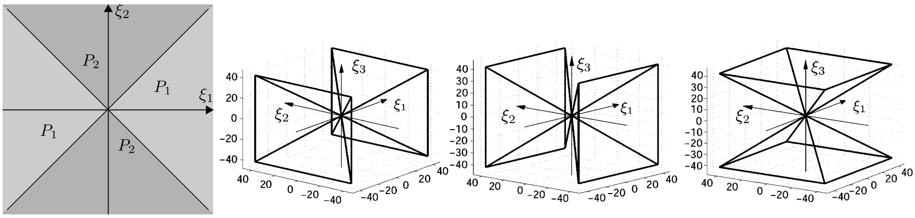


Fig. 3 Symmetric frequency decomposition by cones in $D = 2, 3$

and $\psi_{j,l,k}^{\alpha,(d)}$ has compact support in region

$$\Big\{\xi \in \mathbb{R}^D : \xi_d \in \left[-2^{2j-1}, 2^{2j-1}\right] \setminus \left(-2^{2j-4}, 2^{2j-4}\right),$$

$$\left|\frac{\xi_i}{\xi_d} - l_i 2^{-(2-\alpha)j}\right| \leq 2^{-(2-\alpha)j} \quad \text{for } i = 1, \ldots, d-1, \tag{7}$$

$$\left|\frac{\xi_i}{\xi_d} - l_{i-1} 2^{-(2-\alpha)j}\right| \leq 2^{-(2-\alpha)j} \quad \text{for } i = d+1, \ldots, D\Big\}.$$

Inspired by the cases $D = 2$ and $D = 3$ in [9, 11], we can define the boundary universal shearlets in D-dimension. For $D = 2$, there are 2 pyramidal regions and one set of boundary shearlets. In the case $D = 3$, there are 3 pyramidal regions and 2 sets of boundary universal shearlets. One set corresponding to 2 different pyramidal regions intersect and other set related to 3 different pyramidal regions intersect. In fact, for D-dimensional there are D pyramidal regions and $D - 1$ sets of boundary shearlets. Therefore, we need to introduce several set of boundary universal shearlets. The boundary shearlets associate with 2 different pyramidal regions intersect, are defined for $l_1 = \mp 2^{(2-\alpha)j}$, $|l_2|, \ldots, |l_{D-1}| < 2^{(2-\alpha)j}$ and have the form

$$\widehat{\psi}_{j,l,k,b}^{\alpha}(\xi) = 2^{-(D-1)\frac{D}{2}} \times 2^{-((D-1)\alpha+2)\frac{j}{2}}$$

$$\begin{cases} W(2^{-2j}\xi) V_{(q)}(\xi A_{\alpha,(q)}^{-j} S_{(q)}^{-l}) e^{2\pi i \xi 2^{-(D-1)}} A_{\alpha,(q)}^{-j} S_{(q)}^{-l} k & \text{if } \xi \in P_q \\ W(2^{-2j}\xi) V_{(q')}(\xi A_{\alpha,(q')}^{-j} S_{(q')}^{-l}) e^{2\pi i \xi 2^{-(D-1)}} A_{\alpha,(q)}^{-j} S_{(q)}^{-l} k & \text{if } \xi \in P_{q'}, \end{cases} \tag{8}$$

for all combinations of pyramidal regions P_q and $P_{q'}$. Note that, there are $2C(D, 2) = 2\frac{D!}{(D-2)!2!}$ hyperplanes. The boundary shearlets corresponding to 3 different pyramidal regions intersect, are defined for $l_1 = l_2 = \mp 2^{(2-\alpha)j}$, $|l_3| \leq \ldots |l_{D-1}| < 2^{(2-\alpha)j}$, and have the form

$$\widehat{\psi}^{\alpha}_{j,l,k,b}(\xi) = 2^{-(D-1)\frac{D}{2}} \times 2^{-((D-1)\alpha_j+2)\frac{j}{2}}$$
$$\begin{cases} W(2^{-2j}\xi)V_{(q)}(\xi A^{-j}_{\alpha,(q)} S^{-l}_{(q)}) e^{2\pi i \xi 2^{-(D-1)}} A^{-j}_{\alpha,(q)} S^{-l}_{(q)} k & \text{if } \xi \in P_q \\ W(2^{-2j}\xi)V_{(q')}(\xi A^{-j}_{\alpha,(q')} S^{-l}_{(q')}) e^{2\pi i \xi 2^{-(D-1)}} A^{-j}_{\alpha,(q)} S^{-l}_{(q)} k & \text{if } \xi \in P_{q'} \\ W(2^{-2j}\xi)V_{(q'')}(\xi A^{-j}_{\alpha,(q'')} S^{-l}_{(q'')}) e^{2\pi i \xi 2^{-(D-1)}} A^{-j}_{\alpha,(q)} S^{-l}_{(q)} k & \text{if } \xi \in P_{q''}, \end{cases} \tag{9}$$

for all combinations of pyramidal regions P_q, $P_{q'}$ and $P_{q''}$. The number of them is $2^2 C(D,3)$. We can say that for the boundary universal shearlets associate with L different pyramidal regions intersect, there are $2^{(L-1)}C(D,L)$ hyperplanes, for $L = 2, \dots, D$. Similarly, we proceed for the boundary universal shearlets corresponding to L different pyramidal regions intersect, where $L = 4, \dots, D$.

Before introducing the D-dimensional universal shearlet systems, it is useful to remark that due to the gluing of the boundary shearlets, scaling sequence was defined in [9] as follows:

A sequence $(\alpha_j)_{j\in\mathbb{N}_0} \subseteq \mathbb{R}$ is called a scaling sequence if

$$\alpha_j \in A_j = \left\{ \frac{m}{j} \,\middle|\, m \in \mathbb{Z}, m \le 2j-1 \right\} = \left\{ \dots, \frac{-2}{j}, \frac{-1}{j}, 0, \frac{1}{j}, \frac{2}{j}, \dots, 2 - \frac{1}{j} \right\},$$

for $j \ge 1$ and $\alpha_0 = 0$.

Definition 3 Let $(\alpha_j)_{j\in\mathbb{N}_0}$ be a scaling sequence. Then universal-scaling shearlet system or universal shearlet system is defined by

$$\mathrm{SH}\big(\phi, v, (\alpha_j)_j\big) = \mathrm{SH}_{\mathrm{Low}}(\phi) \cup \mathrm{SH}_{\mathrm{Int}}\big(\phi, v, (\alpha_j)_j\big) \cup \mathrm{SH}_{\mathrm{Bound}}\big(\phi, v, (\alpha_j)_j\big),$$

where

$$\begin{aligned} &\mathrm{SH}_{\mathrm{Low}}(\phi) = \{\psi_{-1,k} | k \in \mathbb{Z}^D\}, \\ &\mathrm{SH}_{\mathrm{Int}}\big(\phi, v, (\alpha_j)_j\big) \\ &\quad = \{\psi^{\alpha_j,(d)}_{j,l,k} | j \ge 0, |l_1|, \dots |l_{D-1}| < 2^{(2-\alpha_j)j}, k \in \mathbb{Z}^D, d = 1, \dots, D\}, \\ &\mathrm{SH}_{\mathrm{Bound}}\big(\phi, v, (\alpha_j)_j\big) \\ &\quad = \{\psi^{\alpha_j}_{j,l,k,b} | j \ge 0, |l_1| = 2^{(2-\alpha_j)j}, |l_2| \dots |l_{D-1}| < 2^{(2-\alpha_j)j}, k \in \mathbb{Z}^d\}, \\ &\qquad \cup \{\psi^{\alpha_j}_{j,l,k,b} | j \ge 0, |l_1| = |l_2| = 2^{(2-\alpha_j)j}, |l_3| \dots |l_{D-1}| < 2^{(2-\alpha_j)j}, k \in \mathbb{Z}^D\}, \\ &\qquad \vdots \\ &\qquad \cup \{\psi^{\alpha_j}_{j,l,k,b} | j \ge 0, |l_1| = |l_2| = \cdots = |l_{D-1}| = 2^{(2-\alpha_j)j}, k \in \mathbb{Z}^D\}. \end{aligned}$$

The next Theorem shows that universal shearlet systems are a smooth Parseval frame for $L^2(\mathbb{R}^D)$. Notice that the following proof is an adaption of the one in [9, 11].

Theorem 2 *With notations as above, the universal shearlet system is a Parseval frame for $L^2(\mathbb{R}^D)$. Moreover, the elements of this system are Schwartz functions and compactly supported in the Fourier domain.*

Proof The smoothness of SH_{Low} and SH_{Int} are concluded by the their smooth defining functions ϕ and υ. It remains to discuss the smoothness of the boundary shearlet elements. We need to analyze the boundary line of the cones which are given by $|\xi_1| = |\xi_2| = \cdots = |\xi_D|$. Let us consider the function $\widehat{\psi}^{\alpha_j}_{j,l,k,b}$, given by (8). Similarly to the $2D$ argument, we observe that the two terms of definition function $\widehat{\psi}^{\alpha_j}_{j,l,k,b}$ are differed by

$$\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_2}{\xi_1}-1\right)\right)\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_3}{\xi_1}-l_2\right)\dots\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1}-l_{D-1}\right)$$

and

$$\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2}-1\right)\right)\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_3}{\xi_2}-l_2\right)\dots\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2}-l_{D-1}\right).$$

Since $\upsilon^{(n)}(0) = 0$ for all $n \geq 1$, it follows that all derivatives of these functions are equal when $\xi_1 = \xi_2 = \cdots = \xi_D$.

Now, it only remains to prove Parseval frame property. For this, let $f \in L^2(\mathbb{R}^D)$, we will first consider the boundary shearlets (8). In the following, we will just investigate the case of $q = 1$, $q' = 2$. The arguments for any q, q' are similar. By Plancherel's theorem, we obtain

$$\begin{aligned}
&\sum_{j\geq 0}\sum_{k\in\mathbb{Z}^D}\left|\langle f, \psi^{\alpha_j}_{j,l,k,1}\rangle\right|^2 \\
&\quad=\sum_{j\geq 0}\sum_{k\in\mathbb{Z}^D}\left|\langle \hat{f}, \widehat{\psi}^{\alpha_j}_{j,l,k,1}\rangle\right|^2 \\
&\quad=\sum_{j\geq 0}\sum_{k\in\mathbb{Z}^D}\bigg|\int_{\mathbb{R}^D} 2^{-(D-1)\frac{D}{2}} \times 2^{-((D-1)\alpha_j+2)\frac{j}{2}}\hat{f}(\xi)e^{2\pi i\xi 2^{-(D-1)}A^{-j}_{\alpha_j,(1)}S^{-l}_{(1)}k} \\
&\quad\times\bigg[\chi_{P_1}\bigg(W(2^{-2j}\xi)\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_2}{\xi_1}-1\right)\right)\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_3}{\xi_1}-l_2\right)\dots \\
&\quad\times\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1}-l_{D-1}\right)\bigg)+\chi_{P_2}\bigg(W(2^{-2j}\xi)\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2}-1\right)\right) \\
&\quad\times\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_3}{\xi_2}-l_2\right)\dots\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2}-l_{D-1}\right)\bigg)\bigg]\bigg|^2.
\end{aligned}$$

To apply Parseval's identity, we will use the change of variable $\eta = \xi 2^{-(D-1)}A^{-j}_{(1),\alpha_j}S^{-l}_{(1)}$ and hence

$$\xi = 2^{(D-1)}\left(2^{2j}\eta_1, 2^{2j}\eta_1 + 2^{\alpha_j j}\eta_2, 2^{\alpha_j j}l_2\eta_1 + 2^{\alpha_j j}\eta_3, \dots, 2^{\alpha_j j}l_{D-1}\eta_1 + 2^{\alpha_j j}\eta_D\right).$$

By using of this variable we have:

$$V_{(1)}\left(\xi A^{-j}_{(1),\alpha_j}S^{-l}_{(1)}\right) = \upsilon\left(\frac{\eta_2}{\eta_1}\right)\upsilon\left(\frac{\eta_3}{\eta_1}\right)\upsilon\left(\frac{\eta_4}{\eta_1}\right)\dots\upsilon\left(\frac{\eta_D}{\eta_1}\right),$$

$$V_{(2)}\left(\xi A^{-j}_{(2),\alpha_j}S^{-l}_{(1)}\right) = \upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2}-1\right)\right)\dots\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2}-l_{D-1}\right),$$

$$= \upsilon\left(\frac{-\eta_1}{\eta_1 + 2^{(\alpha_j-2)j}\eta_2}\right)\upsilon\left(\frac{2^{(2-\alpha_j)j}\eta_3 - l_2\eta_2}{2^{(2-\alpha_j)j}\eta_1 + \eta_2}\right)\dots\upsilon\left(\frac{2^{(2-\alpha_j)j}\eta_D - l_{D-1}\eta_2}{2^{(2-\alpha_j)j}\eta_1 + \eta_2}\right),$$

$$W\big(2^{-2j}\xi\big) = W(2^{(D-1)}\eta_1, \dots, 2^{(2-\alpha_j)j+(D-1)}(\eta_D + l_{D-1}\eta_1)).$$

By the condition on the support of υ and W, the mapping

$$(\eta_1, \eta_2, \dots, \eta_D) \to W\big(2^{(D-1)}\eta_1, \dots, 2^{(2-\alpha_j)j+(D-1)}(\eta_D + l_{D-1}\eta_1)\big),$$

is supported inside the region $|\eta_1| < \frac{1}{2^D}$. Since $\operatorname{supp}\upsilon \subset [-1, 1]$, we can conclude that the functions

$$\Gamma_{1,j}(\eta) = W\big(2^{(D-1)}\eta_1, \dots, 2^{(\alpha_j-2)j+(D-1)}(\eta_D + l_{D-1}\eta_1)\big)\upsilon\left(\frac{\eta_2}{\eta_1}\right)\upsilon\left(\frac{\eta_3}{\eta_1}\right)\dots\upsilon\left(\frac{\eta_D}{\rho_1}\right),$$

are supported inside $Q = [-\frac{1}{2}, \frac{1}{2}]^D$. Now consider the functions

$$\begin{aligned}\Gamma_{2,j}(\eta) = {} & W\big(2^{(D-1)}\eta_1, \dots, 2^{(\alpha_j-2)j+(D-1)}(\eta_D + l_{D-1}\eta_1)\big) \\ & \times \upsilon\left(\frac{-\eta_1}{\eta_1 + 2^{(\alpha_j-2)j}\eta_2}\right)\upsilon\left(\frac{2^{(2-\alpha_j)j}\eta_3 - l_2\eta_2}{2^{(2-\alpha_j)j}\eta_1 + \eta_2}\right)\dots\upsilon\left(\frac{2^{(2-\alpha_j)j}\eta_D - l_{D-1}\eta_2}{2^{(2-\alpha_j)j}\eta_1 + \eta_2}\right).\end{aligned}$$

We will show that the support of $\Gamma_{i,j}$ is contained inside Q. The support condition of υ implies that

$$\begin{aligned}\left|\frac{\eta_2}{\eta_1 + 2^{(\alpha_j-2)j}\eta_2}\right| \le 1 \quad & \Longrightarrow \quad \left|\frac{\eta_2}{\eta_1}\right| \le \left|1 + 2^{(\alpha_j-2)j}\frac{\eta_2}{\eta_1}\right| \le 1 + 2^{(\alpha_j-2)j}\left|\frac{\eta_2}{\eta_1}\right| \\ & \Longrightarrow \quad \left|\frac{\eta_2}{\eta_1}\right| \le \frac{1}{1 - 2^{(\alpha_j-2)j}} \le 2,\end{aligned}$$

where the last estimate is due to $\alpha_j \le 2 - \frac{1}{j}$. This show that, if $|\eta_1| \le \frac{1}{2^D}$, then $|\eta_2| \le 2|\eta_1| \le \frac{1}{2^{(D-1)}}$. Again, by the support condition on υ, we have

$$\left|\frac{2^{(\alpha_j-2)j}\eta_3 - l_2\eta_2}{2^{(\alpha_j-2)j}\eta_1 + \eta_2}\right| \le 1 \quad \Longrightarrow \quad \left|\frac{\eta_3}{\eta_1} - 2^{(\alpha_j-2)j}l_2\frac{\eta_2}{\eta_1}\right| \le \left|1 + 2^{(\alpha_j-2)j}\frac{\eta_2}{\eta_1}\right| \le 1 + 2^{(\alpha_j-2)j}\left|\frac{\eta_2}{\eta_1}\right|.$$

Thus

$$\left|\frac{\eta_3}{\eta_1}\right| \le 1 + 2^{(\alpha_j-2)j}\left|\frac{\eta_2}{\eta_1}\right| + 2^{(\alpha_j-2)j}|l_2|\left|\frac{\eta_2}{\eta_1}\right| \le 4,$$

since $j \ge 1$, $|\frac{\eta_2}{\eta_1}| \le 2$. If $|\eta_1| \le \frac{1}{2^D}$, then $|\eta_3| \le \frac{1}{2^{(D-2)}}$. By the same strategies, we can obtain

$$|\eta_4| \le \frac{1}{2^{(D-3)}}, \dots, |\eta_D| \le \frac{1}{2^1}.$$

Hence, also $\Gamma_{2,j}(\eta)$ is supported inside Q, for each $j \ge 1$. With this, we have that, for $j \ge 1$

$$
\begin{aligned}
&\sum_{k\in\mathbb{Z}^D}\left|\langle\hat f,\widehat{\psi}^{\alpha_j}_{j,l,k,1}\rangle\right|^2\\
&\quad=\sum_{k\in\mathbb{Z}^D}\left|\int_Q 2^{((D-1)\alpha_j+2)\frac{j}{2}}\times 2^{(D-1)\frac{D}{2}}\hat f\big(2^{(D-1)}\eta S^l_{(1)}A^j_{\alpha_j,(1)}\big)\right.\\
&\qquad\left.\times\big(\chi_{P_1}(\eta)\Gamma_{1,j}(\eta)+\chi_{P_2}(\eta)\Gamma_{2,j}(\eta)\big)e^{2\pi i\eta k}d\eta\right|^2\\
&\quad=\int_{P_1 2^{-(D-1)}A^{-j}_{(1),\alpha_j}S^{-l}_{(1)}}2^{((D-1)\alpha_j+2)j+D(D-1)}\left|\hat f\big(2^{(D-1)}\eta S^l_{(1)}A^j_{(1)}\big)\right|^2\\
&\qquad\times\left|W\big(2^{(D-1)}\eta_1,\ldots,2^{(D-1)+(\alpha_j-2)j}(\eta_D+l_{D-1}\eta_1)\big)\right|^2\\
&\qquad\times\left|\upsilon\left(\frac{\eta_2}{\eta_1}\right)\right|^2\left|\upsilon\left(\frac{\eta_3}{\eta_1}\right)\right|^2\ldots\left|\upsilon\left(\frac{\eta_D}{\eta_1}\right)\right|^2d\eta\\
&\qquad+\int_{P_2 2^{-(D-1)}A^{-j}_{(1),\alpha_j}S^{-l}_{(1)}}2^{((D-1)\alpha_j+2)j+D(D-1)}\left|\hat f\big(2^{(D-1)}\eta S^l_{(1)}A^j_{(1)}\big)\right|^2\\
&\qquad\times\left|W\big(2^{(D-1)}\eta_1,\ldots,2^{(D-1)+(\alpha_j-2)j}(\eta_D+l_{D-1}\eta_1)\big)\right|^2\\
&\qquad\times\left|\upsilon\left(\frac{-\eta_2}{\eta_1+2^{(\alpha_j-2)j}\eta_2}\right)\right|^2\ldots\left|\upsilon\left(\frac{2^{(2-\alpha_j)j}\eta_D-l_{D-1}\eta_2}{2^{(2-\alpha_j)j}\eta_1+\eta_2}\right)\right|^2\\
&\quad=\int_{P_1}\left|\hat f(\xi)\right|^2\left|W\big(2^{-2j}\xi\big)\right|^2\left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_2}{\xi_1}-1\right)\right)\right|^2\ldots\left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1}-l_{D-1}\right)\right|^2d\xi\\
&\qquad+\int_{P_2}\left|\hat f(\xi)\right|^2\left|W\big(2^{-2j}\xi\big)\right|^2\left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2}-1\right)\right)\right|^2\ldots\left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2}-l_{D-1}\right)\right|^2d\xi.
\end{aligned}
$$

For $j=0$, observing that $\operatorname{supp} W\subset Q$, hence

$$
\begin{aligned}
&\sum_{k\in\mathbb{Z}^D}\left|\langle f,\psi^{\alpha_j}_{0,k,l,1}\rangle\right|^2\\
&\quad=\int_{P_1}\left|\hat f(\xi)\right|^2\left|W(\xi)\right|^2\left|\upsilon\left(\frac{\xi_2}{\xi_1}\mp 1\right)\right|^2\left|\upsilon\left(\frac{\xi_3}{\xi_1}\right)\right|^2\ldots\left|\upsilon\left(\frac{\xi_D}{\xi_1}\right)\right|^2d\xi\\
&\qquad+\int_{P_2}\left|\hat f(\xi)\right|^2\left|W(\xi)\right|^2\left|\upsilon\left(\frac{\xi_1}{\xi_2}\mp 1\right)\right|^2\left|\upsilon\left(\frac{\xi_3}{\xi_2}\right)\right|^2\ldots\left|\upsilon\left(\frac{\xi_D}{\xi_2}\right)\right|^2d\xi.
\end{aligned}
$$

So for any $f\in L^2(\mathbb{R}^D)$, we conclude that for $l_1=2^{(2-\alpha_j)j}$, $|l_2|\ldots|l_{D-1}|<2^{(2-\alpha_j)j}$, we have

$$
\begin{aligned}
&\sum_{k\in\mathbb{Z}^D}\left|\langle f,\psi^{\alpha_j}_{j,l,k,1}\rangle\right|^2\\
&\quad=\int_{P_1}\left|\hat f(\xi)\right|^2\left|W\big(2^{-2j}\xi\big)\right|^2\left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_2}{\xi_1}-1\right)\right)\right|^2\ldots\left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1}-l_{D-1}\right)\right|^2d\xi\\
&\qquad+\int_{P_2}\left|\hat f(\xi)\right|^2\left|W\big(2^{-2j}\xi\big)\right|^2\left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2}-1\right)\right)\right|^2\ldots\left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2}-l_{D-1}\right)\right|^2d\xi.
\end{aligned}
$$

A similar computation, for $l_1 = l_2 = 2^{(2-\alpha_j)j}$, $|l_3| \dots |l_{D-1}| < 2^{(2-\alpha_j)j}$ yields

$$\sum_{k\in\mathbb{Z}^D} \left|\langle f, \psi^{\alpha_j}_{j,l,k,1}\rangle\right|^2$$

$$= \int_{P_1} \left|\hat{f}(\xi)\right|^2 \left|W\left(2^{-2j}\xi\right)\right|^2 \left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_2}{\xi_1} - 1\right)\right)\right|^2$$

$$\times \left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_3}{\xi_1} - 1\right)\right)\right|^2 \dots \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1} - l_{D-1}\right)\right|^2 d\xi$$

$$+ \int_{P_2} \left|\hat{f}(\xi)\right|^2 \left|W\left(2^{-2j}\xi\right)\right|^2 \left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_2} - 1\right)\right)\right|^2 \dots \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_2} - l_{D-1}\right)\right|^2 d\xi$$

$$+ \int_{P_3} \left|\hat{f}(\xi)\right|^2 \left|W\left(2^{-2j}\xi\right)\right|^2 \left|\upsilon\left(2^{(2-\alpha_j)j}\left(\frac{\xi_1}{\xi_3} - 1\right)\right)\right|^2 \dots \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_3} - l_{D-1}\right)\right|^2 d\xi.$$

By using the same argument we obtain other boundary regions. Using the change of variable $\eta = \xi A^{-j}_{(d)} S^{-l}_{(d)}$ where $d = 1, \dots D$, for $|l_1| \dots |l_{D-1}| < 2^{(2-\alpha_j)j}$ follows that

$$\sum_{d=\{1,\dots,D\}} \sum_{j\in\mathbb{N}_0} \sum_{|l|<2^{(2-\alpha_j)j}} \left|\langle f, \psi^{\alpha_j,(d)}_{j,l,k}\rangle\right|^2$$

$$= \int_{\mathbb{R}^2} \left|\hat{f}(\xi)\right|^2 \sum_{j\in\mathbb{N}_0} \left|W\left(2^{-2j}\xi\right)\right|^2$$

$$\times \Bigg(\sum_{|l_1|,\dots,|l_{D-1}|<2^{(2-\alpha_j)j}} \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_2}{\xi_1} - l_1\right)\right|^2 \dots \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_D}{\xi_1} - l_{D-1}\right)\right|^2 \chi_{P_1}(\xi)$$

$$+$$

$$\vdots$$

$$+ \sum_{|l_1|,\dots,|l_{D-1}|<2^{(2-\alpha_j)j}} \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_1}{\xi_D} - l_1\right)\right|^2 \dots \left|\upsilon\left(2^{(2-\alpha_j)j}\frac{\xi_{D-1}}{\xi_D} - l_{D-1}\right)\right|^2 \chi_{P_D}(\xi)\Bigg) d\xi.$$

Since $\operatorname{supp} \Phi \subset Q$, we have

$$\sum_{k\in\mathbb{Z}^D} \left|\langle f, \psi_{-1,k}\rangle\right|^2 = \sum_{k\in\mathbb{Z}^D} \left|\int_Q \hat{f}(\xi)\hat{\Phi}(\xi)e^{-2\pi i\xi k} d\xi\right|^2 = \int_{\mathbb{R}^D} \left|\hat{f}(\xi)\right|^2 \left|\hat{\Phi}(\xi)\right|^2 d\xi.$$

Summarizing, we conclude that

$$\sum_{\psi\in \mathrm{SH}(\phi,\alpha_j,d)} \left|\langle f, \psi\rangle\right|^2$$

$$= \sum_{\psi\in \mathrm{SH}_{\mathrm{Low}}(\phi,\alpha_j,d)} \left|\langle f, \psi\rangle\right|^2 + \sum_{\psi\in \mathrm{SH}_{\mathrm{Int}}(\phi,\alpha_j,d)} \left|\langle f, \psi\rangle\right|^2 + \sum_{\psi\in \mathrm{SH}_{\mathrm{Bound}}(\phi,\alpha_j,d)} \left|\langle f, \psi\rangle\right|^2$$

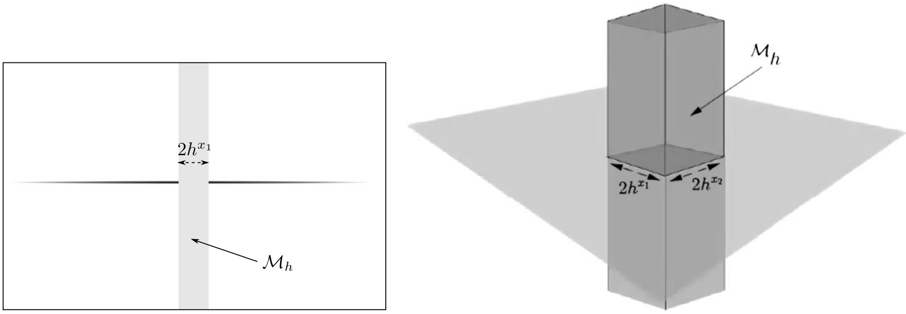


Fig. 4 Sketch of the corrupted modeling image in $D = 2, 3$

$$= \int_{\mathbb{R}^D} \left|\hat{f}(\xi)\right|^2 \sum_{j\geq 0} \left|W\left(2^{-2j}\xi\right)\right|^2 + \int_{\mathbb{R}^D} \left|\hat{f}(\xi)\right|^2 \left|\widehat{\Phi}(\xi)\right|^2 d\xi$$

$$= \int_{\mathbb{R}^D} \left|\hat{f}(\xi)\right|^2 \left(\sum_{j\geq 0} W\left(2^{-2j}\xi\right) + \left|\widehat{\Phi}(\xi)\right|^2\right) = \|f\|^2. \qquad \square$$

3.2 Shearlet Inpainting

In this section an image model which is compactly supported along the $(x_1, \ldots, x_{D-1})$-axes will be analyzed and the mask for missing part of the original image will be introduced.

Let $w : \mathbb{R}^{D-1} \to [0, 1]$ be a smooth function that is supported in $[-\rho_1, \rho_1] \times \ldots [-\rho_{D-1}, \rho_{D-1}]$, for some fixed $\rho_i \geq 0$, $i = 1, \ldots, D-1$. The distribution $w\mathcal{L}$ acting on Schwartz functions $\varphi \in \mathcal{S}(\mathbb{R}^D)$ is given by

$$\langle w\mathcal{L}, \varphi\rangle = \int_{-\rho_1}^{\rho_1} \cdots \int_{-\rho_{D-1}}^{\rho_{D-1}} w(x_1, \ldots, x_{D-1})\varphi(x_1, \ldots, x_{D-1}, 0)dx_{D-1} \ldots dx_1.$$

Note that this distribution is supported on $[-\rho_1, \rho_1] \times \ldots [-\rho_{D-1}, \rho_{D-1}] \times \{0\}$. The Fourier transform on $w\mathcal{L}$ can be computed as follows:

$$\begin{aligned}\langle \widehat{w\mathcal{L}}, \varphi\rangle &= \langle w\mathcal{L}, \hat{\varphi}\rangle \\ &= \int_{-\rho_1}^{\rho_1} \cdots \int_{-\rho_{D-1}}^{\rho_{D-1}} w(x_1, \ldots, x_{D-1}) \left(\int_{\mathbb{R}^D} \varphi(\xi) e^{-2\pi i (x_1\xi_1 + \cdots + x_{D-1}\xi_{D-1})} d\xi\right) dx \\ &= \int_{\mathbb{R}^D} \hat{w}(\xi_1, \ldots, \xi_{D-1})\varphi(\xi_1, \xi_2, \ldots, \xi_D) d\xi.\end{aligned}$$

Let now $F_j \in \mathcal{S}(\mathbb{R}^D)$ be a frequency filter, which is defined to be

$$\widehat{F}_j(\xi) = W_j(\xi) = W\left(2^{-2j}\xi\right), \quad \xi \in \mathbb{R}^D, j \geq 0,$$

where W_j is the corona function from the shearlet construction (2). So we obtain the filtered version of $w\mathcal{L}$ which we denote by $w\mathcal{L}_j$ i.e.,

$$w\mathcal{L}_j(x) = w\mathcal{L} * F_j(x) = \int_{\mathbb{R}^D} w\mathcal{L}(\cdot - t) F_j(t) dt.$$

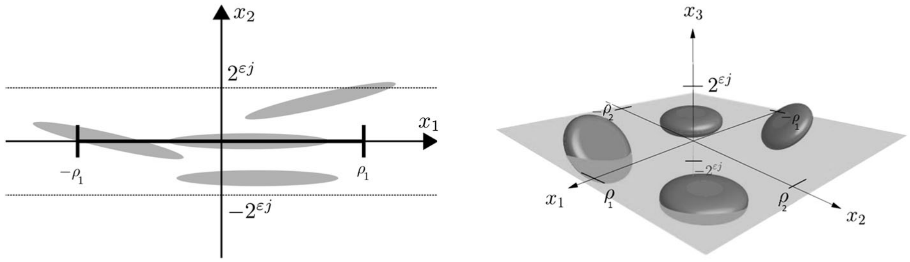


Fig. 5 Some shearlet elements associated with cluster Λ_j in $D = 2, 3$

By computing the Fourier transform $w\mathcal{L}_j$,

$$\widehat{w\mathcal{L}}_j(\xi) = \widehat{w\mathcal{L}}(\xi) \cdot \widehat{F}_j(\xi) = \widehat{w}(\xi_1, \dots, \xi_{D-1}) W_j(\xi_1, \dots, \xi_D), \quad \xi \in \mathbb{R}^D,$$

we can conclude that, the $w\mathcal{L}_j$ is band-limited Schwartz function. To distinguish the destroyed regions of image model, inspired by the missing sensor scenario in seismic data, we need to define the mask of the missing piece of the image

$$\mathcal{M}_h = \left\{(x_1, \dots, x_D) \in \mathbb{R}^D : |x_1| \le h^{x_1}, |x_2| \le h^{x_2}, \dots, |x_{D-1}| \le h^{x_{D-1}}\right\}.$$

The gap space and the orthogonal projection on this space, are now defined by $H_M = L^2(\mathcal{M}_h)$, and $P_M f = \chi_{\mathcal{M}_h} f,\ f \in L^2(\mathbb{R}^D)$. For investigating the Fourier support properties of the model function, we use the notation $\psi_\gamma = \psi_{j,l,k}^{\alpha_j,(d)}$ for $\gamma = (j, l, k, \alpha_j, d) \in \Gamma$ in which the index set Γ satisfies

$$\bigcup_{\gamma \in \Gamma} \psi_\gamma = \mathrm{SH}_{\mathrm{Int}}\big(\phi, \upsilon, (\alpha_j)_j\big) \cup \mathrm{SH}_{\mathrm{Bound}}\big(\phi, \upsilon, (\alpha_j)_j\big).$$

Set $\Gamma_j = \{(j', l, k, \alpha_j, d) \in \Gamma | j' = j\}$, $j \ge 0$, note that, the support of $\widehat{w\mathcal{L}_j}$ is contained in (4) and region $\mathscr{K}_j$ overlaps with $\mathscr{K}_{j-1}$ and $\mathscr{K} j + 1$, but is disjointed from all the other ones. Therefore, for $j \ge 1$, we have that

$$\langle w\mathcal{L}_j, \psi_{-1,k}\rangle = 0, \quad k \in \mathbb{Z}^D$$

$$\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_j,(d)}\big\rangle = 0, \quad \big(j', l, k, \alpha_j, d\big) \in \Gamma_{j'}, \quad |j' - j| > 1.$$

Hence, we may conclude that all non-zero coefficients are contained in sets $\Gamma_j^{\pm 1} = \Gamma_{j-1} \cup \Gamma_j \cup \Gamma_{j+1}$ for $j \ge 0$.

Now, we fix some $\varepsilon > 0$, we choose the set of significant shearlet coefficients to be

$$\Lambda_j = \left\{(j, l, k, \alpha_j, D) : |l_1| \dots |l_{D-1}| \le 1, |l_1 k_1 + \dots + l_{D-1} k_{D-1} - k_D| \le 2^{\varepsilon j}, k \in \mathbb{Z}^D, j \ge 0\right\} \subset \Gamma_j. \tag{10}$$

Since $\operatorname{supp} w\mathcal{L}$ is covered by every Λ_j, we conclude that (10) is an appropriate choice. More precisely, ε controls the behavior between clustered sparsity and cluster coherence in Theorem 1. We put

$$\delta_j^{j'-j} = \big\| 1_{\Gamma_{j'} \setminus \Lambda_{j'}} (T_\Psi w\mathcal{L}_j) \big\|_{\ell^1} = \sum_{\gamma \in \Gamma_{j'} \setminus \Lambda_{j'}} \big|\langle w\mathcal{L}_j, \psi_\gamma\rangle\big|, \quad j, j' \ge 1, |j' - j| \le 1, \tag{11}$$

and then sum up

$$\delta_j = \delta_j^{+1} + \delta_j^0 + \delta_j^{-1} = \left\| 1_{(\Lambda_j^{\mp})^c}(T_\Psi w\mathcal{L}_j) \right\|_{\ell^1} = \sum_{\gamma \in (\Lambda_j^{\pm 1})^c} \left\langle |w\mathcal{L}_j, \psi_\gamma\rangle \right|, \quad j \geq 1, \tag{12}$$

where $\Lambda_j^{\pm 1} = \Lambda_{j-1} \cup \Lambda_j \cup \Lambda_{j+1}$. By definition, $w\mathcal{L}_j$ is δ_j-clustered sparse in $\Psi = \mathrm{SH}(\phi, \upsilon, (\alpha_j)_j)$ with respect to the cluster $\Lambda_j^{\pm 1}$.

It might be helpful to mention that the goal is to show that δ_j is small. Then the abstract Theorem 1 implies a good inpainting result.

The following lemma is needed for estimating the decay coefficients of the shearlet with the singularities. We use of simple notation for the transformed translations,

$$t_d = (t_1, \ldots, t_D) := A_{\alpha_j,(d)}^{-j} S_{(d)}^{-l} k, \quad d = 1, \ldots, D.$$

Lemma 1 *Let* $(j, l, k, \alpha_j, d) \in \Gamma_j$ *with* $j \geq 1$. *If* $\alpha_j \geq 0$, *then the following estimates hold for arbitrary integer* $N_1, \ldots, N_D, N \geq 0$:

(1). *If* $d = D$, *and* $|l_i| > 1$ *for* $i = 1, \ldots, D-1$, *we have*

$$\left\langle \left| w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)} \right\rangle \right| \leq c_{N_1,\ldots,N_D,N} |t_1|^{N_1} \ldots |t_D|^{N_D} \\ \times 2^{-N\alpha_j j} \left(|l_i| - 1 \right)^{-N} \times 2^{(2-(D-1)\alpha_j)\frac{j}{2}} \times 2^{-N_D \alpha_j j}.$$

(2). *If* $d = 1, \ldots, D-1$, *we have*

$$\left| \left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(d)} \right\rangle \right| \leq c_{N_1,\ldots,N_D,N} |t_1|^{-N_1} \ldots |t_D|^{-N} \\ \times 2^{-2N_j} \left(|l_{D-1}| + 1 \right) \times 2^{(2-(D-1)\alpha_j)\frac{j}{2}} \times 2^{-N_D \alpha_j j}.$$

(3). *For boundary shearlets we have*

$$\left\langle \left| w\mathcal{L}_j, \psi_{j,l,k,b}^{\alpha_j} \right\rangle \right| \leq c_{N_1,\ldots,N_D,N} |t_1|^{-N_1} \ldots |t_D|^{-N_D} \\ \times 2^{-2N_j} \left(|l_{D-1}| + 1 \right) \times 2^{(2-(D-1)\alpha_j)\frac{j}{2}} \times 2^{-N_D \alpha_j j}.$$

(4). *If* $d = D$ *and* $|l_i| \leq 1$ *for* $i = 1, \ldots, D-1$ *we have*

$$\left| \left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)} \right\rangle \right| \leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \int_{\mathbb{R}^{D-1}} \left\langle |x_1| \right\rangle^{-N} \times \ldots \left\langle |x_{D-1}| \right\rangle^{-N} \\ \times \tilde{w}_{N,j} \left(2^{-\alpha_j j}(x_1 + k_1), \ldots, 2^{-\alpha_j j}(x_{D-1} + k_{D-1}) \right) \\ \times \left\langle |l_1 x_1 + l_1 k_1 + \cdots - k_D| \right\rangle^{-N} dx,$$

where $\tilde{w}_{N,j} = |w| * \langle |2^{2j}[.]| \rangle^{-N}$.

Proof (1). By the definition of universal shearlets and Plancherel's Theorem we have

$$\left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)} \right\rangle = \left\langle \widehat{w\mathcal{L}}_j, \widehat{\psi}_{j,l,k}^{\alpha_j,(D)} \right\rangle \\ = \int_{\mathbb{R}^D} \widehat{w}(\xi_1, \ldots, \xi_{D-1}) W\left(2^{-2j}\xi\right) \overline{\widehat{\psi}_{j,l,k}^{\alpha_j,(D)}} d\xi$$

$$= \int_{\mathbb{R}} e^{2\pi i t_D \xi_D} \left(\int_{\mathbb{R}^{D-1}} \widehat{w}(\xi_1, \ldots, \xi_{D-1}) W\left(2^{-2j}\xi\right) \widehat{\psi}^{\alpha_j,(D)}_{j,l,0}(\xi) \right.$$
$$\left. \times e^{2\pi i (t_1\xi_1 + \cdots + t_{D-1}\xi_{D-1})} d\xi_1 \ldots d\xi_{D-1} \right) d\xi_D.$$

We differentiate the function $\xi \to \hat{w}(\xi_1, \ldots, \xi_{D-1}) W(2^{-2j}\xi) \psi^{\alpha_j,(D)}_{j,l,0}(\xi)$, N_i-times for $i = 1, \ldots, D$. With partial integration we obtain

$$\begin{aligned} &\left|\left\langle w\mathcal{L}_j, \psi^{\alpha_j,(D)}_{j,l,k}\right\rangle\right| \\ &\quad \le c_{N_1,\ldots,N_D} |t_1|^{-N_1} \ldots |t_D|^{-N_D} \\ &\qquad \times \int_{\mathbb{R}} \underbrace{\int_{\mathbb{R}^{D-1}} |D^{(N_1,\ldots,N_D)}\left(\widehat{w}(\xi_1, \ldots, \xi_{D-1}) W\left(2^{-2j}\xi\right) \psi^{\alpha_j,(D)}_{j,l,0}\right)| d\xi_1 \ldots d\xi_{D-1}}_{h_{N_1 \ldots N_{D-1}}(\xi_D)} d\xi_D, \end{aligned}$$

where the boundary terms vanish because of the compact support of $\xi \to \hat{w}(\xi_1, \ldots, \xi_{D-1}) \times W(2^{-2j}\xi)\psi^{\alpha_j,(D)}_{j,l,0}(\xi)$. With (4) obtain $2^{2j-4} \le |\xi_D| \le 2^{2j-1}$ and

$$\begin{cases} \left(|l_1| - 1\right) 2^{(\alpha_j - 2)j} \le \left|\dfrac{\xi_1}{\xi_D}\right| \le \left(|l_1| + 1\right) 2^{(\alpha_j - 2)j} \\ \left(|l_2| - 1\right) 2^{(\alpha_j - 2)j} \le \left|\dfrac{\xi_2}{\xi_D}\right| \le \left(|l_2| + 1\right) 2^{(\alpha_j - 2)j} \\ \vdots \\ \left(|l_{D-1}| - 1\right) 2^{(\alpha_j - 2)j} \le \left|\dfrac{\xi_{D-1}}{\xi_D}\right| \le \left(|l_{D-1}| + 1\right) 2^{(\alpha_j - 2)j}. \end{cases}$$

We can conclude that

$$\xi_i \in I_{l_i} = \left[-2^{\alpha_j j - 1}\left(|l_i| + 1\right), -2^{\alpha_j j - 4}\left(|l_i| - 1\right)\right] \cup \left[2^{\alpha_j j - 4}\left(|l_i| - 1\right), 2^{\alpha_j j - 1}\left(|l_i| + 1\right)\right],$$

for $i = 1, \ldots, D-1$. The next step is to estimate the term $h_{N_1,\ldots,N_{D-1}}(\xi_D)$. Now, the Leibniz rule and Hölder's Inequality yield

$$\begin{aligned} &h_{N_1,\ldots,N_{D-1}}(\xi) \\ &\quad \le \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \cdots \sum_{n_{D-1}=0}^{N_{D-1}} \binom{N_1}{n_1} \binom{N_2}{n_2} \cdots \binom{N_{D-1}}{n_{D-1}} \\ &\qquad \times \int_{\mathbb{R}^{D-1}} |\widehat{w}^{(n_1,\ldots,n_{D-1})}(\xi_{D-1}, \ldots, \xi_1) D^{N_1 - n_1, \ldots, N_{D-1} - n_{D-1}}\left(W\left(2^{-2j}\xi\right) \psi^{\alpha_j,(D)}_{j,l,0}(\xi)\right) d\xi \\ &\quad \le \sum_{n_1=0}^{N_1} \cdots \sum_{n_{D-1}=0}^{N_{D-1}} \binom{N_1}{n_1} \cdots \binom{N_{D-1}}{n_{D-1}} \|\widehat{w}^{(n_1,\ldots,n_{D-1})}\|_{L^1(I_1 \times \cdots \times I_{D-1})} \\ &\qquad \times \left\| D^{N_1 - n_1, \ldots, N_{D-1}} W\left(2^{-2j}.\right) \psi^{\alpha_j,(D)}_{j,l,0}(.) \right\|_{L^\infty(\mathbb{R}^D)}. \end{aligned}$$

By the support and rapid decay of $\hat{w}$, it can be estimated as

$$\begin{aligned}
&\left\|\widehat{w}^{(n_1,\ldots,n_{D-1})}\right\|_{L^1(I_1\times\cdots\times I_{D-1})}\\
&=\int_{I_{D-1}}\cdots\int_{I_1}\widehat{w}^{(n_1,\ldots,n_{D-1})}(\xi_1,\ldots,\xi_{D-1})d\xi_1\ldots d\xi_{D-1}\\
&\leq c_N\int_{2^{\alpha_j j-4}(|l_{D-1}|-1)}^{2^{\alpha_j j-1}(|l_{D-1}|-1)}\cdots\int_{2^{\alpha_j j-4}(|l_1|-1)}^{2^{\alpha_j j-1}(|l_1|-1)}\langle|\xi_1|\rangle^{-\frac{N}{2}}\ldots\langle|\xi_{D-1}|\rangle^{-\frac{N}{2}}d\xi_1\ldots d\xi_{D-1}\\
&\overset{|l_i|>1}{\leq} c_{N_1,\ldots,N_{D-1}}2^{-N\alpha_j j}\big(|l_i|-1\big)^{-N}.
\end{aligned}$$

Last estimate concludes because

$$\int_{2^{\alpha_j j-4}(|l_k|-1)}^{2^{\alpha_j j-1}(|l_k|-1)}\langle|\xi_k|\rangle^{-\frac{N}{2}}\leq c_N,\quad k=1,\ldots,D-1,k\neq i.$$

So, we can obtain

$$\begin{aligned}
&\left\|D^{N_1-n_1,\ldots,N_D}\big(W\big(2^{-2j}.\big)\psi_{j,l,0}^{\alpha_j,(D)}(.)\big)\right\|_{L^\infty(\mathbb{R}^D)}\\
&\leq c_{N_1,\ldots,N_D}2^{-((D-1)\alpha_j+2)\frac{j}{2}}\times 2^{-(N_1-n_1)\alpha_j j}\ldots\times 2^{-(N_{D-1}-n_{D-1})\alpha_j j}\times 2^{-N_D\alpha_j j}.
\end{aligned}$$

This implies

$$\begin{aligned}
h_{N_1,\ldots,N_D}(\xi_D)&\leq\sum_{n_1=0}^{N_1}\cdots\sum_{n_{D-1}=0}^{N_{D-1}}c_{N_1,\ldots,N_D,N}2^{-N\alpha_j j}\big(|l_i|-1\big)^{-N}\\
&\quad\times 2^{-((D-1)\alpha_j+2)\frac{j}{2}}\times 2^{-(N_1-n_1)\alpha_j j}\ldots 2^{-N_D\alpha_j j}\\
&\leq c_{N_1,\ldots,N_D,N}\times 2^{-N\alpha_j j}\big(|l_i|-1\big)^{-N}\times 2^{-((D-1)\alpha_j+2)\frac{j}{2}}\times 2^{-N_D\alpha_j j}\\
&\quad\times\underbrace{\sum_{n_1=0}^{\infty}\binom{N_1}{n_1}\big(2^{-\alpha_j j}\big)^{(N_1-n_1)}}_{\underset{\alpha_j\geq 0}{=(1+2^{-\alpha_j j})^{N_1}\leq 2^{N_1}}}\cdots\underbrace{\sum_{n_1=0}^{\infty}\binom{N_{D-1}}{n_{D-1}}\big(2^{-\alpha_j j}\big)^{(N_{D-1}-n_{D-1})}}_{\underset{\alpha_j\geq 0}{=(1+2^{-\alpha_j j})^{N_{D-1}}\leq 2^{N_{D-1}}}}\\
&\leq c_{N_1,\ldots,N_D,N}2^{-N\alpha_j j}\big(|l_i|-1\big)^{-N}\times 2^{-((D-1)\alpha_j+2)\frac{j}{2}}\times 2^{-N_D\alpha_j j}.
\end{aligned}$$

Combining this estimate with $|\operatorname{supp} h_{N_1,\ldots,N_D}|\leq 2^{2j}c$ concludes

$$\begin{aligned}
\left|\left\langle w\mathcal{L}_j,\psi_{j,l,k}^{\alpha_j,(D)}\right\rangle\right|&\leq c_{N_1,\ldots,N_D,N}|t_1|^{-N_1}\ldots|t_D|^{-N_D}\\
&\quad\times 2^{-N\alpha_j j}\big(|l_i|-1\big)^{-N}\times 2^{(2-(D-1)\alpha_j)\frac{j}{2}}\times 2^{-N_D\alpha_j j}.
\end{aligned}$$

(2). Similarly the proof of part (1), for $d=1,\ldots,D-1$, we have

$$\begin{aligned}
\left|\left\langle w\mathcal{L}_j,\psi_{j,l,k}^{\alpha_j,(d)}\right\rangle\right|&\leq c_{N_1,\ldots,N_D,N}|t_1|^{-N_1}\ldots|t_D|^{-N_D}\\
&\quad\times 2^{-2N_j}(|l_{D-1}|+1)\times 2^{(2-(D-1)\alpha_j)\frac{j}{2}}\times 2^{-N_D\alpha_j j}.
\end{aligned}$$

(3). The boundary shearlets can be estimated by part (1) and (2).

(4). By the definition $w\mathcal{L}_j$, it follow that

$$\begin{aligned}
\big|w\mathcal{L}_j(x)\big| &= \big|(w\mathcal{L} * F_j)(x)\big| = \left|\int_{\mathbb{R}^{D-1}} w(y) F_j\big(x-(y,o)\big)dy\right| \\
&\le \int_{\mathbb{R}^{D-1}} \big|w(y)\big| 2^{2Dj} \big|\check{W}\big(2^{2j}\big(x-(y,0)\big)\big)\big| dy \\
&= c_N 2^{2Dj} \big\langle \big|2^{2j} x_D\big|\big\rangle^{-N} \underbrace{\big[|w| * \big\langle 2^{2j}[\cdot]\big\rangle^{-N}\big](x_1,\dots,x_{D-1})}_{\tilde{w}_{N,j}(x_1,\dots,x_{D-1})} \\
&= c_N 2^{2Dj} \big\langle \big|2^{2j} x_D\big|\big\rangle^{-N} \tilde{w}_{N,j}(x_1,\dots,x_{D-1}), \quad x=(x_1,\dots,x_{D-1},x_D)\in\mathbb{R}^D.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\big|\psi_{j,l,k}^{\alpha_j,(D)}(x)\big| &\le c_N 2^{((D-1)\alpha_j+2)\frac{j}{2}} \big\langle \big|S_D^l A_{\alpha_j,(D)}^j x - k\big|\big\rangle^{-N} \\
&\le c_N 2^{((D-1)\alpha_j+2)\frac{j}{2}} \big\langle \big|2^{\alpha_j j} x_1 - k_1\big|\big\rangle^{-N} \dots \big\langle \big|2^{\alpha_j j} x_{D-1} - k_{D-1}\big|\big\rangle^{-N} \\
&\times \big\langle \big|l_1 2^{\alpha_j j} x_1 + l_2 2^{\alpha_j j} x_2 + \dots + 2^{2j} x_D k_D\big|\big\rangle^{-N}.
\end{aligned}$$

Now, we can estimate the analysis coefficients.

$$\begin{aligned}
&\big|\big\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\big\rangle\big| \\
&\quad\le 2^{(6-\alpha_j)(D-1)\frac{j}{2}} \int_{\mathbb{R}^D} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1),\dots,2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big) \\
&\qquad\times \langle |x_D|\rangle^{-N} \dots \langle |x_1|\rangle^{-N} \langle |l_1 x_1 + l_1 k_1 + \dots + l_{D-1} x_{D-1} + x_D - k_D|\rangle^{-N} dx.
\end{aligned}$$

Let $\mathbf{k} = l_1 x_1 + l_1 k_1 + \dots + l_{D-1} x_{D-1} + l_{D-1} k_{D-1} - k_D$. It is clear that one of the two factor $\langle |x_D|\rangle^{-N} \langle |x_D + \mathbf{k}|\rangle^{-N}$ has to be smaller than $\langle |\frac{\mathbf{k}}{2}|\rangle^{-N}$. Hence

$$\begin{aligned}
&\int_{\mathbb{R}} \langle |x_D|\rangle^{-N} \langle |x_D + \mathbf{k}|\rangle^{-N} dx_D \\
&\quad= \int_{\mathbb{R}} \underbrace{\max\big\{\langle |x_D|\rangle^{-N}, \langle |x_D + \mathbf{k}|\rangle^{-N}\big\}}_{\le \langle |x_D|\rangle^{-N} + \langle |x_D + \mathbf{k}|\rangle^{-N}} \min\underbrace{\big\{\langle |x_D|\rangle^{-N}, \langle |x_D + \mathbf{k}|\rangle^{-N}\big\}}_{\le \langle |\frac{\mathbf{k}}{2}|\rangle^{-N}} dx_D \qquad (13) \\
&\quad\le c_N \langle |\mathbf{k}|\rangle^{-N} = c_N \langle |l_1 x_1 + l_1 k_1 + \dots - k_D|\rangle^{-N},
\end{aligned}$$

finally, we obtain

$$\begin{aligned}
&\big|\big\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\big\rangle\big| \\
&\quad\le 2^{(6-\alpha_j)(D-1)\frac{j}{2}} \int_{\mathbb{R}^{D-1}} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1),\dots,2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big) \\
&\qquad\times \langle |x_{D-1}|\rangle^{-N} \dots \langle |x_1|\rangle^{-N} \langle |l_1 x_1 + l_1 k_1 + \dots + l_{D-1} x_{D-1} - k_D|\rangle^{-N} dx. \qquad \square
\end{aligned}$$

Proposition 1 *With the notations as above if* $\liminf_{j\to\infty}\alpha_j > 0$, *then*

$$\delta_j \in o\big(2^{-Nj}\big), \quad j \to \infty,$$

for every $N \in \mathbb{N}$.

That is what we call rapid decay, adapted from the behavior of Schwartz functions.

Proof We need to estimate $\delta_j^{j'-j}$ which is defined in (11). If $\mathbf{k} = l_1k_1 + l_2k_2 + \cdots + l_{D-1}k_{D-1}$, then

$$\begin{aligned}\delta_j^{j'-j} = & \sum_{\substack{k\in\mathbb{Z}^D, |l_1|\ldots|l_{D-1}|\leq 1\\ |\mathbf{k}-k_D|>2^{\varepsilon j'}}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_{j'},(D)}\big\rangle\big| + \sum_{\substack{k\in\mathbb{Z}^D\\ |l_1|>1, |l_2|\ldots|l_{D-1}|\leq 1}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_{j'},(D)}\big\rangle\big| \\ & + \sum_{\substack{k\in\mathbb{Z}^D\\ |l_2|>1, |l_1|, |l_3|, \ldots |l_{D-1}|\leq 1}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_{j'},(D)}\big\rangle\big| + \cdots + \sum_{\substack{k\in\mathbb{Z}^D\\ |l_{D-1}|>1, |l_1|, \ldots |l_{D-2}|\leq 1}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_{j'},(D)}\big\rangle\big| \\ & + \sum_{d=1}^{D-1} \sum_{\substack{k\in\mathbb{Z}^D\\ |l_1|\ldots|l_{D-1}|<2^{(2-\alpha_{j'})j'}}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k}^{\alpha_{j'},(d)}\big\rangle\big| + \sum_{\substack{k\in\mathbb{Z}^D, |l_1|=2^{(2-\alpha_{j'})j'}\\ |l_2|\ldots|l_{D-1}|\leq 2^{(2-\alpha_j)j'}}} \big|\big\langle w\mathcal{L}_j, \psi_{j',l,k,b}^{\alpha_{j'}}\big\rangle\big|.\end{aligned}$$

We will compute for $j' = j$, the computation for $j' = j \mp 1$ are exactly the same. We apply Lemma 1 for $N \geq 2$, hence

$$\begin{aligned} & \sum_{\substack{k\in\mathbb{Z}^D, |l_1|\ldots|l_{D-1}|\leq 1\\ |\mathbf{k}-k_D|>2^{\varepsilon j}}} \big|\big\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\big\rangle\big| \\ & \leq c_N 2^{(6-\alpha_j)(D-1)\frac{j}{2}} \sum_{\substack{k\in\mathbb{Z}^D\\ |l_1|\ldots|l_{D-1}|<1}} \int_{\mathbb{R}^{D-1}} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1), \ldots, 2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big) \\ & \quad \times \langle |x_1|\rangle^{-N} \ldots \langle |x_{D-1}|\rangle^{-N} \langle |\mathbf{k}-k_D|\rangle^{-N} dx \\ & = c_N 2^{(6-\alpha_j)(D-1)\frac{j}{2}} \sum_{\substack{k\in\mathbb{Z}^D, |k_D|>2^{\varepsilon j}\\ |l_1|\ldots|l_{D-1}|<1}} \int_{\mathbb{R}^{D-1}} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1), \ldots, 2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big) \\ & \quad \times \langle |x_1|\rangle^{-N} \ldots \langle |x_{D-1}|\rangle^{-N} \langle |l_1x_1 + l_2x_2 + \cdots + l_{D-1}x_{D-1} - k_D|\rangle^{-N} dx \\ & = c_N 2^{(6-\alpha_j)(D-1)\frac{j}{2}} \\ & \quad \times \sum_{\substack{|l_1|\ldots|l_{D-1}|<1\\ k\in\mathbb{Z}^D, |k_D|>2^{\varepsilon j}}} \int_{\mathbb{R}^{D-1}} \Big(\sum_{k_1,\ldots,k_{D-1}\in\mathbb{Z}} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1), \ldots, 2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big) \\ & \quad \times \langle |x_1|\rangle^{-N} \ldots \langle |x_{D-1}|\rangle^{-N} \langle |l_1x_1 + l_2x_2 + \cdots + l_{D-1}x_{D-1} - k_D|\rangle^{-N} dx. \end{aligned}$$

We have the following argumentation:

$$\sum_{k_1,\ldots,k_{D-1}\in\mathbb{Z}} \tilde{w}_{N,j}\big(2^{-\alpha_j j}(x_1+k_1), \ldots, 2^{-\alpha_j j}(x_{D-1}+k_{D-1})\big)$$

$$\leq \sum_{k_1,\ldots,k_{D-1}\in\mathbb{Z}} \int_{\mathbb{R}^{D-1}} |w(y_1,\ldots,y_{D-1})| \langle |2^{(2-\alpha_j)j}(k_1+x_1-2^{\alpha_j j}y_1)| \rangle^{-\frac{N}{2}}$$

$$\times \langle |2^{(2-\alpha_j)j}(k_2+x_2-2^{\alpha_j j}y_2)| \rangle^{-\frac{N}{2}} \ldots \langle |2^{(2-\alpha_j)j}(k_{D-1}+x_{D-1}-2^{\alpha_j j}y_{D-1})| \rangle^{-\frac{N}{2}} dy$$

$$\overset{2^{(2-\alpha_j)j}\geq 1}{\leq} \int_{\mathbb{R}^{D-1}} |w(y_1,\ldots,y_{D-1})| \underbrace{\sum_{k_1\in\mathbb{Z}} \langle |k_1+x_1-2^{\alpha_j j}y_1| \rangle^{-\frac{N}{2}}}_{\leq c_{N_1}}$$

$$\times \underbrace{\sum_{k_2\in\mathbb{Z}} \langle |k_2+x_2-2^{\alpha_j j}y_2| \rangle^{-\frac{N}{2}}}_{\leq c_{N_2}} \ldots \underbrace{\sum_{k_{D-1}\in\mathbb{Z}} \langle |k_{D-1}+x_{D-1}-2^{\alpha_j j}y_{D-1}| \rangle^{-\frac{N}{2}}}_{\leq c_{N_{D-1}}} dy$$

$$\leq c_N.$$

We repeatedly apply a similar computation as in (13), hence

$$\sum_{\substack{k\in\mathbb{Z}^D,\ |l_1|\ldots|l_{D-1}|\leq 1\\ |\mathbf{k}-k_D|>2^{\varepsilon j}}} \left|\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\rangle\right|$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \sum_{\substack{|l_1|\ldots|l_{D-1}|\leq 1\\ k_D\in\mathbb{Z}\ |k_D|>2^{\varepsilon j}}} \int_{\mathbb{R}^{D-1}} \langle |x_1| \rangle^{-N} \ldots \langle |x_{D-1}| \rangle^{-N}$$

$$\times \langle |l_1x_1+\cdots+l_{D-1}x_{D-1}-k_D| \rangle^{-N} dx_1\ldots dx_{D-1}$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \sum_{\substack{|l_2|\ldots|l_{D-1}|\leq 1\\ k_D\in\mathbb{Z}\ |k_D|>2^{\varepsilon j}}} \int_{\mathbb{R}^{D-2}} \langle |x_2| \rangle^{-N} \ldots \langle |x_{D-1}| \rangle^{-N}$$

$$\times \langle |l_2x_2+\cdots+l_{D-1}x_{D-1}-k_D| \rangle^{-N} dx_2\ldots dx_{D-1}.$$

Finally, we obtain

$$\sum_{\substack{k\in\mathbb{Z}^D,\ |l_1|\ldots|l_{D-1}|\leq 1\\ |\mathbf{k}-k_D|>2^{\varepsilon j}}} \left|\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\rangle\right|$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \sum_{\substack{|l_{D-1}|<1\\ |k_D|>2^{\varepsilon j},\ k_D\in\mathbb{Z}}} \int_{\mathbb{R}} \langle |x_{D-1}| \rangle^{-N} \langle |l_{D-1}x_{D-1}-k_D| \rangle^{-N} dx_{D-1}$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \sum_{k_D\in\mathbb{Z},\ |k_D|>2^{\varepsilon j}} \langle |k_D| \rangle^{-N}$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \int_{|x_D|>2^{\varepsilon j}} \langle |x_D| \rangle^{-N} dx_D$$

$$\leq c_N 2^{(D-1)(6-\alpha_j)\frac{j}{2}} \times 2^{-(N-1)\varepsilon j}.$$

For $N \geq 2$, we apply Lemma 1, and conclude the following result

$$\sum_{\substack{1<|l_1|\leq 2^{(2-\alpha_j)j},\ k\in\mathbb{Z}^D \\ |l_2|\ldots|l_{D-1}|\leq 1}} \left|\left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\right\rangle\right|$$

$$\leq c_{N,M} \sum_{\substack{k\in\mathbb{Z}^D,\ 1<|l_1|\leq 2^{(2-\alpha_j)} \\ |l_2|\ldots|l_{D-1}|\leq 1}} |t_1|^{-N}\ldots|t_D|^{-N} 2^{-M\alpha_j j}\big(|l_1|-1\big)^{-M} 2^{(2-(D-1)\alpha_j)j} 2^{-N\alpha_j j}.$$

For some index $(j, l, k, \alpha_j, D) \in \Gamma_j$, we have

$$\sum_{\substack{k\in\mathbb{Z}^D \\ t_1\neq 0,\ldots,t_D\neq 0}} |t_1|^{-N_0}\ldots|t_D|^{N_0}$$

$$= \sum_{\substack{k\in\mathbb{Z}^D\ k_1\neq 0,\ldots,k_{D-1}\neq 0 \\ k_D\neq l_1k_1+\cdots+l_{D-1}k_{D-1}}} \left|2^{-\alpha_j j}k_1\right|^{-N}\ldots\left|2^{-\alpha_j j}k_{D-1}\right|^{-N}\left|2^{-2j}(k_D-l_1k_1\ldots-l_{D-1}k_{D-1})\right|^{-N}$$

$$= 2^{((D-1)N\alpha_j+2N)j} \sum_{\substack{k\in\mathbb{Z}^D \\ k_1,\ldots,k_D\neq 0}} |k_1|^{-N}\ldots|k_{D-1}|^{-N}|k_D|^{-N}$$

$$\leq 2^{((D-1)N\alpha_j+2N)j} \int_{|x_1|\geq 1\ldots|x_D|\geq 1} |x_1|^{-N}\ldots|x_D|^{-N}dx$$

$$\leq c_N 2^{((D-1)N\alpha_j+2N)j}.$$

Similarly, for $t_D = 0$ (or/and $t_i = 0$, $i = 1, \ldots, D-1$) we obtain

$$\sum_{k\in\mathbb{Z}^D} |t_1|^{-N}\ldots|t_D|^{-N} \leq c_N 2^{((D-1)N\alpha_j+2N)j}.$$

Finally, we have

$$\sum_{\substack{1<|l_1|\leq 2^{(2-\alpha_j)j},\ k\in\mathbb{Z}^D \\ |l_2|\ldots|l_{D-1}|\leq 1}} \left|\left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\right\rangle\right|$$

$$\leq c_{N,M} 2^{(6-(D+1)\alpha_j)\frac{j}{2}} \times 2^{-M\alpha_j j} \times 2^{2N_j} \times 2^{DN\alpha_j j} \underbrace{\big(|l_1|-1\big)^{-M}}_{\leq 1}$$

$$\leq c_{N,M} 2^{(6-(D+1)\alpha_j)\frac{j}{2}} 2^{N(D\alpha_j+2)j} \times 2^{-M\alpha_j j}.$$

Using $\liminf_{j\to\infty} \alpha_j > 0$ and choosing M sufficiently large, we obtain desired decay. An analogous argument holds for other interior regions and the boundary elements. □

Now, we investigate the cluster coherence $\mu_c(\Lambda_j^{\mp 1}, \chi_{\mathcal{M}_{h_j}}\Psi)$ and show that, it converges to zero as $j \to \infty$ when $(h)_j \in o(2^{-((D-1)\alpha_j+\varepsilon)j})$.

Lemma 2 *If* $h_j = (h_j^{x_1} \times \cdots \times h_j^{x_{D-1}})_j \in o(2^{-((D-1)\alpha_j+\varepsilon)j})$, *then*

$$\mu_c\big(\Lambda_j^{\mp 1}, \chi_{\mathcal{M}_{h_j}}\Psi\big) \to 0, \quad j \to \infty.$$

Proof We have

$$\mu_c\big(\Lambda_j^{\mp 1}, \chi_{\mathcal{M}_{h_j}}\Psi\big) \le \mu_c(\Lambda_{j-1}, \chi_{\mathcal{M}_{h_j}}\Psi) + \mu_c(\Lambda_j, \chi_{\mathcal{M}_{h_j}}\Psi) + \mu_c(\Lambda_{j+1}, \chi_{\mathcal{M}_{h_j}}\Psi).$$

By the definition of the cluster coherence and considering the main-scale-term, we obtain

$$\mu_c(\Lambda_j, \chi_{\mathcal{M}_{h_j}}\psi) = \max_{\gamma_2\in\Gamma}\sum_{\gamma_1\in\Lambda_j}|\langle\chi_{\mathcal{M}_{h_j}}\psi_{\gamma_1}, \psi_{\gamma_2}\rangle| \le \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d=D}}\sum_{\gamma_1\in\Lambda_j}|\langle\chi_{\mathcal{M}_{h_j}}\psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_D}$$

$$+ \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d\neq D}}\sum_{\gamma_1\in\Lambda_j}|\langle\chi_{\mathcal{M}_{h_j}}\psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_d} + \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ b}}\sum_{\gamma_1\in\Lambda_j}|\langle\chi_{\mathcal{M}_{h_j}}\psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{:=T_b}.$$

We use the rapid decay properties of the universal shearlets to obtain desired result. Set $\gamma_1 = (j, l, k, \alpha_j, D) \in \Lambda_j$, $\gamma_2 \in (j, l', k', \alpha_j, D) \in \Gamma_j$ we have $(N \ge 2)$

$$\begin{aligned}
&\big|\big\langle\chi_{\mathcal{M}_{h_j}}\psi_{j,l,k}^{\alpha_j,(D)}, \psi_{j,l',k'}^{\alpha_j,(D)}\big\rangle\big| \\
&\quad\le c_N 2^{((D-1)\alpha_j+2)j}\int_{-h_j^{x_1}}^{h_j^{x_1}}\cdots\int_{-h_j^{x_{D-1}}}^{h_j^{x_{D-1}}}\int_{\mathbb{R}}\big\langle\big|S_{(D)}^{l}A_{\alpha_j,(D)}^{j}x-k\big|\big\rangle^{-N}\big\langle\big|S_{(D)}^{l'}A_{\alpha_j,(D)}^{j}x-k'\big|\big\rangle^{-N}dx \\
&\quad\le c_N\int_{-h_j^{x_1}}^{h_j^{x_1}}\cdots\int_{-h_j^{x_{D-1}}}^{h_j^{x_{D-1}}}\int_{\mathbb{R}}\big\langle\big|2^{\alpha_j j}x_1-k_1\big|\big\rangle^{-N}\dots\big\langle\big|2^{\alpha_j j}x_{D-1}-k_{D-1}\big|\big\rangle^{-N} \\
&\qquad\times\big\langle\big|2^{\alpha_j j}l_1x_1+\cdots-k_D\big|\big\rangle^{-N}\big\langle\big|2^{\alpha_j j}x_1-k_1'\big|\big\rangle^{-N}\dots\big\langle\big|2^{\alpha_j j}x_{D-1}-k_{D-1}'\big|\big\rangle^{-N} \\
&\qquad\times\big\langle\big|2^{\alpha_j j}l_1'x_1+\cdots+2^{2j}x_D-k_D'\big|\big\rangle^{-N}2^{((D-1)\alpha_j+2)j}dx_D\dots dx_1 \\
&\quad\le c_N\int_{-2^{\alpha_j j}h_j^{x_1}}^{2^{\alpha_j j}h_j^{x_1}}\cdots\int_{-2^{\alpha_j j}h_j^{x_{D-1}}}^{-2^{\alpha_j j}h_j^{x_{D-1}}}\int_{\mathbb{R}}\langle|x_1-k_1|\rangle^{-N}\dots\langle|x_{D-1}-k_{D-1}|\rangle^{-N} \\
&\qquad\times\langle|l_1x_1+\cdots+x_D-k_D|\rangle^{-N}\big\langle\underbrace{|x_1-k_1'|}_{\le 1}\big\rangle^{-N}\dots\big\langle\underbrace{|x_{D-1}-k_{D-1}'|}_{\le 1}\big\rangle^{-N} \\
&\qquad\times\big\langle\big|l_1'x_1+l_2'x_2+\cdots+l_{D-1}'x_{D-1}-k_D'\big|\big\rangle^{-N}dx_D\dots dx_1.
\end{aligned}$$

We may presume that the maximum of T_D is attained for some $\gamma_2 \in \Gamma_j$. Thus

$$\begin{aligned}
T_D &= \max_{\substack{\gamma_2\in\Gamma\\ d=D}}\sum_{\gamma_1\in\Lambda_j}|\langle\chi_{\mathcal{M}_{h_j}}\psi_{\gamma_1}, \psi_{\gamma_2}\rangle| \\
&\le \max_{(j,l',\alpha_j,D)\in\Gamma_j}\sum_{|l_1|\dots|l_{D-1}|\le 1}\sum_{k\in\mathbb{Z}^D}\big|\big\langle\chi_{\mathcal{M}_{h_j}}\psi_{j,l,k}^{\alpha_j(D)}, \psi_{j,l',k}^{\alpha_j,(D)}\big\rangle\big| \\
&\le c_N\max_{(j,l',k',\alpha_j,\upsilon)\in\Gamma_j}\sum_{|l_1|\dots|l_{D-1}|\le 1}\int_{-2^{\alpha_j j}h_j^{x_1}}^{2^{\alpha_j j}h_j^{x_1}}\cdots\int_{-2^{\alpha_j j}h_j^{x_{D-1}}}^{-2^{\alpha_j j}h_j^{x_{D-1}}}\int_{\mathbb{R}}\big\langle\big|l_1'x_1+\cdots-k_D'\big|\big\rangle^{-N}
\end{aligned}$$

$$\times \underbrace{\sum_{k \in \mathbb{Z}^D} \langle |x_1 - k_1| \rangle^{-N} \dots \langle |x_{D-1} - k_{D-1}| \rangle^{-N} \langle |l_1 x_1 + \dots - k_D| \rangle^{-N}}_{\leq \tilde{c}_N} dx_D \dots dx_1$$

$$\leq c_N \max_{(j,l',k',\alpha_j,\upsilon) \in \Gamma_j} \int_{-2^{\alpha_j j} h_j^{x_1}}^{2^{\alpha_j j} h_j^{x_1}} \dots \int_{-2^{\alpha_j j} h_j^{x_{D-1}}}^{-2^{\alpha_j j} h_j^{x_{D-1}}} \underbrace{\int_{\mathbb{R}} \langle |l_1' x_1 + \dots - k_D'| \rangle^{-N} dx_D}_{\leq \tilde{c}_N} \dots dx_1$$

$$\leq c_N 2^{(D-1)\alpha_j j} h_j^{x_1} \times \dots \times h_j^{x_{D-1}} \to 0, \quad j \to \infty.$$

For the factor T_d where $d \neq D$, let use $\gamma_1 = (j, l, k, \alpha_j, D) \in \Lambda_j$ and $\gamma_2 = (j, l', k', \alpha_j, d) \in \Gamma_j$. For $N \geq 2$, we have

$$\begin{aligned}
&\left|\left\langle \chi_{\mathcal{M}_{h_j}} \psi_{j,l,k}^{\alpha_j,(D)}, \psi_{j,l',k'}^{\alpha_j,(d)} \right\rangle\right| \\
&\quad \leq \int_{-h_j^{x_1}}^{h_j^{x_1}} \dots \int_{-h_j^{x_{D-1}}}^{h_j^{x_{D-1}}} \int_{\mathbb{R}} \left|\psi_{j,l,k}^{\alpha_j,(D)}\right| \left|\psi_{j,l',k'}^{\alpha_j,(d)}\right| dx_D \dots dx_1 \\
&\quad \leq c_N 2^{((D-1)\alpha_j+2)j} \int_{-h_j^{x_1}}^{h_j^{x_1}} \dots \int_{\mathbb{R}} \left\langle \left|2^{\alpha_j j} x_1 - k_1\right| \right\rangle^{-N} \dots \left\langle \left|2^{\alpha_j j} x_{D-1} - k_{D-1}\right| \right\rangle^{-N} \\
&\qquad \times \left\langle \left|2^{\alpha_j j} l_1 x_1 + \dots - k_D\right| \right\rangle^{-N} \underbrace{\left\langle \left|2^{\alpha_j j} x_1 - k_1'\right| \right\rangle^{-N} \dots \left\langle \left|2^{\alpha_j j} x_D - k_D'\right| \right\rangle^{-N}}_{\leq 1} dx \\
&\quad \leq c_N 2^{(D-1)\alpha_j j} \int_{-h_j^{x_1}}^{h_j^{x_1}} \dots \int_{-h_j^{x_{D-1}}}^{h_j^{x_{D-1}}} \left\langle \left|2^{\alpha_j j} x_1 - k_1\right| \right\rangle^{-N} \dots \left\langle \left|2^{\alpha_j j} x_{D-1} - k_{D-1}\right| \right\rangle^{-N} \\
&\qquad \times \Bigg(\underbrace{\int_{\mathbb{R}} \left\langle \left|2^{\alpha_j j} l_1 x_1 + \dots + x_D - k_D\right| \right\rangle^{-N} dx_D}_{\leq c_N} \Bigg) dx_{D-1} \dots dx_1.
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
T_d &\leq \max_{(j,l',k',\alpha_j,d) \in \Gamma_j} \sum_{|l_1|,\dots,|l_{D-1}| \leq 1} \sum_{\substack{k \in \mathbb{Z}^D \\ |l_1 k_1 + \dots l_{D-1} k_{D-1} - k_D| < 2^{\varepsilon j}}} \left|\left\langle \chi_{\mathcal{M}_{h_j}} \psi_{j,l,k}^{\alpha_j,(D)}, \psi_{j,l',k'}^{\alpha_j,(d)} \right\rangle\right| \\
&\leq c_N 2^{(D-1)\alpha_j j} 2^{\varepsilon j} \\
&\quad \times \int_{-h_j^{x_1}}^{h_j^{x_1}} \underbrace{\sum_{k_1 \in \mathbb{Z}} \left\langle \left|2^{\alpha_j j} x_1 - k_1\right| \right\rangle^{-N} dx_1}_{\leq \tilde{c}_N} \dots \int_{-h_j^{x_{D-1}}}^{h_j^{x_{D-1}}} \underbrace{\sum_{k_{D-1} \in \mathbb{Z}} \left\langle \left|2^{\alpha_j j} x_{D-1} - k_{D-1}\right| \right\rangle^{-N}}_{\leq \tilde{c}_N} dx_{D-1} \\
&\leq c_N 2^{((D-1)\alpha_j + \varepsilon)j} h_j^{x_1} \times \dots \times h_j^{x_{D-1}} \to 0, \quad j \to \infty.
\end{aligned}$$

For the boundary shearlet elements, it is easy to display that there exists a constant $c > 0$ such that $T_b \leq c(T_d + T_D)$. So the aim obtain from the first two cases. □

Now, we apply the abstract error estimate of Theorem 1 to obtain the fallowing theorem, which is one of the main results of this paper and proves the success of image inpainting via a high-dimensional universal shearlet. Theorem 3 relates the degree of anisotropic scaling to the admissible gap sizes: When α_j becomes smaller, the asymptotical condition $(h_j)_j \in o(2^{-((D-1)\alpha_j+\varepsilon)j})$ is satisfied for larger h_j.

We choose $(\alpha_j)_j$ as large possible, namely $\alpha_j := 2 - \frac{1}{j}$ for $j \geq 1$, then we have $2^{\alpha_j j} \in \Theta(2^{2j})$. This particularly concludes that the element of $\mathrm{SH}(\phi, \upsilon, (\alpha_j)_j)$ scale in an isotropic manner and can be viewed as a special kind of wavelet frame. Therefore, in terms of inpainting, shearlet frames are superior to isotropic wavelet systems, increasing the degree of anisotropy improving the recovery error of inpainting. Indeed, the necessary condition provides a deep insight into the relation between the degree of anisotropy of the underlying system and the admissible gap size and it shows the structural difference between wavelets and shearlets.

Theorem 3 *Let $(\alpha_j)_j$ be a scaling sequence and $\Psi = \mathrm{SH}(\phi, \upsilon, (\alpha_j)_j)$ be a universal shearlet system. We assume that the following conditions are satisfied*:

(*INP1*) *There exist $\delta > 0$ and $j_0 \in \mathbb{N}$ such that $\alpha_j > \delta$ for all $j \geq j_0$, i.e.*,

$$\lim_{j\to\infty} \inf \alpha_j > 0.$$

(*INP2*) *For a fixed $\varepsilon > 0$, the sequence of gap widths satisfies*

$$h_j = \left(h_j^{x_1} \times \cdots \times h_j^{x_{D-1}}\right)_j \in o\left(2^{-((D-1)\alpha_j+\varepsilon)j}\right).$$

Then

$$\left(\frac{\|w\mathcal{L}_j^* - w\mathcal{L}_j\|_{1,\Psi}}{\|w\mathcal{L}_j\|_{1,\Psi}}\right)_j \in o\left(2^{-Nj}\right), \quad \text{as } j \to \infty$$

for every $N \in \mathbb{N}_0$. Here, the recovery provided by Algorithm 1 is denoted by $w\mathcal{L}_j^*$.

Proof A similar computations as in the proof the Lemma 1 show that there exists an $N \geq 0$ such that $\|w\mathcal{L}_j\|_{1,\Psi} \in o(2^{Nj})$ as $j \to \infty$. In addition, it is easy to obtain a lower bound for this term. Now we apply the abstract error estimate of Theorem 1, Proposition 1 and Lemma 2 to obtain the desired result. □

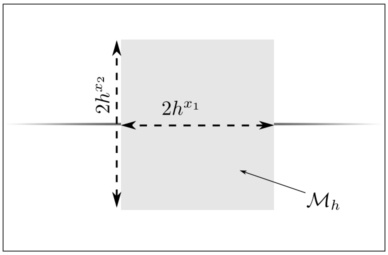
We would like to remark at this point that the condition (INP2) on the gap width in Theorem 3 is only the sufficient conditions for asymptotically perfect inpainting, for instance, the superiority of shearlets over wavelets also need a necessary condition. To tackle this problem one could use another inpainting which was analyzed in [16].

Also, we believe that a weaker version of $(h_j)_j \in o(2^{-((D-1)\alpha_j+\varepsilon)j})$ is necessary to achieve asymptotically perfect inpainting. In order to give an argument for this conjecture, we observe that $\psi_{j,l,k}^{\alpha_j,D}$ is concentrated inside of $\mathcal{M}_{h_j}$ provided that $|k_i| \leq 2^{\alpha_j j} h_j^i$ for $i = 1, \ldots, D-1$. Therefore the missing part might be estimated by

$$\begin{aligned}\|P_M w\mathcal{L}_j\|_{1,\Psi} &= \sum_{\psi\in\Psi} |\langle P_M w\mathcal{L}_j, \psi\rangle| \gtrsim \sum_{(j,k,\alpha,D)\in\Lambda_j,\ l_1\ldots l_D=0} \left|\left\langle w\mathcal{L}_j, \psi_{j,l,k}^{\alpha_j,(D)}\right\rangle\right| \\ &\simeq 2^{\varepsilon j} 2^{\alpha_j j} h_j 2^{(D-1)(6-\alpha_j)\frac{j}{2}} = 2^{((D-1)6+2\varepsilon)\frac{j}{2}} 2^{(D-1)\alpha_j \frac{j}{2}} h_j\end{aligned}$$

This display that enlarging h_j produces a lack of information and will cause inappropriate error for recovery by ℓ^1-minimization.

Fig. 6 Sketch of the corrupted modeling image



4 Inpainting with Specific Image Model on $L^2(\mathbb{R}^2)$

The general approach in this section is the same as in the previous one. We investigate the inpainting results of ℓ^1 minimization by chosen proper index set Λ_j and estimate the relative sparsity and cluster coherence respect to this set.

At first, we would like to analyze a specific mathematical model which is the model of corrupted line segments. Let $w \in C^\infty(\mathbb{R}^2)$ be a function that is supported in $[-\rho, \rho] \times [-\eta, \eta]$ where $\rho, \eta > 0$. A whole sequence of models $(w_j)_{j \geq 0}$ is given by

$$w_j(x) = w * F_j(x) = \langle w, F_j(x - [\cdot])\rangle, \quad x \in \mathbb{R}^2,$$

where filters F_j are defined by the inverse Fourier transform of the corona functions (2).

Now, we define the mask of a missing part of image as follows. The mask $\mathcal{M}_h$ is the intersection of a small vertical strip around the x_2-axis and a small horizontal strip around the x_1-axis which is given by

$$\mathcal{M}_h = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| \leq h^{x_1}, |x_2| \leq h^{x_2}\}.$$

For fix some $\varepsilon > 0$, we define the clusters

$$\Lambda_j = \{(j, l, k, \alpha_j, d) : \ |l| < 1, |k_2| < 2^{\varepsilon j}, k \in \mathbb{Z}^2, d = 1, 2\}, \quad j \geq 0.$$

We may determine the relative sparsity of the shearlet coefficient with respect to the cluster $\Lambda_j^{\pm 1}$.

Lemma 3 *Assuming that* $\liminf_{j\to\infty} \alpha_j > 0$ (*recall that* α_j *is scaling parameter*), *then*

$$\delta_j \in o(2^{-Nj}), \quad j \to \infty,$$

for every $N \in \mathbb{N}$.

Proof By the definition, we have

$$\begin{aligned}\delta_j = &\sum_{\substack{k\in\mathbb{Z}^2, |l|\leqslant 1\\ |k_2|>2^{\varepsilon j}}} |\langle w_j, \psi_{j,l,k}^{\alpha_j,(1)}\rangle| + \sum_{\substack{k\in\mathbb{Z}^2\\ |l|>1}} |\langle w_j, \psi_{j,l,k}^{\alpha_j,(1)}\rangle| \\ &+ \sum_{\substack{k\in\mathbb{Z}^2, |l|\leqslant 1\\ |k_2|>2^{\varepsilon j}}} |\langle w_j, \psi_{j,l,k}^{\alpha_j,(2)}\rangle| + \sum_{\substack{k\in\mathbb{Z}^2\\ |l|>1}} |\langle w_j, \psi_{j,l,k}^{\alpha_j,(2)}\rangle| + \sum_{\substack{k\in\mathbb{Z}^2\\ |l|=2^{(2-\alpha_j)j}}} |\langle w_j, \psi_{j,l,k}^{\alpha_j,(b)}\rangle| \\ = &\, T_1 + T_2 + T_3 + T_4 + T_5.\end{aligned}$$

Let $\alpha_j \geqslant o$ and $|l| > 1$. By Plancherel's Theorem, we obtain

$$\langle w_j, \psi_{j,l,k}^{\alpha_j,(1)} \rangle = \int_{\mathbb{R}} e^{2\pi i t_2 \xi_2} \int_{\mathbb{R}} \widehat{w}(\xi) \underbrace{W(2^{-2j}\xi)\widehat{\psi}_{j,l,0}^{\alpha_j,(1)}(\xi)}_{=\sigma_{j,l}(\xi)} e^{2\pi i t_1 \xi_1} d\xi_1 d\xi_2.$$

We repeatedly apply integration by parts, hence

$$|\langle w_j, \psi_{j,l,k}^{\alpha_j(1)} \rangle| \le c_{L,M} |t_1|^{-L} |t_2|^{-M} \|h_{L,M}\|_{L^1(\mathbb{R}^2)},$$

where $h_{L,M}(\xi_1, \xi_2) = \int_{\mathbb{R}^2} |D^{(L,M)} \widehat{w}(\xi_1, \xi_2) \sigma_{j,l}(\xi_1, \xi_2)| d\xi_1 d\xi_2$. By a similar argumentation as proof Lemma 1 in part (1), we can estimate the term $h_{L,M}(\xi_1, \xi_2)$ as of the of following

$$h_{L,M}(\xi_1, \xi_2) \le c_{L,N} 2^{-(2+\alpha_j)\frac{j}{2}} \langle |2^{\alpha_j j}| \rangle^{-N} \langle |2^{2j}| \rangle^{-N} 2^{-L\alpha_j j}.$$

The final estimate gives

$$|\langle w_j, \psi_{j,l,k}^{\alpha_j(1)} \rangle| \le c_{L,M,N} |t_1|^{-L} |t_2|^{-M} 2^{-(2+\alpha_j)\frac{j}{2}} \langle |2^{\alpha_j j}| \rangle^{-N} \langle |2^{2j}| \rangle^{-N} 2^{-L\alpha_j j}.$$

Now we can estimate T_2. For fix $j > o$ and some index $(j, l, k, \alpha_j, 1) \in \Gamma_j$, we have $(N \ge 2)$

$$\sum_{k \in \mathbb{Z}^2} |t_1|^{-L} |t_2|^{-M} \le c_N 2^{(L\alpha_j + 2M)j}.$$

Now, we can compute T_2:

$$\begin{aligned} T_2 &\le c_{L,M,N} \sum_{\substack{k \in \mathbb{Z}^2 \\ 1 < |l| \le 2^{(2-\alpha_j)j}}} |t_1|^{-L} |t_2|^{-M} 2^{-(2+\alpha_j)\frac{j}{2}} \langle |2^{\alpha_j j}| \rangle^{-N} \langle |2^{2j}| \rangle^{-N} 2^{-L\alpha_j j} \\ &\le c_{M,N} 2^{-(2+\alpha_j)\frac{j}{2}} \langle |2^{\alpha_j j}| \rangle^{-N} \langle |2^{2j}| \rangle^{-N} 2^{2Mj}. \end{aligned}$$

The assumption $\liminf_{j \to \infty} \alpha_j > 0$ and choice N sufficiently large, imply the desired result. Note that, the strategy for T_4 and T_5 are analogous, therefore we remove at this point. In the following, we estimate T_3 and T_4 are done similarly.

To estimate T_3, we benefit of rapid decay of shearlet elements and model w_j. Let $\alpha_j \ge 0$ and $|l| \le 1$,

$$|\langle w_j, \psi_{j,l,k}^{\alpha_j(1)} \rangle| \le c_N 2^{(2+\alpha_j)\frac{j}{2}} \int_{\mathbb{R}^2} \langle |x_1| \rangle^{-N} \langle |x_2| \rangle^{-N} \langle |x_1 - k_1| \rangle^{-N} \langle |l x_1 + x_2 - k_2| \rangle^{-N} dx,$$

with this, we have

$$\begin{aligned} T_3 &\le c_N 2^{(2+\alpha_j)\frac{j}{2}} \sum_{\substack{k_2 \in \mathbb{Z}, |l| \le 1 \\ |k_2| > 2^{\varepsilon j}}} \int_{\mathbb{R}^2} \langle |x_1| \rangle^{-N} \langle |x_2| \rangle^{-N} \underbrace{\sum_{k_1 \in \mathbb{Z}} \langle |x_1 - k_1| \rangle^{-N}}_{\le c_N} \langle |l x_1 + x_2 - k_2| \rangle^{-N} dx \\ &\le c_N 2^{(2+\alpha_j)\frac{j}{2}} \sum_{\substack{k_2 \in \mathbb{Z}, |l| \leqslant 1 \\ |k_2| > 2^{\varepsilon j}}} \int_{\mathbb{R}} \langle |x_1| \rangle^{-N} \int_{\mathbb{R}} \langle |x_2| \rangle^{-N} \langle |l x_1 + x_2 - k_2| \rangle^{-N} dx. \end{aligned}$$

The factors of the integral can be compute by (13), so we obtain

$$
\begin{aligned}
T_3 &\leq c_N 2^{(2+\alpha_j)\frac{j}{2}} \sum_{\substack{k_2\in\mathbb{Z},|l|\leqslant 1\\ |k_2|>2^{\varepsilon j}}} \int_{\mathbb{R}} \langle|x_1|\rangle^{-N} \langle|lx_1 - k_2|\rangle^{-N} dx_1 \\
&\leq c_N 2^{(2+\alpha_j)\frac{j}{2}} \sum_{\substack{k_2\in\mathbb{Z}\\ |k_2|>2^{\varepsilon j}}} \langle|k_2|\rangle^{-N} \\
&\leq c_N 2^{(2+\alpha_j)\frac{j}{2}} 2^{-(N-1)\varepsilon j}.
\end{aligned}
$$

Choosing N arbitrary large proving the claim. □

Next we estimate the cluster coherence and show that the size of the gaps which can be filled by the geometric shape of clusters $\Lambda_j^{\pm 1}$.

Lemma 4 *Presume that* $(h_j^{x_1} \times h_j^{x_2})_j \in o(2^{-(\alpha_j+2+\varepsilon)j})$, *we have*

$$
\mu_c\big(\Lambda_j^{\mp 1}, \chi_{\mathcal{M}_{h_j}} \Psi\big) \to 0, \quad j \to \infty.
$$

Proof We will investigate the main-scale-term:

$$
\begin{aligned}
\mu_c(\Lambda_j, \chi_{\mathcal{M}_{h_j}} \psi) \leq{} & \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d=1}} \sum_{\substack{\gamma_1\in\Lambda_j\\ d=1}} |\langle \chi_{\mathcal{M}_{h_j}} \psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_1} + \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d=1}} \sum_{\substack{\gamma_1\in\Lambda_j\\ d=2}} |\langle \chi_{\mathcal{M}_{h_j}} \psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_2} \\
& + \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d=2}} \sum_{\substack{\gamma_1\in\Lambda_j\\ d=1}} |\langle \chi_{\mathcal{M}_{h_j}} \psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_3} + \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ d=2}} \sum_{\substack{\gamma_1\in\Lambda_j\\ d=2}} |\langle \chi_{\mathcal{M}_{h_j}} \psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{=:T_4} \\
& + \underbrace{\max_{\substack{\gamma_2\in\Gamma\\ b}} \sum_{\gamma_1\in\Lambda_j} |\langle \chi_{\mathcal{M}_{h_j}} \psi_{\gamma_1}, \psi_{\gamma_2}\rangle|}_{:=T_b}.
\end{aligned}
$$

Let $\gamma_1 = (j, l, k, \alpha_j, 1) \in \Lambda_j$, $\gamma_2 \in (j, l', k', \alpha_j, 1) \in \Gamma_j$. We obtain $(N \geq 2)$

$$
\begin{aligned}
&\big|\big\langle \chi_{\mathcal{M}_{h_j}} \psi_{j,l,k}^{\alpha_j,(1)}, \psi_{j,l',k'}^{\alpha_j,(1)}\big\rangle\big| \\
&\quad\leq c_N \int_{-2^j h_j^{x_1}}^{2^j h_j^{x_1}} \int_{-2^{\alpha_j j} h_j^{x_2}}^{-2^{\alpha_j j} h_j^{x_2}} \langle|x_1 + lx_2 - k_1|\rangle^{-N} \\
&\qquad \times \langle|x_2 - k_2|\rangle^{-N} \underbrace{\big\langle|x_1 + l'x_2 - k_1'|\big\rangle^{-N}}_{\leq 1} \big\langle|x_2 - k_2'|\big\rangle^{-N} dx_2 dx_1.
\end{aligned}
$$

Hence, we have

$$T_1 \le c_N \max_{(j,l',k',\alpha_j,1)\in\Gamma_j} \sum_{|l|\le 1} \int_{-2^j h_j^{x_1}}^{2^j h_j^{x_1}} \int_{-2^{\alpha_j j} h_j^{x_2}}^{-2^{\alpha_j j} h_j^{x_2}} \left\langle |x_2 - k_2'| \right\rangle^{-N}$$

$$\times \underbrace{\sum_{k\in\mathbb{Z}^2} \langle |x_1 + l x_2 - k_1| \rangle^{-N} \langle |x_2 - k_2| \rangle^{-N} \, dx_2 dx_1}_{\le c_N}$$

$$\le c_N \max_{(j,l',k',\alpha_j,1)\in\Gamma_j} \int_{-2^j h_j^{x_1}}^{2^j h_j^{x_1}} \int_{-2^{\alpha_j j} h_j^{x_2}}^{-2^{\alpha_j j} h_j^{x_2}} \underbrace{\left\langle |x_2 - k_2'| \right\rangle^{-N}}_{\le 1} dx_2 dx_1$$

$$\le c_N 2^{(\alpha_j+2)j} h_j^{x_1} \times h_j^{x_2} \to 0, \quad j \to \infty.$$

For the factor T_2, let $\gamma_1 = (j, l, k, \alpha_j, 2) \in \Lambda_j$, $\gamma_2 \in (j, l', k', \alpha_j, 1) \in \Gamma_j$. For $N \ge 2$, we can compute as following:

$$\left| \left\langle \chi_{\mathcal{M}_{h_j}} \psi_{j,l,k}^{\alpha_j,(2)}, \psi_{j,l',k'}^{\alpha_j,(1)} \right\rangle \right|$$

$$\le c_N 2^{(\alpha_j+2)j} \int_{h_j^{x_1}}^{h_j^{x_1}} \int_{h_j^{x_2}}^{h_j^{x_2}} \left\langle \left| 2^{\alpha_j j} x_1 - k_1 \right| \right\rangle^{-N}$$

$$\times \underbrace{\left\langle \left| l 2^{\alpha_j j} x_1 + 2^{2j} x_2 - k_2 \right| \right\rangle^{-N} \left\langle \left| l' 2^{\alpha_j j} x_2 + 2^{2j} x_1 - k_1' \right| \right\rangle^{-N}}_{\le 1} \left\langle \left| 2^{\alpha_j j} x_2 - k_2' \right| \right\rangle^{-N} dx.$$

Finally, we obtain

$$T_2 \le c_N 2^{(\alpha_j+2)j} \max_{(j,l',k',\alpha_j,1)\in\Gamma_j} \sum_{\substack{|l|\le 1 \\ |k_2| \le 2\varepsilon j}} \int_{-h_j^{x_1}}^{h_j^{x_1}} \int_{-h_j^{x_2}}^{h_j^{x_2}} \left\langle \left| 2^{\alpha_j j} x_2 - k_2' \right| \right\rangle^{-N}$$

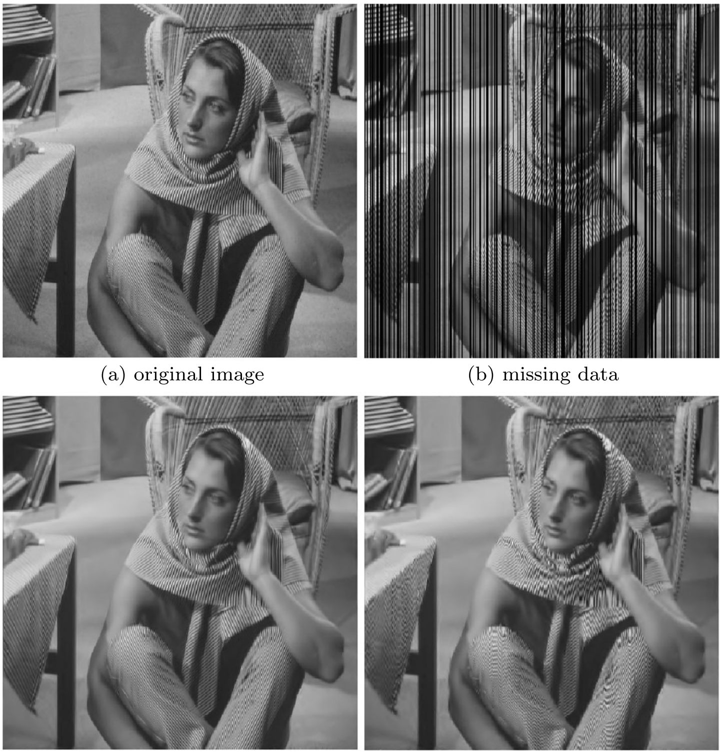
$$\times \underbrace{\sum_{k_1\in\mathbb{Z}} \left\langle \left| 2^{\alpha_j j} x_1 - k_1 \right| \right\rangle^{-N} dx_2 dx_1}_{\le c_N}$$

$$\le c_N 2^{(\alpha_j+2)j} 2^{\varepsilon j} \max_{(j,l',k',\alpha_j,1)\in\Gamma_j} \int_{-h_j^{x_1}}^{h_j^{x_1}} \int_{-h_j^{x_2}}^{h_j^{x_2}} \underbrace{\left\langle \left| 2^{\alpha_j j} x_2 - k_2' \right| \right\rangle^{-N}}_{\le 1} dx$$

$$\le c_N 2^{(\alpha_j+2+\varepsilon)j} h_j^{x_1} \times h_j^{x_2} \to 0, \quad j \to \infty.$$

Not that, the estimate for T_3 and T_4 are done similar to T_1 and T_2, respectively. By the definition of the boundary shearlets, it can estimate easily. □

Now, we can apply the error estimate of Theorem 1 to show the success of image inpainting with special image model. By considering the gap size as asymptotically smaller than the length of the corresponding shearlet elements, asymptotically perfect inpainting is achieved. Inpainting result for shearlets and wavelets in special cases can be found in [9, 16].



(a) original image (b) missing data

(c) shearlet (PSNR=**29.52**) (d) wavelet (PSNR=26.01)

Fig. 7 (**a**) Original image. (**b**) Missing data. (**c**) Masked image inpainted with shearlets using iterative thresholding. (**d**) Masked image inpainted with wavelets using iterative thresholding

Theorem 4 *Let $(\alpha_j)_j$ be a scaling sequence and $\Psi = \mathrm{SH}(\phi, \upsilon, (\alpha_j)_j)$ be a universal shearlet system. If*

$$(INP1) \quad \lim_{j\to\infty} \inf \alpha_j > 0.$$

$$(INP2) \quad \text{For a fixed } \varepsilon > 0,\ h_j = \left(h_j^{x_1} \times h_j^{x_2}\right)_j \in o\left(2^{-(2+\alpha_j+\varepsilon)j}\right).$$

Then

$$\left(\frac{\|w_j^* - w_j\|_{1,\Psi}}{\|w_j\|_{1,\Psi}}\right)_j \in o\left(2^{-Nj}\right), \quad \text{as } j \to \infty$$

for every $N \in \mathbb{N}_0$, where the recover provided by Algorithm 1 is denoted by w_j^*.

Proof The claim follows from the abstract error estimate of Theorem 1, Lemma 3 and Lemma 4. □

Note that our goal was to show that results make sense and provide insight. In Theorem 4, we focused on the recovery of anisotropic fashion. However, one might ask about isotropic method such as wavelet which may lead to improve error estimate. This certainly

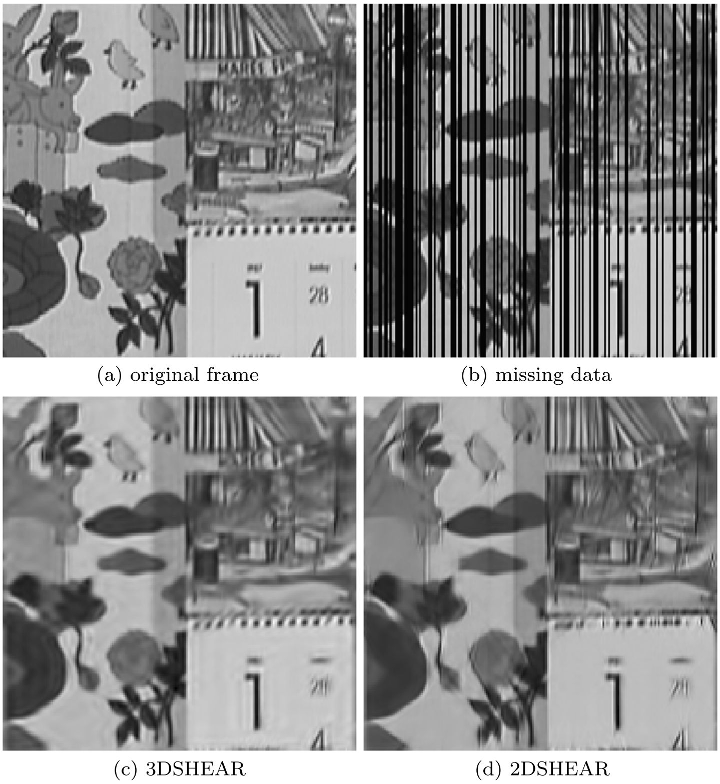


Fig. 8 Video inpainting of mobile video sequence. Starting from the *top left*: original frame, missing data, inpainted frame using 3DSHEAR (PSNR = **27.32**) and 2DSHEAR (PSNR = 24.12)

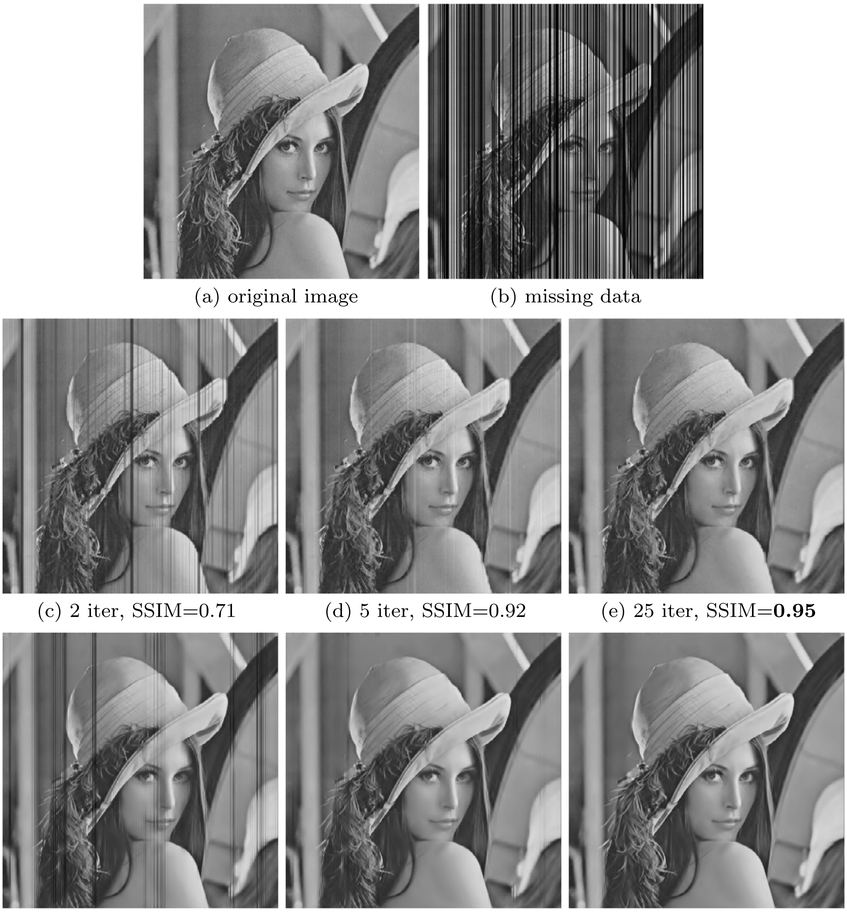
requires a careful adaption of the argument which is being investigated in our current research.

5 Numerical Experiments

This section is devoted to a set of numerical experiments. We focus on the inpainting problem and present a number of examples comparing various inpainting methods. In Fig. 7, vertical lines were removed from a grayscale Barbara image. The lines were inpainted using the methods from [15]. When using iterative thresholding to inpaint an image which consists almost completely of curvilinear features, shearlets outperform wavelets.

In Fig. 8 we compare the performance of the various video inpainting on a typical frame extracted from the inpainted video sequence. Although more subjective in nature, the figures show that the visual quality of the 3D shearlet frame is superior [18].

Our results are also of interest in the case considering the algorithm developed in [20] together with shearlet transforms (SIRL1 + TGV). We give a precise description of the crucial multiscale transform we used and the parameters. For shearlets, we have used the shearlet transform available at www.ShearLab.org. The discrete shearlet system is generated by using 4 scales and [1 1 2 2] for directional parameters. The chosen parameters for the proposed algorithm are:



(a) original image (b) missing data

(c) 2 iter, SSIM=0.71 (d) 5 iter, SSIM=0.92 (e) 25 iter, SSIM=**0.95**

(f) 10 iter, SSIM=0.74 (g) 20 iter, SSIM=0.89 (h) 400 iter, SSIM=**0.93**

Fig. 9 Comparison of iterative thresholding (IT) algorithm as used in the ShearLab package and proposed method. Both methods use the same shearlet system with 4 scales. Image (**a**) shows the original test image “Lena” and (**b**) displays the image distorted image. *Second row* shows intermediate results of the proposed algorithm combining iterative thresholding and TGV. *Last row* shows intermediate results of different iterations of the IT algorithm

- $\beta = 10^3$,
- $\alpha = [1, 2]$,
- $\mu = [10, 10, 20]$,
- $\varepsilon = 5 \times 10^{-5}$.

We have compared the results and parameters provided in the ShearLab package, based on iterative thresholding (IT) [18] with shearlets and total generalized variation (SIRL1 + TGV) [20]. To visualize the difference between the two methods, we demonstrate results of various number of iterations in Figs. 9 and 10. While both methods in total yield a similarly good performance, differences are clearly visible. The IT method smooths the signal more than the method of SIRL1 + TGV. There are still shadows of inpainting gaps visible in the results of SIRL1 + TGV, but the texture and sharp edges are preserved in a better way,


(a) original image (b) missing data

(c) 2 iter, SSIM=0.68 (d) 5 iter, SSIM=0.84 (e) 25 iter, SSIM=**0.98**

(f) 5 iter, SSIM=0.76 (g) 10 iter, SSIM=0.96 (h) 25 iter, SSIM=**0.97**

Fig. 10 Comparison of iterative thresholding (IT) algorithm as used in the ShearLab package and proposed method. Both methods use the same shearlet system with 4 scales. Image (**a**) shows the original test image and (**b**) displays the image distorted image. *Second row* shows intermediate results of the proposed algorithm combining iterative thresholding and TGV. *Last row* shows intermediate results of different iterations of the IT algorithm

which is in particular due to the additional TGV regularizer. Although proposed method is not optimized at all, it takes 440 s, for 25 iterations, while the lean method of IT needs 462 s, for computing the final result after 400 iterations.

Acknowledgements The first author would like to thank Professor Gitta Kutyniok for stimulating discussions and pointing out various references. She is grateful to the department of Mathematics, Technische Universität Berlin for hospitality and Martin Genzel for proofreading the paper.

References

1. Bertalmio, M., Bertozzi, A., Sapiro, G.: Navier-Stokes, fluid dynamics, and image and video inpainting. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, pp. I355–I362 (2001)

2. Cai, J.-F., Chan, R.H., Shen, Z.: Simultaneous cartoon and texture inpainting. Inverse Probl. Imaging **4**, 379–395 (2010)
3. Davenport, M., Duarte, M., Eldar, Y., Kutyniok, G.: Introduction to compressed sensing. In: Compressed Sensing: Theory and Applications, pp. 1–68. Cambridge University Press, Cambridge (2012)
4. Donoho, D., Kutyniok, G.: Microlocal analysis of the geometric separation problem. Commun. Pure Appl. Math. **66**, 1–47 (2013)
5. Easley, G., Labate, D., Lim, W.: Sparse directional image representations using the discrete shearlet transform. Appl. Comput. Harmon. Anal. **25**, 25–46 (2008)
6. Easley, G.R., Labate, D., Colonna, F.: Shearlet based total variation for denoising. IEEE Trans. Image Process. **18**(2), 260–268 (2009)
7. Elad, M., Starck, J.-L., Querre, P., Donoho, D.L.: Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). Appl. Comput. Harmon. Anal. **19**, 340–358 (2005)
8. Foucart, S., Rauhut, H.: A Mathematical Introduction to Compressive Sensing. Birkhäuser, Basel (2013)
9. Genzel, M., Kutyniok, G.: Symptotic analysis of inpainting via universal shearlet systems. SIAM J. Imaging Sci. **7**, 2301–2339 (2014)
10. Guo, K., Labate, D.: Optimally sparse multidimensional representation using shearlets. SIAM J. Math. Anal. **39**, 298–318 (2007)
11. Guo, K., Labate, D.: The construction of smooth Parseval frames of shearlets. Math. Model. Nat. Phenom. **8**, 82–105 (2013)
12. Guo, K., Labate, D.: Geometric separation of singularities using combined multiscale dictionaries. J. Fourier Anal. Appl. **21**(4), 667–693 (2015)
13. Hennenfent, G., Fenelon, L., Herrmann, F.J.: Nonequispaced curvelet transform for seismic data reconstruction: a sparsity-promoting approach. Geophysics **75**(6), WB203–WB210 (2010)
14. Herrmann, F.J., Hennenfent, G.: Non-parametric seismic data recovery with curvelet frames. Geophys. J. Int. **173**(1), 233–248 (2008)
15. King, E.J., Kutyniok, G., Zhuang, X.: Analysis of data separation and recovery problems using clustered sparsity. In: Wavelets and Sparsity XIV. Proceedings of the SPIE, vol. 8138, p. 813818 (2011)
16. King, E.J., Kutyniok, G., Zhuang, X.: Analysis of inpainting via clustered sparsity and microlocal analysis. J. Math. Imaging Vis. **48**, 205–234 (2014)
17. Kutyniok, G.: Geometric separation by single-pass alternating thresholding. Appl. Comput. Harmon. Anal. **36**, 23–50 (2014)
18. Kutyniok, G., Lim, W.-Q., Reisenhofer, R.: Shearlab 3D: faithful digital shearlet transforms based on compactly supported shearlets. ACM Trans. Math. Softw. **42**(1), 5 (2015)
19. Labate, D., Lim, W.-Q., Kutyniok, G., Weiss, G.: Sparse multidimensional representation using shearlets. In: Papadakis, M., Laine, A.F., Unser, M.A. (eds.) Wavelets XI. SPIE Proc., vol. 5914, pp. 254–262. SPIE, Bellingham (2005)
20. Ma, J., März, M.: A multilevel based reweighting algorithm with joint regularizers for sparse recovery. Preprint
21. Nam, S., Davies, M.E., Elad, M., Gribonval, R.: Cosparse analysis modeling-uniqueness and algorithms. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 5804–5807. IEEE Press, New York (2011)
22. Nam, S., Davies, M.E., Elad, M., Gribonval, R.: The cosparse analysis model and algorithms. Appl. Comput. Harmon. Anal. **34**, 30–56 (2013)
23. Negi, P.S., Labate, D.: 3D discrete shearlet transform and video processing. IEEE Trans. Image Process. **21**(6), 2944–2954 (2012)