# Shearlet-based Analysis of Image Inpainting and Convolutional Framelets

by
José Pedro Rodríguez Ayllón

A dissertation submitted to the Department of Mathematics,

College of Natural Sciences and Mathematics

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Mathematics

Chair of Committee: Dr. Demetrio Labate

Committee Member: Dr. Emanuel Papadakis

Committee Member: Dr. Bernhard G. Bodmann

Committee Member: Dr. Thomas J. Hebert

University of Houston
August 2020

## DEDICATION/EPIGRAPH

*A mi madre Cecilia y a mi bebé Matías.*

José Pedro Rodríguez Ayllón

August, 2020

## ACKNOWLEDGMENTS

I am very grateful to my Advisor, Prof. Demetrio Labate, who guided me in this research work. I also wish to thank the members of my dissertation committee, Prof. Papadakis, Prof. Bodmann and Prof. Hebert for their observations and suggestions to improve the quality of my dissertation. And last but not least I want to acknowledge my family for their continuous support.

## ABSTRACT

The main part of my dissertation deals with image inpainting - a classical problem in image analysis that I analyze from the point of view of microlocal analysis and the theory of sparse approximation. My most important result provides a new set of theoretical performance guarantees for the exact recovery of missing data in images where the information is dominated by curvilinear singularities. In fact, my study shows that a shearlet-based approach for the recovery of missing curvilinear edges in images is provably superior to methods based on conventional wavelets in a precises sense and gives a quantitative assessment on the size of the region that can reliably recovered. As a consequence, this result offers the theoretical underpinning for algorithms based on directional multiscale methods such as shearlets in applications to image inpainting. The arguments in my proofs rely on a new application of the microlocal properties of shearlets and techniques from oscillatory integrals that are inspired in part by a seminal paper by Donoho and Kutyniok, who first introduced methods from microlocal analysis in combination with ideas from sparse representations for problems of image analysis. The second part of my dissertation is a new study of convolutional framelets - a method recently introduced to provide a mathematical framework for the patch-based analysis of images - using tools from tensor analysis. This method gives an alternative approach to analyze framelets and a deeper insight into the mathematical properties of convolutional framelets. The first part of the dissertation follows rather closely some material published by the author in his first journal paper.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# 1 Introduction

Today's world uses images extensively in many fields like science, engineering, medicine, architecture, art to name a few. Those fields rely on electronics to acquire, process, transmit and store those images. And due to many factors such as the inherent noise from electronic components, radio interference and defects in manufacture or malfunctions, images may get corrupted and some information is lost. So, there is the need to recover the missing part of the damaged image.

Human vision is remarkably able to fill these missing parts. Especially in art, conservators use a series of techniques called *inpainting* to restore and repair paintings. This word is now commonly used to describe a range of signal processing techniques for recovering missing blocks of data in digital images, video and audio. Typical examples of inpainting include the removal of overlaid text in images, repair of scratched photos and recovery of missing blocks in a streamed video such as those illustrated in Figure 1.

Many different ideas and methods have been proposed to deal with image inpainting and their performance may vary depending on the type of images considered as well as the geometry of the region to be recovered.

A very common approach to inpainting is to apply variational methods such in the work found in [1,3,4] which motivate the design of filling-in algorithms by geometric considerations. This leads to the use of partial differential equations (PDE) models which propagate information from the boundaries of the missing regions in an image while guaranteeing smoothness of some sort. The variational approach to inpainting has been shown to perform well on piecewise smooth images (e.g., those idealized images called *cartoon image*) and carry only geometric information. Yet, images also contains texture and variational methods do not perform well in such settings. On the other hand, local statistical analysis and prediction have been shown to perform well at filling texture content [16,23]. Additional works that use PDEs or variational principles to recover missing data from the close neighborhood of a region impose additional criteria of regularity to fill in holes [7,10,27,54].

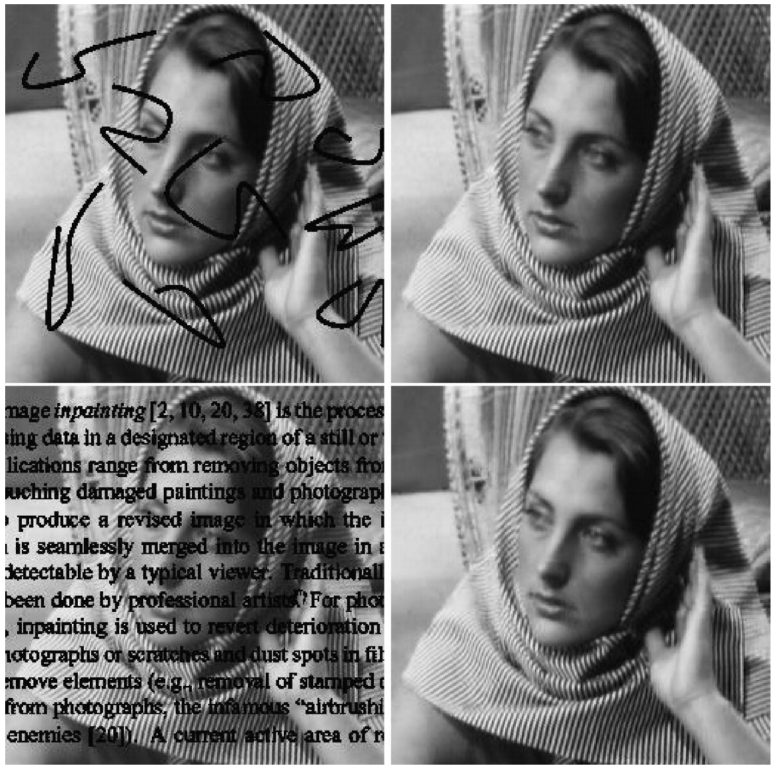


Figure 1: Examples of image inpainting. Left: corrupted images. Right: images restored using the inpainting algorithm in [26]

Since real images contain typically edges, smooth regions and texture, there is a major interest in developing techniques that can handle all these features. Besides that, approaches based on segmentation which label pixels as cartoon or texture should be avoided since some areas in the image contain contributions from both texture and cartoon. So, methods that decompose an image additively into layers may offer some advantages by combining layer-specific methods for filling in image holes, as done for instance in [5, 44, 60] where an image is separated into a cartoon and a texture component. Then the inpainting is done separately in each layer and the completed layers are superimposed to form the output image. This layer decomposition can also be based on variation methods leading to extend the notion of total-variation in [53].

Adapting the same multi-layer idea, Stark et al. introduced in [57, 58] a novel method of image decomposition that optimizes sparsity at each layer of the representation. The core idea is to use two adapted dictionaries, one to represent textures and the other to represent cartoons that also

also are mutually incoherent. That is, each dictionary is sparse for its target type of content but not for the other one. Building on this idea, Elad et al. introduced in [26] an inpainting method based on this sparse decomposition, called *morphological component analysis*, (MCA) that is capable to fill in holes in overlapping texture and cartoon image layers.

The idea of applying sparsity in combination with convex optimization has been also explored in other work, e.g., [6, 18]. These methods try to recover an image from highly incomplete linear measurements by $\ell_1$-minimization under the assumption that the image admits a sparse representation in a dictionary such as wavelets. Following this idea, several methods for inpainting have adopted representations such as wavelets, curvelets or shearlets to frame the inpainting problem as an optimization problem [6,18,26,30,56]. However, while these papers contain theoretical analyses of the convergence of their algorithms to the minimizers of specific optimization problems, they lack a theoretical analysis of how well those optimizers actually inpaint. Other results, especially in the engineering literature, examine the problem in the *discrete* setting and do not allow to take into account the geometry of the problem. By contrast, variational methods are built on continuous methods and may be analyzed using a continuous model like in [55]. By comparing inpainting methods performed through variational approaches with those built on $\ell_1$-minimization [7, 52] one can obtain useful insights. For instance, works such as [42, 43] provide an intuitive explanation of why directional representation systems such as curvelets ans shearlets can outperform wavelets when inpainting images governed by curvilinear singularities.

Perhaps the first paper attempting to formulate the inpainting problem according to a rigorous mathematical setting is the work by King et al. [46, 47] where inpainting is examined in the continuous domain as a function interpolation problem in a Hilbert space. Namely, the inpainting problem consists in recovering an unknown image $x$ in a Hilbert space $\mathcal{H}$ under the assumption that only a masked object $x_K = P_K x$ is known; here $P_K$ denotes the orthogonal projection into a known subspace $\mathcal{H}_K \subset \mathcal{H}$. To solve this problem, King et al. [47] propose an approach relying on microlocal analysis and sparse approximations based on methods originally introduced in [20]. Under the assumption that the unknown image $x$ is sparse with respect to a certain representation system $\Phi$, they search among all possible solutions $x^*$ such that $P_K x^* = x_K$ for the one that minimizes the $\ell^1$-norm of the representation coefficients of $x^*$ with respect to $\Phi$. Since images found in many applications are dominated by edges, it is reasonable to consider an image model consisting of distributions supported on curvilinear singularities. King et al. [47] proved that, if the missing information is a line segment, an $\ell^1$-norm minimization approach in combination with an appropriate function representation $\Phi$ is able to recover the missing information, asymptotically, provided the gap size is not too large. Remarkably, the theoretical performance of the recovery depends on the sparsifying and microlocal properties of the representation system $\Phi$, namely, asymptotically perfect recovery is achieved *if the gap size in the line singularity is asymptotically smaller than the size of the structure elements in* $\Phi$. In particular, it is proved that inpainting using the shearlet system – a multiscale anisotropic system that provides nearly optimally sparse representation of cartoon-like images [31, 48] – outperforms wavelets and similar conventional multiscale systems. A generalization using a more general shearlet system is given in [29].

The result by King et al. offers a rigorous theoretical assessment of the expected performance of a representation-based inpainting method. However, their approach makes a strong simplifying assumptions on the image model, namely, that the singularity to be inpainted is linear. This is a clearly major limitation since edges found in images are not necessarily linear.

One major contribution of this dissertation it to remove the image model restriction of King et al. [47] and consider more realistic images containing *general curvilinear singularities* while adopting the same continuous-domain formulation of the inpainting problem. Handling this more general type of singularities requires to develop several new technical tools and a significantly new proof.

While our arguments involve the same concept of clustered sparsity employed in [47] and originally introduced in [20], the fundamental technical elements of the proofs are novel, and rely critically on microlocal properties of shearlets and techniques from the analysis of oscillatory integrals associated with the continuous shearlet transform developed. Our main result generalizes and extends the result of King et al. to images containing curvilinear singularities where a section of the singularity curve is missing. Similar to [47], we consider two strategies for inpainting: one based on $\ell^1$ minimization and one based on thresholding. Using $\ell^1$ minimization in combination with a shearlet representation, our result recovers the same rate found by King et al. [47] in the case of linear singularities.

We conclude this introduction by recalling that, besides the above inpainting techniques described, deep neural networks have also been applied to image inpainting with promising results [11,17,61,63] following their success in other data restoration problems. In fact, more and more such methods are emerging by the day claiming state of the art results. Despite their remarkable performance though, such methods have some critical drawbacks. First, they rely on an extensive training procedure and their performance tend to degrade when the test images are moving away from the image type and format used during training. The other drawback is that deep neural networks are still difficult to interpret and currently provide no performance guarantees.

This rest of the dissertation is organized as follows. The reminder of this section gives notation and definitions. Section 2 introduces the mathematical model and gives the main results of this work. Section 3 gives technical tools to prove the results in Section 2. Section 4 proves the inpainting results using wavelets and Section 5 using shearlets. These first five sections follow closely some material in [39]. Section 6 gives the introduction to convolutional framelets which deals with the second part of this work. Section 7 introduces the tensor approach studied with these convolutional framelets. And finally, Section 8 tries to extend the notion of Hankel matrix to a tensor and presents related and future work to do.

## 1.1 Notation and basic definitions

In the following, we adopt the convention that $x \in \mathbb{R}^2$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^2$ (in the frequency domain) is a row vector, i.e., $\xi = (\xi_1, \xi_2)$. A vector $x$ multiplying a matrix $A \in GL_2(\mathbb{R})$ on the right is understood to be a column vector, while a vector $\xi$ multiplying $A$ on the left is a row vector. Thus, $Ax \in \mathbb{R}^2$ and $\xi A \in \widehat{\mathbb{R}}^2$.

Given two sequences $a = \{a_j\}_{j=1}^{\infty}$, $b = \{b_j\}_{j=1}^{\infty}$, we write $a \simeq b$ if there are constants $C_1 \neq$

0, $C_2 \neq 0$ such that $C_1\, b_j \leq a_j \leq C_2\, b_j$ for all large $j$. We write $a = O(b)$ is the limit $\lim_{j\to\infty} \frac{a_j}{b_j}$ exists and $a = o(b)$ is the limit $\lim_{j\to\infty} \frac{a_j}{b_j} = 0$.

From [28], we have

**Definition 1.1.** *The Fourier transform of $f \in L^1(\mathbb{R}^2)$ is defined as*

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

*where $\xi \in \widehat{\mathbb{R}}^2$, and the inverse Fourier transform is*

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

From [41], we have

**Definition 1.2.** *A set $E = \{e_\lambda : \lambda \in \Lambda\}$ in a Hilbert space $\mathcal{H}$ is a frame if there a constants $0 < A \leq B < \infty$ such that $A\,\|f\|^2 \leq \sum_{\lambda\in\Lambda} |\langle f, e_\lambda\rangle|^2 \leq B\,\|f\|^2$ for all $f \in \mathcal{H}$. A frame is tight if $A = B$ and is a Parseval frame if $A = B = 1$.*

Given a frame $E \subset \mathcal{H}$, the frame *synthesis operator* $F$ is the operator

$$F : \ell_2(I) \to \mathcal{H}, \quad F(\{c_\lambda\}_{\lambda\in\Lambda}) = \sum_{\lambda\in\Lambda} c_\lambda e_\lambda.$$

The dual operator of $F$, denoted by $F^*$, is the frame *analysis operator*

$$F^* : \mathcal{H} \to \ell_2(I), \quad F^* f = \{\langle f, e_\lambda\rangle : \lambda \in \Lambda\}.$$

We recall that if $E$ is a Parseval frame then, for any $f \in \mathcal{H}$,

$$FF^* f = \sum_{\lambda\in\Lambda} \langle f, e_\lambda\rangle\, e_\lambda = f.$$

For any measurable set $Q$ in $\mathbb{R}^2$ and any $f$ in $L^2(\mathbb{R}^2)$, we define $P_Q\, f$, the orthogonal projection of $f$ onto the set $Q$, that is,

$$P_Q f(x) = \mathbb{1}_Q(x) f(x) = \begin{cases} f(x) & \text{if } x \in Q \\ 0 & \text{if } x \notin Q \end{cases}.$$

Finally, we use the convention that same symbol $c$ or $C$ can be used denote a different generic constants in different expressions.

## 1.2 Multiscale representations: wavelets and shearlets

In this section, we introduce appropriate multiscale representations for the images we want to inpaint. Namely we consider (i) a Parseval frames of smooth bandlimited wavelets and (ii) a Parseval frames of smooth band-limited shearlets.

Images found in most applications are typically dominated by edges and other anisotropic structures. While wavelets have been very successful in signal processing applications, they have a geometric bias and are not very efficient to represent edges [51]. Curvelets [8] and shearlets [48, 49] were introduced precisely with the aim to overcome the limitations of conventional multiscale systems in the representation of edges. As we will show below, shearlets are a collection of functions defined not only over a range of locations and scale, like wavelets, but also over a range of orientations. Thanks to their increased directional sensitivity, shearlets are much more efficient than wavelets in representing images with edges [31, 33]. It was in fact shown that shearlets provide optimally sparse approximations, in a precise sense, in the class of *cartoon-like images*, a function class that was introduced to model idealize images with edges [22].

### 1.2.1 Wavelets

Meyer wavelets are some of the earliest known examples of orthonormal wavelets and they have high regularity [15]. We begin our construction with a smooth function $\varphi \in C^\infty$ such that its Fourier transform satisfies $0 \leq \widehat{\phi} \leq 1$ , $\widehat{\varphi} = 1$ on $[-\frac{1}{16}, \frac{1}{16}]$ and $\widehat{\varphi} = 0$ outside $[-\frac{1}{8}, \frac{1}{8}]$. Then for

$\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$ we define

$$\widehat{\Phi}(\xi) = \widehat{\Phi}(\xi_1, \xi_2) = \widehat{\phi}(\xi_1)\widehat{\phi}(\xi_2) \tag{1}$$

and

$$W(\xi) = W(\xi_1, \xi_2) = \sqrt{\widehat{\Phi}(2^{-2}\xi_1, 2^{-2}\xi_2)^2 - \widehat{\Phi}(\xi_1, \xi_2)^2}. \tag{2}$$

So, we have that $\operatorname{supp}(W) \subset [-\frac{1}{2}, \frac{1}{2}]^2 \setminus [-\frac{1}{16}, \frac{1}{16}]^2$ and the condition $\sum_{j \in \mathbb{Z}} |W(2^{-2j}\xi)|^2 = 1$, for a.e. $\xi \in \mathbb{R}^2$. We also define $W_j := W(2^{-2j}\cdot)$ with $j \in \mathbb{Z}$ and support inside the Cartesian coronae

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

that are illustrated in Figure 2. And for $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^2$, we define in the Fourier domain

$$\widehat{\phi}_{j,k}(\xi) = 2^{-2j}\, W(2^{-2j}\xi)\, e^{2\pi i 2^{-2j}\xi k}. \tag{4}$$

From [12,13,15,45], we have that $\overline{\Phi} = \{\phi_\lambda : \lambda \in \Lambda\} \subset L^2(\mathbb{R}^2)$ is a *Parseval frame of Meyer wavelets* where $\Lambda = \bigcup_{j \in \mathbb{Z}} \Lambda_j$, $\Lambda_j = \{\lambda = (j,k),\, k \in \mathbb{Z}^2\}$ and $\phi_\lambda = \phi_{j,k}$.

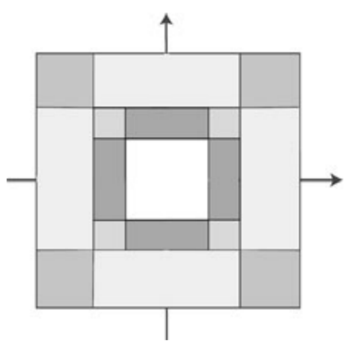

Figure 2: Frequency tiling of wavelets. Meyer wavelets are supported in the Cartesian coronae $Q_j$ in (3). The figure shows some additional partitions inside each corona which can be used to define an orthonormal wavelet system rather than a Parseval frame.

### 1.2.2 Shearlets

We now construct our *Parseval frame of shearlets* like in [34, 38] which is a modification of cone-adapted shearlets in [48]. We do this essentially by adding a directional refinement to the elements in (4). First, let us consider the following cone-shaped regions in the Fourier domain $\widehat{\mathbb{R}}^2$

$$\mathcal{C}_1 = \left\{ (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\frac{\xi_2}{\xi_1}| \leq 1 \right\}, \ \mathcal{C}_2 = \left\{ (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\frac{\xi_2}{\xi_1}| > 1 \right\},$$

and let $V \in C_0^\infty(\mathbb{R})$ be chosen so that $\operatorname{supp} V \subset [-1, 1]$ and

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

Let $G_{(1)}(\xi_1, \xi_2) = V(\frac{\xi_2}{\xi_1})$ and $G_{(2)}(\xi_1, \xi_2) = V(\frac{\xi_1}{\xi_2})$, and let $W \in C_0^\infty(\mathbb{R}^2)$ be the same window function as in (2). Then, our *shearlet system* for $L^2(\mathbb{R}^2)$ is given by

$$\Psi = \{\psi_{-1,k} : k \in \mathbb{Z}^2\} \cup \{\psi_{j,\ell,k}^{(\nu)} : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^2, \nu = 1, 2\} \cup \{\psi_{j,\ell,k} : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^2\} \tag{5}$$

consisting of:

- *coarse-scale shearlets* $\{\psi_{-1,k} : k \in \mathbb{Z}^2\}$ given by $\psi_{-1,k} = \widehat{\Phi}(\cdot - k)$ where $\widehat{\Phi}$ is given by (1).
- *interior shearlets* $\{\psi_{j,\ell,k}^{(\nu)} : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^2, \nu = 1, 2\}$ given by

$$\hat{\psi}_{j,\ell,k}^{(\nu)}(\xi) = |\det A_{(\nu)}|^{-j/2} \, W(2^{-2j}\xi) \, G_{(\nu)}(\xi A_{(\nu)}^{-j} B_{(\nu)}^{-\ell}) \, e^{2\pi i \xi A_{(\nu)}^{-j} B_{(\nu)}^{-\ell} k}, \quad \xi \in \mathcal{C}_\nu \tag{6}$$

- and *boundary shearlets* $\{\psi_{j,\ell,k} : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^2\}$ given by

$$\widehat{\psi}_{j,\ell,k}(\xi) = \begin{cases} 2^{-\frac{3}{2}j - \frac{1}{2}} W(2^{-2j}\xi) V\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \xi \in \mathcal{C}_1 \\ 2^{-\frac{3}{2}j - \frac{1}{2}} W(2^{-2j}\xi) V\left(2^j \frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \xi \in \mathcal{C}_2 \end{cases}$$

where

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

From [34], we see that (5) is a *Parseval frame of shearlets* for $L^2(\mathbb{R}^2)$ whose elements are $C^\infty$ and band-limited.

From now on, we will write (5) as $\Psi = \{\psi_\eta : \eta \in M\}$, and the index set $M$ is expressed as $M = M_C \cup M_F$, where $M_C = \{k \in \mathbb{Z}^2\}$ are the index set associated with coarse-scale shearlets and $M_F = \{\eta = (j, \ell, k, \nu) : j \geq 0, |\ell| \leq 2^j, k \in \mathbb{Z}^2, \nu = 1, 2\}$ is the set associated with fine-scale shearlets. See [34] for additional details about this construction.

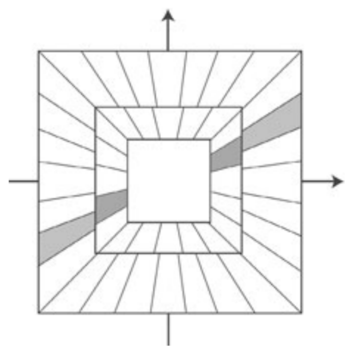

Figure 3: Frequency tiling of shearlet system

We recall here that shearlets offer nearly optimally sparse approximations properties, in a precise sense, for the class of cartoon-like images – an idealized model of images with edges [31, 33].

In Figure 3, we can see that the shearlet decomposition provides a finer partition of the Fourier domain that is associated with directional sub-bands as opposed to wavelets in Figure 2.

Another remarkable property is that the continuous shearlet transform associated with the shearlet representation provides a precise characterization of curvilinear singularities due to its microlocal properties [32, 36, 37, 50]. These properties of shearlets underpin several results derived in this work.

## 2 Main results

Here we state the main results of this dissertation after discussing the mathematical model of cartoon-like images and the inpainting algorithms.

### 2.1 Mathematical model of inpainting

A simplified model of natural images are cartoon-like images, which emphasizes anisotropic features, most notably edges. Since these images basically consists of smooth regions separated by edges, it is suggestive to use a model consisting of piecewise regular functions, such as the one illustrated in Figure 4.

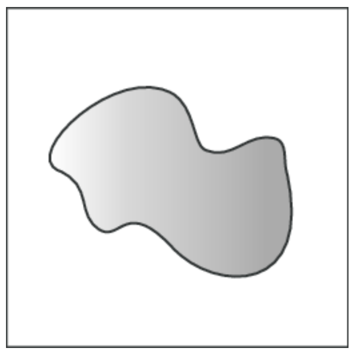

Figure 4: Example of a cartoon-like image (function values represented using a gray scale map) extracted from [48]

We follow a continuous image model like in [47] and we first introduce the mathematical model of the image we want to inpaint. In [48], a *cartoon-like image* is defined as a function $f : \mathbb{R}^2 \to \mathbb{C}$ of the form $f = f_0 + \mathcal{X}_B f_1$ where $B \subset [0, 1]^2$ and $\partial B$ is a closed $C^2$ curve with bounded curvature and $f_i \in C^2(\mathbb{R}^2)$ are functions with support $\operatorname{supp}(f_i) \in [0, 1]^2$ and $\|f_i\|_{C^2} \leq 1$ for $i = 1, 2$.

The simplest **edge model** is the step function or Heaviside distribution $H(x) = \mathbb{1}(x)$ in $\mathbb{R}$ where the singularity is at 0. Since $\widehat{H}(\xi) = \delta(\xi) + \frac{1}{\pi i} p.v. \left(\frac{1}{\xi}\right)$ where $\delta$ is the Dirac distribution and $p.v.$ is the principal value distribution, then for any test function $\phi$

$$\langle H, \phi \rangle = \langle \widehat{H}, \widehat{\phi} \rangle = \int_{\widehat{\mathbb{R}}} \delta(\xi)\widehat{\phi}(\xi) d\xi + \frac{1}{\pi i} \int_{\widehat{\mathbb{R}}} \frac{\widehat{\phi}(\xi)}{\xi} d\xi.$$

Thus, if we parametrize $\partial B$ with a $C^2$ curve $\tau : [a, b] \to \mathbb{R}^2$ , then the function $\mathcal{X}_B$ acting on a function $\phi$ in $\mathbb{R}^2$ would be

$$\int_{\mathbb{R}^2} \delta_\tau(x)\phi(x)dx = \int_{\partial B} \phi(\tau)d\tau$$

So, we can model a cartoon like-image as a distribution like in [20]. We now formalize this concept. From [28] we have

**Definition 2.1.** *The Schwartz space is* $\mathcal{S}(\mathbb{R}^n) = \{f \in \mathbb{C}^\infty(\mathbb{R}^n) : \|f\|_{(N,\alpha)} < \infty, N \in \mathbb{N}, \alpha\}$ *where* $\alpha = (\alpha_1, \ldots, \alpha_r)$ *is a multi-index,* $\|f\|_{(N,\alpha)} = \sup_{x\in\mathbb{R}^n}(1 + |x|)^N |\partial^\alpha f(x)|$ *and* $\partial^\alpha f(x) = \partial_{\alpha_1} \cdots \partial_{\alpha_r} f(x)$.

So, the Schwartz space consists of those $C^\infty$ functions, which together with all their derivatives, vanish at infinity faster for any power of $|x|$.

**Definition 2.2.** *A distribution on* $U \subset \mathbb{R}^n$ *is a continuous linear functional on* $C_c^\infty(U)$. *The space of all distributions on* $U$ *is denoted by* $\mathcal{D}'(U)$ *with the topology of point-wise convergence on* $C_c^\infty(U)$. *A tempered distribution is a distribution defined on* $\mathcal{S}$. *We denote by* $\mathcal{S}'$ *the space of tempered distributions.*

Let $S$ be a closed smooth curve contained in $[-1, 1]^2 \subset \mathbb{R}^2$ that has nonvanishing curvature everywhere. We define a tempered distribution $\mathcal{T} \in \mathcal{S}'(\mathbb{R}^2)$ acting on the class of Schwartz functions $\phi \in \mathcal{S}(\mathbb{R}^2)$ and supported on $S$ by

$$\langle \mathcal{T}, \phi\rangle = \int_S \phi(s)\, g(s)\, d\sigma(s)$$

where $g$ is a real-valued smooth function defined on the curve $S$. $\mathcal{T}$ is the model of a cartoon-like image we use in this work.

We want to recover a missing portion of this curve model $\mathcal{T}$. We model this missing portion using the concept of a *mask*. For $h > 0$, we denote as $\mathcal{M}_h$ the *horizontal strip domain*

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

This horizontal stripe is the mask, and correspondingly, we consider the masked function

$$f = P_{\mathbb{R}^2 \setminus \mathcal{M}_h} \mathcal{T}.$$

This is the model of the image we wish to inpaint. Notice we could also assume that the region to be inpainted is contained in a vertical strip domain of width $h$ and our proofs below would be very similar.

## 2.2 Recovery algorithms

Our argument follows an approach similar to [47], namely we use $\ell_1$-minimization and thresholding algorithms for inpainting as they provide an efficient theoretical and computational framework to take advantage of the sparsity properties of the wavelet and shearlet decomposition.

Our approach carries out an asymptotic scale-dependent analysis. So, we start by decomposing the image into subbands in the Fourier domain and we achieve this by projecting $\mathcal{T}$ into the Fourier regions associated with the Cartesian coronae $Q_j$, $j \in \mathbb{Z}$, given by (3). Since $\mathcal{T}$ is supported in the curve $S$, we have that $\widehat{\mathcal{T}}(\xi) = \langle \mathcal{T}, e^{-2\pi\xi\cdot x} \rangle$ is a $C^\infty$ function, see [28]. So for $j \in \mathbb{Z}$, we let $\mathcal{T}_j \in L^2(\mathbb{R}^2) \subset \mathcal{S}'(\mathbb{R}^2)$ be characterized by

$$\widehat{\mathcal{T}_j}(\xi) = \widehat{\mathcal{T}}(\xi) W(2^{-2j}\xi).$$

Notice that $\widehat{\mathcal{T}_j}$ is smooth and compactly supported since $W(2^{-2j}\cdot)$ is band-pass filter that appears in (4). Correspondingly, we have a sequence of masked images

$$f_j = P_{\mathbb{R}^2 \setminus \mathcal{M}_{h_j}} \mathcal{T}_j \in L^2(\mathbb{R}^2), \tag{7}$$

where now $h_j$ depends on the scale parameter $j$.

#### 2.2.1 Inpainting via $\ell_1$-minimization

The $\ell_1$ minimization process to recover an approximate solution has the form

$$R_j^{\ell} = \text{argmin}_{T_j} \|F^* \mathcal{T}_j\|_1 \quad \text{subject to } f_j = P_{\mathbb{R}^2 \setminus \mathcal{M}_{h_j}} \mathcal{T}_j,$$

where $F$ is the frame operator associated with a Parseval frame of wavelets or shearlets.

Notice that the norm is placed on the *analysis* coefficients rather than the *synthesis* coefficients as in [21, 25]. This is because we are using a frame operator $F$ which is not necessarily associated with a basis. If we discretize the problem, thne we have $F \in \mathbb{R}^{N\times L}$ where $L >> N$, $F$ has full rank and $FF^* = I_N$. To solve for the synthesis coefficients $x = Fc$, we need to minimize $f(c) = \|x - Fc\|_2^2 + \lambda\|c\|_1$ where $\lambda > 0$. So, the Hessian matrix of $f$ contains a term of the form $F^*F \in \mathbb{R}^{L\times L}$ which is not of full rank. In addition, $F^*F$ is positive semidefinite. Whereas, to solve for the analysis coefficients $c = F^*x$, we have to minimize $g(x) = \|c - F^*x\|_2^2 + \lambda\|F^*x\|_1$ and its Hessian matrix contains a term of the form $FF^* = I_N$ which is positive definite. That is, $g$ has a unique global minimum. Therefore, the solution is unique [2]. Several inpainting algorithms are based on the idea of $\ell_1$-minimization and have been shown heuristically to be successful [6, 18, 26].

**Remark 2.3.** *When we apply $\ell_1$ minimization with shearlets, for a technical reason, we slightly modify the setting in [47] by choosing the missing part to be $P_{\mathcal{M}_{h_j}} f(x) = h_j^{\Delta_0} \mathbb{1}_{|x_2|\le h_j} f(x)$, for some fixed small $\Delta_0 > 0$. For $\ell_1$ minimization with wavelets and thresholding with both wavelets and shearlets, we follow [47] to choose the missing part to be $P_{\mathcal{M}_{h_j}} f(x) = \mathbb{1}_{|x_2|\le h_j} f(x)$.*

#### 2.2.2 Inpainting via thresholding

For the thresholding strategy, given a Parseval frame of wavelets or shearlets $E = \{e_\lambda\}_{\lambda\in\Lambda}$ and a sequence of thresholds $\sigma_j$, $j \in \mathbb{Z}$, we let $I_j = \{\lambda \in \Lambda : |\langle f, e_\lambda\rangle| \ge \sigma_j\}$. Then the reconstructed image with respect to $E$ is

$$R_j^{\tau} = F\,(\mathbb{1}_{I_j} F^* \mathcal{T}_j).$$

## 2.3 Asymptotic analysis

The width of the area to be inpainted is a useful indication of the power of the inpainting method. For instance, Chan and Kan [9] proposed a variational inpainting methods and demonstrated that the local thickness of the region to be inpainted affects the success of the inpainting method more than the overall size of the area to be inpainted.

In our approach, since we apply a multiscale analysis framework, we examine the impact of the width of the area to be inpainted by making the gap size $h$ dependent on the scale $j$, as shown in equation (7). As we indicated above, we consider a strategy based on $\ell^1$ minimization and another one based on thresholding to recover $\mathcal{T}_j$, $j \in \mathbb{Z}$. In both cases, we establish a procedure to construct an approximate solution $R_j \in L^2(\mathbb{R}^2)$ and show that

$$\frac{\|R_j - \mathcal{T}_j\|_2}{\|\mathcal{T}_j\|_2} \to 0, \text{ as } j \to \infty$$

provided $h_j = o(2^{-\alpha j})$ for an appropriate $\alpha > 0$. That is, we show the relative $L^2$-error of the reconstructed signal $R_j$ to the band-passed image $\mathcal{T}_j$ approaches to 0 as $j$ increases. We remark that, since $R_j - \mathcal{T}_j \in L^2(\mathbb{R}^2)$, for any $\phi$ with compact support, it follows that $|\langle R_j - \mathcal{T}_j, \phi\rangle| \leq \|R_j - \mathcal{T}_j\|_2 \|\phi\|_2$. Thus, $\frac{1}{\|\mathcal{T}\|_2}(R_j - \mathcal{T}_j) \to 0$ in $\mathcal{D}'$.

We will prove that if the reconstruction approach is based on shearlets then the parameter $\alpha$ controlling the gap size can be taken significantly smaller than in an analogous scheme based on wavelets. That is, shearlets can asymptotically recover missing data with a significantly larger spatial support. As we remarked above, it was proved [47] that the wavelet estimate cannot be improved in the thresholding case, implying that the shearlet result is provably superior to the wavelet case. It is conjectured that the same holds in the $\ell^1$ case but there is no proof at this time that the wavelet estimate cannot be improved.

**Remark 2.4.** *In this work, we applied parabolic scaling for the shearlet system (that is the parameters in the dilation matrix $A$ are scaled parabolically). A different scaling approach is used* *in [29] where a universal shearlet system is defined, using scaling matrices* $A_{\alpha_j,(h)} = \begin{pmatrix} 4 & 0 \\ 0 & 2^{\alpha_j} \end{pmatrix}$, $A_{\alpha_j,(v)} = \begin{pmatrix} 2^{\alpha_j} & 0 \\ 0 & 4 \end{pmatrix}$ *where* $(\alpha_j) \subset (-\infty, 2)$ *is called a scaling sequence. We remark that the analysis carried out in [29] only consider the case of inpainting a linear singularity, as in [47], and concludes that, for the* $\ell_1$*-minimization,* $\frac{\|R_j - \mathcal{T}_j\|_1}{\|\mathcal{T}_j\|_1}$ *decays as* $o(2^{-Nj})$ *provided that* $h_j$ *is of order* $o(2^{-\alpha_j j})$.

## 2.4 Main results of inpainting analysis

Following our previous analysis, our main results are the following theorems which we wish to prove.

**Theorem 2.5.** *Let* $\Phi$ *be a Parseval frame of wavelets on* $L^2(\mathbb{R}^2)$ *as defined in Section 1.2 and let* $R_j^\ell$ *be the reconstructed image of* $\mathcal{T}_j$ *obtained via* $\ell_1$ *minimization where we assume that* $h_j = o(2^{-2j})$. *Then*

$$\frac{\|R_j^\ell - \mathcal{T}_j\|_2}{\|\mathcal{T}_j\|_2} \to 0, \ \text{as } j \to \infty.$$

**Theorem 2.6.** *Let* $\Phi$ *be a Parseval frame of wavelets on* $L^2(\mathbb{R}^2)$ *as defined in Section 1.2 and let* $R_j^\tau$ *be the reconstructed image of* $\mathcal{T}_j$ *obtained via thresholding where we assume that* $0 \leq \sigma_j \leq 2^{-4j}$ *and* $h_j = o(2^{-j})$. *Then*

$$\frac{\|R_j^\tau - \mathcal{T}_j\|_2}{\|\mathcal{T}_j\|_2} \to 0, \ \text{as } \to \infty.$$

**Theorem 2.7.** *Let* $\Psi$ *be a Parseval frame of shearlets on* $L^2(\mathbb{R}^2)$ *as defined in Section 1.2 and let* $R_j^\ell$ *be the reconstructed image of* $\mathcal{T}_j$ *obtained via* $\ell_1$ *minimization where we assume that* $h_j = o(2^{-j})$. *Then*

$$\frac{\|R_j^\ell - \mathcal{T}_j\|_2}{\|\mathcal{T}_j\|_2} \to 0, \ \text{as } j \to \infty.$$

**Theorem 2.8.** *Let* $\Psi$ *be a Parseval frame of shearlets on* $L^2(\mathbb{R}^2)$ *as defined in Section 1.2 and let* $R_j^\tau$ *be the reconstructed image of* $\mathcal{T}_j$ *obtained via thresholding where we assume that* $0 \leq \sigma_j \leq 2^{-4j}$

*and* $h_j = o(2^{-\frac{3}{4}j})$. *Then*

$$\frac{\|R_j^\tau - \mathcal{T}_j\|_2}{\|\mathcal{T}_j\|_2} \to 0, \ \text{as } j \to \infty.$$

We remark that our estimates for the $\ell_1$ minimization case (Theorems 2.5 and 2.7) extend the results from King et al. [47] to the more challenging setting where the missing information is curvilinear. As we show below our proof uses a very different approach from the one in the original paper and it relies in part on properties of the shearlet representation explored in [38] and [37].

In the thresholding case (Theorems 2.6 and 2.8), our estimates improve those found by King et al. [47] indicating a better inpainting performance (i.e., the size of the missing gap can be larger) than $\ell_1$ minimization for both wavelets and shearlets. Furthermore, proofs of Theorems 2.6 and 2.8 do not require the assumption of nonvanishing curvature. Hence our result includes the situation where the missing region is a line segment and, thus, improves the result in [47].

Our estimates show that the size of the gap that can be filled by shearlets with asymptotically high precision is larger than the corresponding one for wavelets. King et al. [47] prove that, in the thresholding case, the wavelet rate cannot be improved for linear gaps. This shows that shearlets perform better than wavelets. There is currently no proof of a similar negative wavelet result in the $\ell_1$ case.

## 3 Preparation: Useful technical results

Here we introduce some constructions that will be needed for the proofs of our main results.

We consider a smooth curve $S \subset [-1,1]^2$. Using a smooth partition of unity, we can decompose $S$ as $S = \bigcup_1^M S_m$ where each $S_m$ has non vanishing curvature (the case where $S_m$ is a straight line was already considered in [47]). Each $S_m$ can be parametrized either as vertical curve $(f(u), u)$ or horizontal curve $(u, f(u))$ where $u \in (a_m, b_m)$ and $f \in C^\infty(a_m, b_m)$, $m = 1, \ldots, M$. In either case we assume there is a constant $k > 0$ such that $|f''(u)| \geq k > 0$ for all $u \in [a_m, b_m]$. We also assume a *vertical* curve is defined if the slope of the tangent line to the curve is greater or equal than 2 so $|f'(u)| \leq 1/2$. And similarly, a *horizontal* curve is defined if the slope to its tangent line is less than 2 so that $|f'(u)| \leq 2$. With this assumptions, the function $y = \frac{1}{2}x^2$, $x \in (-1,1)$ is a horizontal curve. Whereas $y^2 = 8x$, $y \in (-1,1)$, is a vertical curve and it may be written as $(\frac{1}{8}u^2, u)$ for $u \in (-1,1)$.

For each curve $S_m$, $1 \leq m \leq M$, there is a smooth density function $g_m \in \mathbb{C}_0^\infty(S_m)$. So, for any $\phi \in \mathcal{S}(\mathbb{R}^2)$,

$$\langle \mathcal{T}, \phi \rangle = \int_S \phi(s)\, g(s)\, d\sigma(s) = \sum_{m=1}^M \int_{S_m} \phi(s)\, g_m(s)\, d\sigma(s) = \sum_{m=1}^M \langle \mathcal{T}_m, \phi \rangle,$$

where, for each $m$, $\mathcal{T}_m$ is a distribution defined either by

$$\langle \mathcal{T}_m, \phi \rangle = \int_{a_m}^{b_m} \phi(f(u), u)\, g_m(u)\, du \quad \text{if } S_m \text{ is a vertical curve}$$

or by

$$\langle \mathcal{T}_m, \phi \rangle = \int_{a_m}^{b_m} \phi(u, f(u))\, g_m(u)\, du \quad \text{if } S_m \text{ is a horizontal curve.}$$

To be consistent with the above notation, we define $\mathcal{T}_{m,j}$ in the Fourier domain as $\widehat{\mathcal{T}}_{m,j}(\xi) = W(2^{-2j}\xi)\widehat{\mathcal{T}}_m(\xi)$. Hence $\widehat{\mathcal{T}}_j(\xi) = \sum_{m=1}^M \widehat{\mathcal{T}}_{m,j}(\xi)$. It is also helpful to use polar coordinates and we make a change of variable. That is, for any $\xi = (\xi_1, \xi_2)$, we write $\xi = \rho\Theta(\theta)$ where $\rho = |\xi| = \sqrt{\xi_1^2 + \xi_2^2}$ and $\Theta(\theta) = (\cos(\theta), \sin(\theta))$ where $\Theta(0) = (1,0)$ because by convention, the angle at origin is zero. Thus, for a vertical curve $S_m$ we may write in polar coordinates as

$$\widehat{\mathcal{T}}_{m,j}(\rho,\theta) = W(2^{-2j}\rho\Theta(\theta)) \int_{a_m}^{b_m} e^{-2\pi i\rho\Theta(\theta)\cdot(f(u),u)} g_m(u)\,du. \tag{8}$$

Similarly for a horizontal curve we have

$$\widehat{\mathcal{T}}_{m,j}(\rho,\theta) = W(2^{-2j}\rho\Theta(\theta)) \int_{a_m}^{b_m} e^{-2\pi i\rho\Theta(\theta)\cdot(u,f(u))} g_m(u)\,du.$$

Now we establish some useful technical results which will be needed to prove our main results.

**Lemma 3.1.** *Assume that the local curve $S_m$ is vertical and let $\beta^{(2)}_{j,\ell,k} = \langle \psi^{(2)}_{j,\ell,k}, \mathcal{T}_{m,j}\rangle$, where $\mathcal{T}_{m,j}$ is given above and $\psi^{(2)}_{j,\ell,k}$ is given by (6). Then, for any $N \in \mathbb{N}$, there exists a constant $C_N$, independent of $j, \ell, k$ such that $|\beta^{(2)}_{j,\ell,k}| \leq C_N\, 2^{\frac{5}{2}j}\, 2^{-2Nj}$.*

*Proof.*

$$\beta^{(2)}_{j,\ell,k} = \langle \widehat{\psi}^{(2)}_{j,\ell,k}, \widehat{\mathcal{T}}_{m,j}\rangle = \int_{\widehat{\mathbb{R}}^2} \widehat{\psi^{(2)}_{j,\ell,k}}(\xi)\overline{\widehat{\mathcal{T}}_{m,j}}(\xi)d\xi.$$

Let $\xi = (\xi_1,\xi_2) = (\rho\cos(\theta), \rho\sin(\theta)) = \rho\Theta(\theta)$. From (6) we get:

$$\widehat{\psi^{(2)}_{j,\ell,k}}(\xi) \;=\; 2^{-3j/2} W(2^{-2j}\rho\Theta(\theta)) V(2^j\cot(\theta)-\ell) e^{2\pi i\rho\Theta(\theta)A_{(2)}^{-j}B_{(2)}^{-\ell}k}.$$

Notice $\operatorname{supp}(W) \subset ([-\frac{1}{2},\frac{1}{2}] - [-\frac{1}{16},\frac{1}{16}])^2$. Thus, $\frac{1}{16}2^{2j} \leq \rho \leq \frac{1}{\sqrt{2}}2^{2j}$. Also, $\operatorname{supp}(V) \subset [-1,1]$, so $|2^j\cot(\theta)-\ell| \leq 1$, thus $|\cot(\theta)| \leq 2^{-j}(1+|\ell|) \leq 1+2^{-j}$. Hence $|\theta-\pi/2| \leq \pi/4+\epsilon_j$ or $|\theta-3\pi/2| \leq \pi/4+\epsilon_j$ where $\epsilon_j \to 0$ as $j\to\infty$. Also, let $\phi(u) = \Theta(\theta)\cdot(f(u),u) = \cos(\theta)f(u)+\sin(\theta)u$.

For $\theta \in [\pi/4-\epsilon_j, 3\pi/4+\epsilon_j] \cup [5\pi/4-\epsilon_j, 7\pi/4+\epsilon_j]$ we have $|\sin(\theta)| \geq \frac{1}{2} \geq \frac{1}{2}|\cos(\theta)|$. So, there is $c>0$ such that $c \leq |\sin(\theta)| - \frac{1}{2}|\cos(\theta)| \leq |\cos(\theta)f'(u)+\sin(\theta)| = |\phi'(u)|$ since $|f'(u)| \leq \frac{1}{2}$.

$$\begin{aligned}
\beta^{(2)}_{j,\ell,k} &= \int_{\widehat{\mathbb{R}}^2} \widehat{\psi^{(2)}_{j,\ell,k}}(\xi)\overline{\widehat{\mathcal{T}_{m,j}}}(\xi)d\xi \\
&= \int_{\frac{1}{16}2^{2j}}^{\frac{1}{\sqrt{2}}2^{2j}} \left[\int_{\pi/4-\epsilon_j}^{3\pi/4+\epsilon_j} + \int_{5\pi/4-\epsilon_j}^{7\pi/4+\epsilon_j}\right] 2^{-3j/2}W(2^{-2j}\rho\Theta(\theta))V(2^j\cot(\theta)-\ell)
\end{aligned}$$

$$\times \quad e^{(2\pi i\rho\Theta(\theta)A_{(2)}^{-j}B_{(2)}^{-\ell}k)}\Big(\overline{W}(2^{-2j}\rho\Theta(\theta))\int_a^b e^{(2\pi i\rho\Theta(\theta)\cdot(f(u),u))}g_m(u)du\Big)\rho d\theta d\rho$$

$$= \quad 2^{5j/2}\int_{\frac{1}{16}2^{2j}}^{\frac{1}{\sqrt{2}}2^{2j}}\Big[\int_{\pi/4-\epsilon_j}^{3\pi/4+\epsilon_j}+\int_{5\pi/4-\epsilon_j}^{7\pi/4+\epsilon_j}\Big]\big|W(\rho\Theta(\theta))\big|^2V(2^j\cot(\theta)-\ell)$$

$$\times \quad e^{(2\pi i2^{2j}\rho\Theta(\theta)A_{(2)}^{-j}B_{(2)}^{-\ell}k)}\Big(\int_a^b e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}g_m(u)du\Big)|\rho|d\theta d\rho.$$

Integrating by parts with respect to the variable $u$ we get:

$$\begin{aligned}
&\int_a^b e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}g_m(u)du\\
&=\int_a^b \frac{g_m(u)}{2\pi i2^{2j}\rho\phi'(u)}e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}(2\pi i2^{2j}\rho\phi'(u))du\\
&=\frac{g_m(u)}{2\pi i2^{2j}\rho\phi'(u)}e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}\Big|_a^b-\int_a^b\frac{d}{du}\Big(\frac{g_m(u)}{2\pi i2^{2j}\rho\phi'(u)}\Big)e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}du\\
&=\frac{1}{-2\pi i2^{2j}\rho}\int_a^b g_{m,1}(u)e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}du
\end{aligned}$$

since $g_m$ is compactly supported on $(a,b)$. Also, notice $0<c\le|\phi'(u)|$. Repeating the above process $N$ times we get

$$\int_a^b e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}g_m(u)du=\frac{2^{-2Nj}}{(-2\pi i\rho)^N}\int_a^b g_{m,N}(u)e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}du.$$

Hence

$$\begin{aligned}
|\beta^{(2)}_{j,\ell,k}| = \quad &\Big|2^{5j/2}\int_{\frac{1}{16}2^{2j}}^{\frac{1}{\sqrt{2}}2^{2j}}\Big[\int_{\pi/4-\epsilon_j}^{3\pi/4+\epsilon_j}+\int_{5\pi/4-\epsilon_j}^{7\pi/4+\epsilon_j}\Big]\big(W(\rho\Theta(\theta))\big)^2V(2^j\cot(\theta)-\ell)\\
\times \quad & e^{(2\pi i2^{2j}\rho\Theta(\theta)A_{(2)}^{-j}B_{(2)}^{-\ell}k)}\Big(\int_a^b e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}g_m(u)du\Big)\rho d\theta d\rho\Big|\\
\le \quad & 2^{5j/2}\int_{\frac{1}{16}2^{2j}}^{\frac{1}{\sqrt{2}}2^{2j}}\Big[\int_{\pi/4-\epsilon_j}^{3\pi/4+\epsilon_j}+\int_{5\pi/4-\epsilon_j}^{7\pi/4+\epsilon_j}\Big]\big|W(\rho\Theta(\theta))\big|^2\big|V(2^j\cot(\theta)-\ell)\big|\\
\times \quad & \Big|\int_a^b e^{\left[2\pi i2^{2j}\rho\phi(u)\right]}g_m(u)du\Big|\rho d\theta d\rho\\
\le \quad & 2^{5j/2}\int_{\frac{1}{16}2^{2j}}^{\frac{1}{\sqrt{2}}2^{2j}}\Big[\int_{\pi/4-\epsilon_j}^{3\pi/4+\epsilon_j}+\int_{5\pi/4-\epsilon_j}^{7\pi/4+\epsilon_j}\Big]\big|W(\rho\Theta(\theta))\big|^2\big|V(2^j\cot(\theta)-\ell)\big|
\end{aligned}$$

$$\times \quad \frac{2^{-2Nj}}{(2\pi\rho)^N} \int_a^b \left|g_{m,N}(u)\right| du \, \rho d\theta d\rho \leq 2^{5j/2} C_N 2^{-2Nj}. \quad \square$$

The following lemma is a special case of the classical method of stationary phase (cf. Proposition 3 in [59, Chapter VIII]).

**Lemma 3.2.** *Let $\varphi$ and $\psi$ be smooth functions. Suppose $\varphi'(u_0) = 0$ and $\varphi''(u_0) \neq 0$. If $\psi$ is supported in a sufficiently small neighborhood of $u_0$, then*

$$J(\lambda) = \int_{\mathbb{R}} e^{i\,\lambda\,\varphi(u)}\,\psi(u)\,du = \lambda^{-1/2}\,e^{i\,\lambda\,\varphi(u_0)}\left(a(u_0) + O(\lambda^{-\frac{1}{2}})\right),$$

*as $\lambda \to \infty$, where $a(u_0) = \left(\frac{2\pi i}{\varphi''(u_0)}\right)^{\frac{1}{2}} \psi(u_0)$.*

We remark that, in the following, we will apply Lemma 3.2 for estimates where $a(u_0)$ appears in absolute value. Thus, in the statement above it is irrelevant the choice of a particular square root.

We will also need the following classical lemma, known as Van der Corput Lemma, from the theory of oscillatory integrals (cf. Proposition 2 and its corollary in [59, Chapter VIII]) that has been a key tool in harmonic analysis for finding decay rates of solutions to differential equations (e.g., Bessel functions).

**Lemma 3.3.** *Let $k \geq 2$, $\lambda > 0$, and $\phi(x)$ be a real-valued function defined on $[a, b]$ such that $|\phi^{(k)}(x)| \geq 1$ for all $x \in [a, b]$. Also, let $\psi$ be smooth and compactly supported in $[a, b]$. Then*

$$\left|\int_a^b e^{i\lambda\phi(x)}\psi(x)dx\right| \leq C_k \lambda^{-\frac{1}{k}}\left(|\psi(b)| + \int_a^b |\psi'(x)|dx\right),$$

*where $C_k$ only depends only on $k$.*

Finally, we will need the following observation whose proof relies on Lemmata 3.2 and 3.3.

**Lemma 3.4.** *With the notation introduced above, for any $j \in \mathbb{Z}$, we have*

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

*Proof.* If $\|\mathcal{T}_{m,j}\|_2 \simeq 2^j$ for each $m = 1, \ldots, N$. Then using the decomposition $\mathcal{T}_j = \sum_{m=1}^{M} \mathcal{T}_{m,j}$, we have $\|\mathcal{T}_j\|_2 \leq \sum_{m=1}^{M} \|\mathcal{T}_{m,j}\|_2$. Now for $M = 2$, we have $\|\mathcal{T}_j\|_2 \geq \big|\|\mathcal{T}_{1,j}\|_2 - \|\mathcal{T}_{2,j}\|_2\big|$. Therefore, we have $\|\mathcal{T}_j\|_2 \geq \|\mathcal{T}_{1,j}\|_2 - \|\mathcal{T}_{2,j}\|_2$ and $\|\mathcal{T}_j\|_2 \geq \|\mathcal{T}_{2,j}\|_2 - \|\mathcal{T}_{1,j}\|_2$. Since $\|\mathcal{T}_{m,j}\|_2 \simeq 2^j$, then $C_1\, 2^j \leq \|\mathcal{T}_{m,j}\|_2 \leq C_2\, 2^j$ for $C_1, C_2 \neq 0$. Thus,

$$(C_2 - C_1)2^j \leq \|\mathcal{T}_j\|_2 \leq (C_2 + C_2)2^j.$$

Therefore, $\|\mathcal{T}_j\|_2 \simeq \|\mathcal{T}_{1,j}\|_2 + \|\mathcal{T}_{2,j}\|_2$. And inductively, $\|\mathcal{T}_j\|_2 \simeq \sum_{m=1}^{M} \|\mathcal{T}_{m,j}\|_2$. So, it is sufficient to show that $\|\mathcal{T}_{m,j}\|_2 \simeq 2^j$ for any $m$. We will consider below the case where $S_m$ is a vertical curve. The case where $S_m$ is a horizontal curve can be treated similarly.

By a suitable translation and rotation in the definition of $S_m$, we may assume that there is an $\epsilon > 0$ such that curve $S_m$ is vertical with $a_m = -\epsilon$, $b_m = \epsilon$, and that $f(0) = 0$, $f'(0) = 0$ and $g_m(0) = c \neq 0$ for some constant $c$. Letting $\phi(u) = -2\,\pi\, \cos\theta\,(f(u) + \tan\theta\, u)$, for $u \in (-\epsilon, \epsilon)$, and using equation 8 we can write

$$\widehat{\mathcal{T}}_{m,j}(\rho, \theta) = W(2^{-2j}(\rho\cos\theta, \rho\sin\theta)) \int_{-\epsilon}^{\epsilon} e^{i\rho\phi(u)} g_m(u)\, du$$

where $\phi'(u) = -2\pi(\cos(\theta)f'(u) + \sin(\theta)) = -2\pi\cos(\theta)(f'(u) + \tan(\theta))$ and $\phi''(u) = -2\pi\cos(\theta)f''(u)$.

We choose $\epsilon_0 > 0$ small enough so that $\epsilon_0 < \frac{1}{2}\epsilon$ and $g_m(u) \neq 0$ on $[-\epsilon_0, \epsilon_0]$. Let

$$\theta_0 = \min\{|\tan^{-1}(-f'(-\epsilon_0))|, |\tan^{-1}(-f'(\epsilon_0))|\}.$$

Remember that $\tan^{-1}$ is increasing. And also, since $f'' \neq 0$ on its domain, $f'$ is either increasing or decreasing. Therefore, $\tan^{-1}(-f')$ is either increasing or decreasing, hence bijective from $[-\epsilon_0, \epsilon_0]$ to

$$[\tan^{-1}(-f'(-\epsilon_0)), \tan^{-1}(-f'(\epsilon_0))] \supseteq [-\theta_0, \theta_0].$$

Therefore, for any $\theta \in [-\theta_0, \theta_0]$ or $\theta \in [\pi - \theta_0, \pi + \theta_0]$ there is a unique $u_\theta \in [-\epsilon_0, \epsilon_0]$ such that $\theta =$

$\tan^{-1}(-f'(u_\theta))$. Also since $\lim_{\theta\to(\pi/2+k\pi)}\tan(\theta)=\pm\infty$, then for $\theta\in[-\theta_0,\theta_0]$ or $\theta\in[\pi-\theta_0,\pi+\theta_0]$, we see that $cos(\theta)\neq 0$. Thus for $|\theta|\leq\theta_0$, or $|\theta-\pi|\leq\theta_0$,we can apply Lemma 3.2 to get,

$$\widehat{\mathcal{T}}_{m,j}(\rho,\theta)=W(2^{-2j}(\rho\cos\theta,\rho\sin\theta))\,\rho^{-\frac{1}{2}}\left(a(u_\theta)\,e^{-2\pi i\rho\phi(u_\theta)}+O(\rho^{-\frac{1}{2}})\right)$$

where

$$a(u_\theta)\quad=\quad\left(\frac{2\pi i}{\phi''(u_\theta)}\right)^{\frac{1}{2}}g_m(u_\theta)=(i\cos\theta\,f''(u_\theta))^{-\frac{1}{2}}\,g_m(u_\theta)\neq 0.$$

Since $0<c_1\leq|a(u_\theta)|\leq c_2$ for all $u_\theta\in[-\epsilon_0,\epsilon_0]$, from the conditions on the support of $W(2^{-2j}\xi)$ and omitting the higher order decay term in $\widehat{\mathcal{T}}_{m,j}(\rho,\theta)$ , we have that

$$\begin{aligned}
I_1 &= \int_{\mathbb{R}}\left[\int_{|\theta|\leq\theta_0}+\int_{|\theta-\pi|\leq\theta_0}\right]|\widehat{\mathcal{T}}_{m,j}(\rho,\theta)|^2\,d\theta\,\rho\,d\rho\\
&\simeq \int_{\frac{1}{16}2^{2j}}^{2^{2j}}\left[\int_{|\theta|\leq\theta_0}+\int_{|\theta-\pi|\leq\theta_0}\right]|W(2^{-2j}(\rho\cos\theta,\rho\sin\theta))|^2\,|a(u_\theta)|^2 d\theta\,\rho^{-1}\,\rho\,d\rho\\
&\simeq \int_{2^{2j-4}}^{2^{2j}}d\rho\simeq 2^{2j}.
\end{aligned}$$

For $\theta_0\leq|\theta|\leq\frac{\pi}{4}$ and $\theta_0\leq|\theta-\pi|\leq\frac{\pi}{4}$ and for $|u|\leq\epsilon$, we have $|\phi''(u)|=2\pi|\cos\theta||f''(u)|\geq c>0$. In this case, we apply Lemma 3.3 with $k=2$ to get

$$|\widehat{\mathcal{T}}_{m,j}(\rho,\theta)|\leq C|W(2^{-2j}(\rho\cos\theta,\rho\sin\theta))|\,\rho^{-\frac{1}{2}}.$$

We have

$$\begin{aligned}
I_2 &= \int_{\mathbb{R}}\left[\int_{\theta_0\leq|\theta|\leq\frac{\pi}{4}}+\int_{\theta_0\leq|\theta-\pi|\leq\frac{\pi}{4}}\right]|\widehat{\mathcal{T}}_{m,j}(\rho,\theta)|^2\,d\theta\,\rho\,d\rho\\
&\leq C\int_{\frac{1}{16}2^{2j}}^{2^{2j}}\left[\int_{\theta_0\leq|\theta|\leq\frac{\pi}{4}}+\int_{\theta_0\leq|\theta-\pi|\leq\frac{\pi}{4}}\right]|W(2^{-2j}(\rho\cos\theta,\rho\sin\theta))|^2 d\theta\,\rho^{-1}\,\rho\,d\rho\\
&\leq C\,2^{2j}.
\end{aligned}$$

For $\frac{\pi}{4} \leq |\theta| \leq \frac{\pi}{2}$ and $\frac{\pi}{4} \leq |\theta - \pi| \leq \frac{\pi}{2}$ and for $|u| \leq \epsilon$, we have $|\phi'(u)| = 2\pi(|\cos\theta f'(u) + \sin\theta|) \geq c > 0$, where we used the assumption that $|f'(u)| \leq \frac{1}{2}$ for $|u| \leq \epsilon$. Thus integration by parts gives

$$\left| \int_{-\epsilon}^{\epsilon} e^{i\rho\phi(u)} g_m(u)\, du \right| \leq C\, \rho^{-1}.$$

Then we have

$$\begin{aligned} I_3 &= \int_{\mathbb{R}} \left[ \int_{\frac{\pi}{4} \leq |\theta| \leq \frac{\pi}{2}} + \int_{\frac{\pi}{4} \leq |\theta - \pi| \leq \frac{\pi}{2}} \right] |\widehat{\mathcal{T}}_{m,j}(\rho,\theta)|^2\, d\theta\, \rho\, d\rho \\ &\leq C \int_{\frac{1}{16}2^{2j}}^{2^{2j}} \left[ \int_{\frac{\pi}{4} \leq |\theta| \leq \frac{\pi}{2}} + \int_{\frac{\pi}{4} \leq |\theta - \pi| \leq \frac{\pi}{2}} \right] |W(2^{-2j}(\rho\cos\theta, \rho\sin\theta))|^2 d\theta\, \rho^{-2}\, \rho\, d\rho \\ &\leq C. \end{aligned}$$

Since $\|\mathcal{T}_{m,j}\|_2^2 = I_1 + I_2 + I_3$, we finally have $\|\mathcal{T}_{m,j}\|_2^2 \simeq 2^{2j}$ and hence $\|\mathcal{T}_{m,j}\|_2 \simeq 2^j$.

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

We now present some preparation for the $\ell_1$-minimization algorithm. We recall that $L^2(\mathbb{R}^2) = \mathcal{H}$ where $\mathcal{H} = \mathcal{H}_K \oplus \mathcal{H}_M$, $\mathcal{H}_M = L^2(\mathcal{M}_{h_j})$ and $\mathcal{H}_K = L^2(\mathbb{R}^2 \setminus \mathcal{M}_{h_j})$. We want to recover $R_j^\ell = \operatorname{argmin}_{T_j} \|F^* \mathcal{T}_j\|_1$ subject to $f_j = P_{\mathbb{R}^2 \setminus \mathcal{M}_{h_j}} \mathcal{T}_j$. So, we require the following notion from [47]

**Definition 3.5.** *Let $F$ be a Parseval frame, and let $\Lambda$ be an index set of coefficients. We then define the concentration on $P_{\mathcal{M}_{h_j}}$ by*

$$\kappa = \kappa(\Lambda, \mathcal{H}_M) = \sup_{f \in \mathcal{H}_M} \frac{\|\mathbb{1}_\Lambda F^* f\|_1}{\|F^* f\|_1}.$$

This notion measures the total $\ell_1$ norm which can be concentrated to the index set $\Lambda$ restricted to functions in $\mathcal{H}_M$. Another important notion is that of *clustered sparsity.*

**Definition 3.6.** *Fix $\delta > 0$. Given a Hilbert space $\mathcal{H}$ with a Parseval frame $F$, $x \in \mathcal{H}$ is $\delta-$clustered sparse in $F$ (with respect to $\Lambda$) if*

$$\|\mathbb{1}_{\Lambda^c} F^* x\|_1 \leq \delta.$$

And from [46] we have

**Lemma 3.7.** *Fix $\delta > 0$ and suppose $x_o$ is $\delta -$ clustered sparse in $F$. Let $x^*$ be a solution of the $\ell_1$-minimization problem. Then,*

$$\|x^* - x^o\|_2 \leq \frac{2\delta}{1 - 2\kappa}$$

Now on, we will abuse notation slightly. For a Parseval frame $F$, we will write $P_M F = \{P_M \phi_i\}_i$ and$P_K F = \{P_K \phi_i\}_i$ for the projected frame elements. The following is a modified version of the notion of *mutual coherence* introduced by [19, 24].

**Definition 3.8.** *Let $F_1 = \{\phi_{1,i}\}_{i\in I}$ and $F_2 = \{\phi_{2,j}\}_{j\in J}$ lie in a Hilbert space $\mathcal{H}$ and let $\Lambda \subset I$. Then the cluster coherence $\mu_c(\Lambda, F_1 : F_2)$ of $F_1$ and $F_2$ with respect to $\Lambda$ is defined as*

$$\mu_c(\Lambda, F_1; F_2) = \max_{j\in J} \sum_{i\in\Lambda} |\langle \phi_{1,i}, \phi_{2,j}\rangle|.$$

And similarly from [46, 47],

**Lemma 3.9.**

$$\kappa(\Lambda, \mathcal{H}_M) \leq \mu_c(\Lambda, P_M F; P_M F) = \mu_c(\Lambda, P_M F; F).$$

Therefore, combining Lemmata 3.7 and 3.9 we have

**Lemma 3.10.** *Fix $\delta > 0$ and suppose that $x^o$ is $\delta -$ clustered sparse in $F$. Let $x^*$ solve the $\ell_1$-minimization problem, then*

$$\|x^* - x^o\|_2 \leq \frac{2\delta}{1 - 2\mu_c(\Lambda, P_M F; F)}.$$

**Remark 3.11.** *Notice that error decreases as linearly with the $\delta -$ clustered sparsity. Also we emphasize that both $\delta -$clustered sparsity and clustered coherence depend on the chosen "geometric set of indices" $\Lambda$. This helps us to determine whether the frame $F$ is a good dictionary for inpainting. However, $\Lambda$ is just an analysis tool and it is not needed explicitly in the $\ell_1$-minimization algorithm. The larger the set $\Lambda$ is, the smaller $\|\mathbb{1}_{\Lambda^c} F^* x^o\|_1$ is. That is, $x^o$ is $\delta -$ clustered sparse and the* *larger the cluster coherence. So, if $F$ sparsifies $x^o$ well, then a small set $\Lambda$ can be chosen to keep $\|\mathbb{1}_{\Lambda^c} F^* x^o\|_1$ small.*

**Remark 3.12.** *$\Lambda$ is the set of indices of the cluster of significant frame coefficients. Either for wavelet or shearlet.*

For the *thresholding part*, let $\mathcal{H}$ be a Hilbert space and fix $x^0 \in \mathcal{H}$. Let $E = \{e_\lambda : \lambda \in \Lambda\}$ be a Parseval frame on $\mathcal{H}$ and $P_K$, $P_M$ be projection operators on $\mathcal{H}$ such that $x_0 = P_K x^0 + P_M x^0$. Here $P_K x^0$ models the known part of the signal $x^0$ and $P_M x^0$ the missing part of $x^0$.

The *one-step-thresholding* algorithm from [47, Section 2.3] (version without noise) is the following.

**Algorithm 1.**

- **Input:** *The incomplete signal $\overline{x} = P_K x_0$; the Parseval frame $E = \{e_\lambda : \lambda \in \Lambda\}$; the thresholding parameter $\sigma$.*
- **Algorithm:**
  1. *Compute $\langle \overline{x}, e_i \rangle$ for all $i$;*
  2. *build the set $I = \{\lambda \in \Lambda : |\langle \overline{x}, e_\lambda \rangle| \geq \sigma\}$;*
  3. *compute $x^* = F\, \mathbb{1}_I\, F^* \overline{x}$.*
- **Output:** *The set $I$ of significant coefficients; the approximation $x^*$ to $x^0$.*

The next lemma, from [47, Proposition 3], gives an estimate o the approximation error of algorithm 1. We show the proof for completeness.

**Lemma 3.13.** *Let $I$ and $x^*$ be computed via Algorithm 1 with the assumption that all elements of the Parseval frame $E = \{e_\lambda : \lambda \in \Lambda\}$ have equal norm $\|e_i\| = e$ for all $\lambda \in \Lambda$. Then*

$$\|x^* - x^0\|_2 \leq e\,(\|\mathbb{1}_{I^c} F^* x^0\|_1 + \|\mathbb{1}_I F^* P_M x^0\|_1).$$

*Proof.* Since $x^* = F\mathbb{1}_I F^* P_K x^0$ and

$$\begin{aligned} x^0 &= P_K x^0 + P_M x^0 \\ &= FF^* P_K x^0 + FF^* P_M x^0 \\ &= F\mathbb{1}_I F^* P_K x^0 + F\mathbb{1}_{I^c} F^* P_K x^0 + F\mathbb{1}_I F^* P_M x^0 + F\mathbb{1}_{I^c} F^* P_M x^0, \end{aligned}$$

so,

$$\begin{aligned} \|x^* - x^0\|_2 &= \|F\mathbb{1}_I F^* P_K x^0 - \Big(F\mathbb{1}_I F^* P_K x^0 + F\mathbb{1}_{I^c} F^* P_K x^0 + F\mathbb{1}_I F^* P_M x^0 \\ &+ \ F\mathbb{1}_{I^c} F^* P_M x^0\Big)\|_2 \\ &= \|F\mathbb{1}_{I^c} F^* P_K x^0 + F\mathbb{1}_{I^c} F^* P_M x^0 + F\mathbb{1}_I F^* P_M x^0\|_2 \\ &= \|F\mathbb{1}_{I^c} F^* x^0 + F\mathbb{1}_I F^* P_M x^0]\|_2 \\ &\leq \|F\mathbb{1}_{I^c} F^* x^0\|_2 + \|F\mathbb{1}_I F^* P_M x^0\|_2. \end{aligned}$$

Remember $F : \ell^2 \to \mathcal{H}$ with $F(\{\alpha_\lambda\}_{\lambda\in\Lambda}) = \sum_{\lambda\in\Lambda} \alpha_\lambda e_\lambda$. Also we have $||e_\lambda||_2 = e$ for all $j$. So,

$$\left\|\sum_{n=1}^{N} \alpha_{\lambda_n} e_{\lambda_n}\right\|_2 \leq \sum_{n=1}^{N} |\alpha_{\lambda_n}| \, \|e_{\lambda_n}\|_2 = e \sum_{n=1}^{N} |\alpha_{\lambda_n}|.$$

Taking $N \to \infty$, we have

$$\|F(\{\alpha_\lambda\}_{\lambda\in\Lambda})\|_2 = \left\|\sum_{\lambda\in\Lambda} \alpha_\lambda e_\lambda\right\|_2 \leq e \sum_{\lambda\in\Lambda} |\alpha_\lambda| = e\|\{\alpha_\lambda\}_{\lambda\in\Lambda}\|_1.$$

$$\begin{aligned} Therefore, \|x^* - x^o\|_2 \leq & \quad \|F\mathbb{1}_I F^* P_M x^o\|_2 + ||F\mathbb{1}_{IC} F^* x^o\|_2 \\ \leq & \quad e\|\mathbb{1}_I F^* P_M x^o\|_1 + e\|\mathbb{1}_{IC} F^* x^o\|_1. \quad \square \end{aligned}$$

Given a Hilbert space $\mathcal{H}$ and a Parseval frame $E = \{e_\lambda : \lambda \in \Lambda\}$, a vector $x \in \mathcal{H}$ is $\delta$ *clustered* *sparse* in $E$ with respect to $I \subset \Lambda$ if there is a $\delta > 0$ such that $\|\mathbb{1}_{I_c} F^* x^0\|_1 = \sum_{\lambda \in I^c} |\langle x^0, e_\lambda \rangle| \leq \delta$, where $F^*$ is the frame analysis operator. For the approximation error in Lemma 3.13 to be small, the signal $x_0$ must be $\delta$ clustered sparse in $E$ with respect to $I$.

## 4 Inpainting using wavelets

In this section, we examine image inpainting using the wavelet system $\Phi = \{\phi_\lambda : \lambda \in \Lambda\}$ defined in section 1.2.

In all arguments below, it is enough to analyze a section of the curve $S_m$, with a fixed $m \in [1, M]$ because $\mathcal{T} = \sum_{m=1}^{M} \mathcal{T}_m$. Hence, to simplify the notation, in the following we denote $S_m$ by $S$ and $\mathcal{T}_{m,j}$ by $\mathcal{T}_j$. In addition, we will only consider the case where the curve section is locally vertical; the horizontal case can be treated in a very similar way.

### 4.1 Proof of Theorem 2.5 ($\ell_1$ minimization)

We denote the indices of the wavelet coefficients as $\Lambda = \bigcup_{j\in\mathbb{Z}} \Lambda_j$ where $\Lambda_j = \{(j,k) : k \in \mathbb{Z}^2\}$ for each level $j \in \mathbb{Z}$. We also denote as $S_{w,j} \subset \Lambda_j$ *the indices of the cluster of significant wavelet coefficients* and we define as

$$S_{w,j} := \{k = (k_1, k_2)\,,\, |k_1| \le 2 \cdot 2^{2j},\, |k_2| \le 2 \cdot 2^{2j}\}. \tag{9}$$

As in [20], for each of the sets $S_{w,j} \subset \Lambda_j$, we define the *wavelet approximation error* at the level $j$ as

$$\delta_j^w := \sum_{\lambda \in S_{w,j}^c} |\langle \mathcal{T}_j, \phi_\lambda \rangle| \tag{10}$$

and the *cluster coherences* as

$$\mu_c(S_{w,j}, P_{\mathcal{M}_{h_j}}\Phi; \Phi) := \max_{\lambda'} \sum_{\lambda \in S_{w,j}} |\langle P_{\mathcal{M}_{h_j}}\phi_\lambda, \phi_{\lambda'} \rangle|. \tag{11}$$

Using Lemma 3.10, we have that for any $j \in \mathbb{Z}$

$$\|R_j^\ell - \mathcal{T}_j\|_2 \le \frac{2\,\delta_j^w}{1 - 2\mu_c(S_{w,j}, P_{\mathcal{M}_{h_j}}\Phi; \Phi)}, \tag{12}$$

where $R_j^\ell$, $\mathcal{T}_j$ are defined as in Theorem 2.6, $\delta_j^w$ is given by (10) and $\mu_c$ by (11).

Therefore, Theorem 2.5 follows directly from the two propositions below and (12).

**Proposition 4.1.** *For any $j \in \mathbb{Z}$*

$$\delta_j^w = o(2^j) = o(\|\mathcal{T}_j\|_2).$$

**Proposition 4.2.** *Assume that $h_j = o(2^{-2j})$. Then*

$$\mu_c(S_{w,j}, P_{\mathcal{M}_{h_j}}\Phi; \Phi) \to 0 \quad \textit{as } j \to \infty.$$

The respective proofs follow in the following subsections.

#### 4.1.1 Proof of Proposition 4.1

Let $\beta_{j,k} = \langle \widehat{\phi}_{j.k}, \widehat{\mathcal{T}}_j \rangle$. Thus, we can write (10) as $\delta_j^w = \sum_{k \in S_{w,j}^c} |\beta_{j,k}|$. So, we have

$$\begin{aligned}
\beta_{j,k} &= 2^{-2j} \int_a^b \left[ \int_{\widehat{\mathbb{R}}^2} |W(2^{-2j}\xi)|^2 e^{(2\pi i \xi 2^{-2j}(k+2^{2j}(f(u),u)))} d\xi \right] g(u) du \\
&= 2^{2j} \int_a^b \left[ \int_{\widehat{\mathbb{R}}^2} |W(\xi)|^2 e^{(2\pi i \xi (k+2^{2j}(f(u),u)))} d\xi \right] g(u) du.
\end{aligned}$$

We define the operator

$$L = \Big(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial z_1^2}\Big)\Big(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial z_2^2}\Big). \tag{13}$$

Hence, using (55) in the Appendix

$$\begin{aligned}
|\beta_{j,k}| &= 2^{2j} \left| \int_a^b \int_{\widehat{\mathbb{R}}^2} L^N \Big(|W(\xi)|^2\Big) L^{-N} e^{2\pi i \xi \left(k+2^{2j}(f(u),u)\right)} d\xi \, g(u) \, du \right| \\
&= 2^{2j} \Big| \int_a^b \Big(1 + (k_1 + 2^{2j} f(u))^2\Big)^{-N} \Big(1 + (k_2 + 2^{2j} u)^2\Big)^{-N} \\
&\times \int_{\widehat{\mathbb{R}}^2} L^N \Big(|W(\xi)|^2\Big) e^{2\pi i \xi \left(k+2^{2j}(f(u),u)\right)} d\xi g(u) du \Big|
\end{aligned}$$

$$\leq \quad 2^{2j} C_N \int_a^b \left(1 + (k_1 - 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 - 2^{2j} u)^2\right)^{-N} |g(u)| du$$

We know $|u|, |f(u)| \leq 1$. If $k \in S_{w,j}^c$, then from 9, either $|k_1| > 2 \cdot 2^{2j}$ or $|k_2| > 2 \cdot 2^{2j}$. So,

$$1 + (k_1 + 2^{2j} f(u))^2 \geq (k_1 + 2^{2j} f(u))^2 \geq (|k_1| - 2^{2j} |f(u)|)^2 \geq 2^{4j}.$$

Similarly, we see $1 + (k_2 - 2^{2j}|u|)^2 \geq 2^{4j}$. Let $A = \{(k_1, k_2 : |k_1| > 2 \cdot 2^{2j}\}$ and let $B = \{(k_1, k_2 : |k_2| > 2 \cdot 2^{2j}\}$. Notice that $S_{w,j}^c = A \cup B$. For $N = 2$ we have

$$\begin{aligned}
&\sum_{k \in S_{w,j}^c} |\beta_{j,k}| \\
&\leq \sum_{k \in A} |\beta_{j,k}| + \sum_{k \in B} |\beta_{j,k}| \\
&\leq C \sum_{k \in A} 2^{2j} \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-2} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)|\, du \\
&+ C \sum_{k \in B} 2^{2j} \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-2} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)|\, du \\
&\leq C\, 2^{2j} \int_a^b \sum_{|k_1| \geq 2 \cdot 2^{2j}} 2^{-4j} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-1} \sum_{k_2 \in \mathbb{Z}} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)|\, du \\
&+ C\, 2^{2j} \int_a^b \sum_{|k_2| \geq 2 \cdot 2^{2j}} 2^{-4j} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-1} \sum_{k_1 \in \mathbb{Z}} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-2} |g(u)|\, du \\
&\leq C\, 2^{2j}\, 2^{-2j} = o(2^j). \quad \Box
\end{aligned}$$

### 4.1.2 Proof of Proposition 4.2

Using Fourier Transform, Plancherel's formula and A.2, we see

$$\begin{aligned}
\langle P_{\mathcal{M}_{h_j}} \phi_{j,k}, \phi_{j,k'} \rangle &= \langle \widehat{\mathbb{1}_{\mathcal{M}_{h_j}}} * \widehat{\phi_{j,k}}, \widehat{\phi_{j,k'}} \rangle \\
&= 2h_j \int_{\widehat{\mathbb{R}}^2} \int_{\mathbb{R}} \operatorname{sinc}(2\pi h_j \eta_2)\, \widehat{\phi_{j,k}} \left(\xi - (0, \eta_2)\right) d\eta_2\, \overline{\widehat{\phi_{j,k'}}}(\xi)\, d\xi \\
&= 2h_j 2^{-4j} \int_{\widehat{\mathbb{R}}^2} \int_{\mathbb{R}} W(2^{-2j}(\xi_1, \xi_2 - \eta_2)) \operatorname{sinc}(2\pi h_j \eta_2) \\
&\times \; e^{-2\pi i 2^{-2j} \eta_2 k_2} d\eta_2\, \overline{W}(2^{-2j} \xi)\, e^{2\pi i \xi 2^{-2j} (k - k')}\, d\xi.
\end{aligned}$$

Making the change of variables $\tau = 2^{-2j}\xi$ and $\gamma_2 = 2^{-2j}\eta_2$, we have

$$\langle P_{\mathcal{M}_{h_j}}\phi_{j,k}, \phi_{j,k'}\rangle = 2h_j\, 2^{2j} \int_{\widehat{\mathbb{R}}^2} g(\tau)\, e^{2\pi i\tau(k-k')}\, d\tau.$$

where the function

$$g(\tau) = \int_{\mathbb{R}} W(\tau - (0,\gamma_2))\, \overline{W}(\tau)\, \mathrm{sinc}(2\pi h_j 2^{2j}\gamma_2)\, e^{-2\pi i\gamma_2 k_2}\, d\gamma_2$$

is compactly supported because W is. Additionally, by dominated convergence theorem, [28], $g$ is smooth; and therefore,

$$L(g(\tau)) = \int_{\mathbb{R}} L\Big(W(\tau - (0,\gamma_2))\, \overline{W}(\tau)\Big)\, \mathrm{sinc}(2\pi h_j 2^{2j}\gamma_2)\, e^{-2\pi i\gamma_2 k_2}\, d\gamma_2.$$

So, we see

$$\begin{aligned} |L(g(\tau))| &\leq \int_{\mathbb{R}} \Big|L\Big(W(\tau - (0,\gamma_2))\, \overline{W}(\tau)\Big)\Big|\, |\mathrm{sinc}(2\pi h_j 2^{2j}\gamma_2)|\, |e^{-2\pi i\gamma_2 k_2}|\, d\gamma_2 \\ &\leq \int_{\mathbb{R}} \Big|L\Big(W(\tau - (0,\gamma_2))\, \overline{W}(\tau)\Big)\Big|\, d\gamma_2 \leq C. \end{aligned}$$

is bounded because $W$ is smooth and compactly supported and $|\mathrm{sinc}| \leq 1$. Notice also that $C$ is independent from $k$, $k'$ and $h_j$. Hence we can apply Lemma A.1 with (55) and the differential operator $L$ given by (13) to get

$$\begin{aligned} &|\langle P_{\mathcal{M}_{h_j}}\phi_{j,k}, \phi_{j,k'}\rangle| \\ &= 2h_j 2^{2j} \left| \int_{\widehat{\mathbb{R}}^2} L\left(g(\tau)\right) L^{-1}\left(e^{2\pi i\tau(k-k')}\right) d\tau \right| \\ &= 2h_j 2^{2j}(1+(k_1-k_1')^2)^{-1}(1+(k_2-k_2')^2)^{-1} \left| \int_{\widehat{\mathbb{R}}^2} L\left(g(\tau)\right)\, e^{2\pi i\tau(k-k')} d\tau \right| \\ &\leq C 2h_j 2^{2j}(1+(k_1-k_1')^2)^{-1}(1+(k_2-k_2')^2)^{-1}. \end{aligned}$$

From there, we see that

$$\begin{aligned}\sum_{k\in S_{w,j}} |\langle P_{\mathcal{M}_{h_j}}\phi_{j,k},\phi_{j,k'}\rangle| &\le C\,2^{2j}\,h_j \sum_{k\in\mathbb{Z}^2}(1+(k_1-k_1')^2)^{-1}(1+(k_2-k_2')^2)^{-1}\\ &\le C\,2^{2j}\,h_j.\end{aligned}$$

Since $h_j = o(2^{-2j})$, then $\mu_c(S_{w,j}, P_{\mathcal{M}_{h_j}}\Phi;\Phi)\to 0$, as $j\to\infty$. $\square$

Therefore, we have proved Theorem 2.5.

## 4.2 Proof of Theorem 2.6 (Thresholding)

We now apply Algorithm 1 to the signal $\mathcal{T}$ using the Parseval frame of wavelets $\Phi = \{\phi_{j,k} : j \in \mathbb{Z}, k\in\mathbb{Z}^2\}$ defined in Section 1.2. Note that $\left\|\phi_{j,k}\right\|_2 = \|\phi\|_2$ for all $j\in\mathbb{Z}, k\in\mathbb{Z}^2$.

For any $j\in\mathbb{Z}$, $k\in\mathbb{Z}^2$, let $\gamma_{j,k} = \langle\phi_{j,k}, P_{\mathcal{M}_{h_j}}\mathcal{T}_j\rangle$, $\beta_{j,k} = \langle\phi_{j,k},\mathcal{T}_j\rangle$ and $\alpha_{j,k} = \beta_{j,k}-\gamma_{j,k}$. For $j\ge 0$, $0\le\sigma_j\le 2^{-4j}$, we set $I_j = \{k\in\mathbb{Z}^2 : |\alpha_{j,k}|\ge\sigma_j\}$. By applying Lemma 3.13, we obtain the following estimate.

**Proposition 4.3.** *Fix $j\in\mathbb{Z}$ and let the set of significant coefficients $I_j$ be given as above. Let the approximation $R_j$ of the function $\mathcal{T}_j$ be computed according to Algorithm 1. Then*

$$\|R_j^\tau - \mathcal{T}_j\|_2 \le \|\phi\|_2\,(\|\mathbb{1}_{I_j^c}F^*\mathcal{T}_j\|_1 + \|\mathbb{1}_{I_j}F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1).$$

Observe also that

$$\|\mathbb{1}_{I_j^c}F^*\mathcal{T}_j\|_1 = \sum_{k\in I_j^c}|\langle\phi_{j,k},\mathcal{T}_j\rangle| = \sum_{k\in I_j^c}|\beta_{j,k}|,$$

$$\|\mathbb{1}_{I_j}F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 = \sum_{k\in I_j}|\langle\phi_{j,k},P_{\mathcal{M}_{h_j}}\mathcal{T}_j\rangle| = \sum_{k\in I_j}|\gamma_{j,k}|$$

and $R_j^\tau = F[\mathbb{1}_{I_j}F^*P_{\mathbb{R}^2\setminus\mathcal{M}_{h_j}}\mathcal{T}_j]$. Theorem 2.6 is proved using the fact that $\|\mathbb{1}_{I_j}F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 \le \|F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1$, Proposition 4.3 above and the following result.

**Proposition 4.4.** *Fix $j \in \mathbb{Z}$. For any $0 \leq \sigma_j \leq 2^{-4j}$ and $h_j = o(2^{-j})$, we have*

$$(i) \quad \|F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1 = \sum_{k \in \mathbb{Z}^2} |\gamma_{j,k}| \leq C\, 2^{2j}\, h_j = o(\|\mathcal{T}_j\|_2); \tag{14}$$

$$(ii) \quad \sum_{k \in I_j^c} |\beta_{j,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2), \quad \text{as } j \to \infty. \tag{15}$$

### 4.2.1 Proof of Proposition 4.4

We again use Plancherel Theorem and the change of variable $2^{-j}\xi = \eta$. So, we have

$$\begin{aligned} \gamma_{j,k} &= \langle \widehat{\phi}_{j,k}, \widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j} \rangle \\ &= 2^{-2j} \int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi)\, e^{2\pi i 2^{-2j}\xi k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(\xi)}\, d\xi \\ &= 2^{2j} \int_{\widehat{\mathbb{R}}^2} W(\eta)\, e^{2\pi i \eta k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(2^{2j}\eta)}\, d\eta. \end{aligned} \tag{16}$$

We also see that

$$\begin{aligned} \mathcal{T}_j(x) &= \int_{\widehat{\mathbb{R}}^2} \widehat{\mathcal{T}_j(\xi)}\, e^{2\pi i \xi x}\, d\xi \\ &= \int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi) \left( \int_a^b e^{-2\pi i \xi \cdot (f(u),u)}\, g(u)\, du \right) e^{2\pi i \xi x}\, d\xi \\ &= \int_a^b \left( 2^{4j} \int_{\widehat{\mathbb{R}}^2} W(\eta)\, e^{2\pi i 2^{2j} \eta \cdot \left(x - (f(u),u)\right)}\, d\eta \right) g(u) du \\ &= 2^{4j} \int_a^b \check{W}\left( 2^{2j} \left( x - (f(u),u) \right) \right) g(u)\, du \end{aligned}$$

where $\check{W}$ is the inverse Fourier Transform of $W$. Therefore,

$$\begin{aligned} \widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(2^{2j}\eta) &= \int_{\mathbb{R}^2} \mathbb{1}_{\mathcal{M}_{h_j}}(x) \mathcal{T}_j(x) e^{2\pi i 2^{2j} \eta \cdot x} dx \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{\mathcal{M}_{h_j}}(x)\, 2^{4j} \int_a^b \check{W}\left( 2^{2j} \left( x - (f(u),u) \right) \right) g(u)\, du\, e^{-2\pi i 2^{2j} \eta \cdot x} dx \\ &= \int_{\mathbb{R}^2} \left( \int_a^b \mathbb{1}_{\mathcal{M}_{h_j}} \left( x + (f(u),u) \right) e^{-2\pi i 2^{2j} \eta\, (x + (f(u),u))} g(u) du \right) 2^{4j} \check{W}(2^{2j} x) dx. \end{aligned}$$

Therefore, in (16) we have

$$
\begin{aligned}
\gamma_{j,k} &= 2^{2j}\int_{\widehat{\mathbb{R}}^2} W(\eta)\, e^{2\pi i\eta k}\,\overline{\widehat{P_{\mathcal{M}_{h_j}}\mathcal{T}_j}(2^{2j}\eta)}\, d\eta \\
&= 2^{2j}\int_{\widehat{\mathbb{R}}^2}\int_{\mathbb{R}^2}\left(\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)\, e^{2\pi i 2^{2j}\eta\,(x+(f(u),u))} g(u)du\right) \\
&\times\ 2^{4j}\overline{\check{W}(2^{2j}x)}\, dx\; W(\eta)e^{2\pi i\eta k}d\eta \\
&= 2^{2j}\int_{\mathbb{R}^2}\left[\int_a^b\left(\int_{\widehat{\mathbb{R}}^2} W(\eta)e^{2\pi i\eta\cdot\left(k+2^{2j}(x+(f(u),u))\right)}d\eta\right)\right. \\
&\times\ \left.\mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)g(u)du\right]2^{4j}\overline{\check{W}(2^{2j}x)}dx.
\end{aligned}
$$

From proposition A.1, we have

$$
\begin{aligned}
\int_{\widehat{\mathbb{R}}^2} W(\eta)e^{2\pi i\eta\cdot\left(k+2^{2j}(x+(f(u),u))\right)}d\eta &= \int_{\mathbb{R}^2} L\Big[W(\eta)\Big]L^{-1}\Big[e^{2\pi i\eta\cdot\left(k+2^{2j}(x+(f(u),u))\right)}\Big]d\eta \\
&= \Big(1+\big(k_1+2^{2j}(x_1+f(u))\big)^2\Big)^{-1} \\
&\times\ \Big(1+\big(k_2+2^{2j}(x_2+u)\big)^2\Big)^{-1} \\
&\times\ \int_{\mathbb{R}^2} L\Big[W(\eta)\Big]e^{2\pi i\eta\cdot\left(k+2^{2j}(x+(f(u),u))\right)}d\eta
\end{aligned}
$$

So, we get

$$
\begin{aligned}
\sum_{k\in\mathbb{Z}^2}|\gamma_{j,k}| &\le 2^{2j}\int_{\mathbb{R}^2}\left[\int_a^b\sum_{k\in\mathbb{Z}^2}\Big(1+\big(k_1+2^{2j}(x_1+f(u))\big)^2\Big)^{-1}\Big(1+\big(k_2+2^{2j}(x_2+u)\big)^2\Big)^{-1}\right. \\
&\times\ \left.\left(\int_{\mathbb{R}^2}\Big|L\left[W(\eta)\right]\Big|d\eta\right)\mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)|g(u)|du\right]2^{4j}\Big|\check{W}\Big(2^{2j}x\Big)\Big|dx \\
&\le 2^{2j}\int_{\mathbb{R}^2}\left[C\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)\big|g(u)\big|du\right]2^{4j}\Big|\check{W}\Big(2^{2j}x\Big)\Big|dx \\
&= 2^{2j}\int_{|x|\le 2^{-(2-\Delta_0)j}}\left[C\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)\big|g(u)\big|du\right]2^{4j}\Big|\check{W}\Big(2^{2j}x\Big)\Big|dx \\
&+\ 2^{2j}\int_{|x|> 2^{-(2-\Delta_0)j}}\left[C\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x+(f(u),u)\big)\big|g(u)\big|du\right]2^{4j}\Big|\check{W}\Big(2^{2j}x\Big)\Big|dx \\
&= I_1+I_2.
\end{aligned}
\tag{17}
$$

Notice that $\mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))=1$ whenever $-h_j-x_2\le u_2\le h_j-x_2$. So, for any $\Delta_0>0$ and

$|x| \leq 2^{-(2-\Delta_0)j}$ we have:

$$\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x + (f(u),u)\big)\big|g(u)\big|du \leq C_1 h_j$$

where $C_1$ is independent of $j$. Thus

$$\begin{aligned}
I_1 &= 2^{2j} C \int_{|x|\leq 2^{-(2-\Delta_0)j}} \left[\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\left(x + (f(u),u)\right)|g(u)|du\right] 2^{4j} \left|\check{W}\left(2^{2j}x\right)\right| dx \\
&\leq 2^{2j} C \int_{|x|\leq 2^{-(2-\Delta_0)j}} C_1 h_j 2^{4j} \left|\check{W}\left(2^{2j}x\right)\right| dx
\end{aligned}$$

Remember $W$ is compactly supported and $\widehat{\check{W}} = W$. So, using Lemma A.3, for each $N \in \mathbb{N}$ there is $C_N$ such that $\left|\check{W}(z)\right| \leq C_N(1+|z|^2)^{-N}$. Therefore,

$$\begin{aligned}
I_1 &\leq 2^{2j} C \int_{|x|\leq 2^{-(2-\Delta_0)j}} C_1 h_j 2^{4j} \left|\check{W}\left(2^{2j}x\right)\right| dx \\
&\leq 2^{2j} C \int_{|x|\leq 2^{-(2-\Delta_0)j}} C_1 h_j 2^{4j} C_N (1+|2^{2j}x|^2)^{-N} dx \\
&\leq 2^{2j} C h_j C_N \int_{|x|\leq 2^{\Delta_0 j}} (1+|x|^2)^{-N} dx \\
&\leq 2^{2j} C h_j \qquad (18)
\end{aligned}$$

for a suitable value of $N$ where $C$ is an independent constant. For $|x| > 2^{-(2-\Delta_0)j}$, notice $\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\big(x + (f(u),u)\big)\big|g(u)\big|du \leq C$ where $C$ is an independent constant. Therefore, applying again Lemma A.3, we have

$$\begin{aligned}
I_2 &= 2^{2j} \int_{|x|>2^{-(2-\Delta_0)j}} \left[C \int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}\left(x + (f(u),u)\right)|g(u)|du\right] 2^{4j} \left|\check{W}\left(2^{2j}x\right)\right| dx \\
&\leq 2^{2j} \int_{|x|>2^{\Delta_0 j}} C C_N (1+|x|^2)^{-N} dx \\
&= 2^{2j} C C_N \frac{\pi}{N-1} (1+2^{2\Delta_o j})^{-N+1}. \qquad (19)
\end{aligned}$$

So, for a suitable value of $N$, we get $I_2 \leq 2^{2j} C h_j$ where $j$ is large enough. Therefore, combining (18) and (19) into (17), we have

$$\sum_{k\in\mathbb{Z}^2} |\gamma_{j,k}| \;\; \le \;\; I_1 + I_2 \le C 2^{2j} h_j$$

where $C$ is an independent constant. This proves (14). Now, for (15), we again use Plancherel's formula, the change of variable $\eta = 2^{-2j}\xi$ and Lemma A.1 in the Appendix. So, we have for any $N \in \mathbb{N}$

$$\begin{aligned}
\beta_{j,k} &= \langle \widehat{\mathcal{T}}_j, \widehat{\phi}_{j,k} \rangle \\
&= \int_{\widehat{\mathbb{R}}^2} \left( \int_a^b e^{-2\pi i \xi\cdot(f(u),u)} g(u) du W(2^{-2j}\xi) \right) \left( 2^{-2j}\overline{W}(2^{-2j}\xi) e^{-2\pi i 2^{-2j}\xi\cdot k} \right) d\xi \\
&= 2^{2j} \int_a^b \left[ \int_{\widehat{\mathbb{R}}^2} |W(\eta)|^2 e^{-2\pi i \eta\cdot\left(k+2^j(f(u),u)\right)} d\eta \right] g(u) du \\
&= 2^{2j} \int_a^b \left[ \int_{\mathbb{R}^2} L^N \left( |W(\eta)|^2 \right) \right. \\
&\times \left. \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} e^{-2\pi i \eta\cdot\left(k+2^j(f(u),u)\right)} d\eta \right] g(u) du.
\end{aligned}$$

Since $W$ is smooth and compactly supported, for any $N \in \mathbb{N}$, there is a constant $C_N$ such that

$$|\beta_{j,k}| \;\; \le \;\; 2^{2j} C_N \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du. \tag{20}$$

Let $K_j = \{k : |k_1| \le 2^{2j+1}, |k_2| \le 2^{2j+1}\}$. If $k \in K_j^c$, then either $|k_1| > 2^{2j+1}$ or $|k_2| > 2^{2j+1}$. As before, if $|k_1| > 2^{2j+1}$, for all $|f(u)|, |u| \le 1$, it follows that

$$1 + (k_1 + 2^{2j} f(u))^2 \ge (k_1 + 2^{2j} f(u))^2 \ge 2^{4j}.$$

Similarly, if $|k_2| > 2^{2j+1}$, we have $1 + (k_1 + 2^{2j} f(u))^2 \ge 2^{4j}$. We proceed similarly to Proposition 4.1. From (20), we have

$$\sum_{k\in K_j^c} |\beta_{j,k}| \;\; \le \;\; C_N \, 2^{2j} \int_a^b \sum_{k\in K_j^c} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du$$

$$
\begin{aligned}
&\le C_N\, 2^{2j} \int_a^b \sum_{|k_1|>2^{2j+1}, k_2\in\mathbb{Z}} \Big(1+(k_1+2^{2j}f(u))^2\Big)^{-N} \Big(1+(k_2+2^{2j}u)^2\Big)^{-N} \\
&+ \sum_{k_1\in\mathbb{Z}, |k_2|>2^{2j+1}} \Big(1+(k_1+2^{2j}f(u))^2\Big)^{-N} \Big(1+(k_2+2^{2j}u)^2\Big)^{-N} |g(u)|du \\
&\le C_N\, 2^{2j} \int_a^b 2^{(1-N)4j} \sum_{k\in\mathbb{Z}^2} \Big(1+(k_1+2^{2j}f(u))^2\Big)^{-1} \Big(1+(k_2+2^{2j}u)^2\Big)^{-N} \\
&+ 2^{(1-N)4j} \sum_{k\in\mathbb{Z}^2} \Big(1+(k_1+2^{2j}f(u))^2\Big)^{-N} \Big(1+(k_2+2^{2j}u)^2\Big)^{-N} |g(u)|du \\
&\le C_N 2^{-2(2N-3)j}. \qquad (21)
\end{aligned}
$$

Recall that $I_j = \{k \in \mathbb{Z}^2 : |\alpha_{j,k}| \ge \sigma_j\}$ and $\beta_{j,k} = \alpha_{j,k} + \gamma_{j,k}$. So, using (21) and (14), we have

$$
\begin{aligned}
\sum_{k\in I_j^c} |\beta_{j,k}| &\le \sum_{k\in I_j^c \bigcap K_j} |\beta_{j,k}| + \sum_{k\in I_j^c \bigcap K_j^c} |\beta_{j,k}| \\
&\le \sum_{k\in I_j^c \bigcap K_j} |\alpha_{j,k}| + \sum_{k\in I_j^c \bigcap K_j} |\gamma_{j,k}| + \sum_{k\in I_j^c \bigcap K_j^c} |\beta_{j,k}| \\
&\le \sum_{k\in I_j^c \bigcap K_j} |\alpha_{j,k}| + \sum_{k\in \mathbb{Z}^2} |\gamma_{j,k}| + \sum_{k\in K_j^c} |\beta_{j,k}| \\
&\le \sum_{k\in I_j^c \bigcap K_j} |\alpha_{j,k}| + C\, 2^{2j}\, h_j + C_N 2^{-2(2N-3)j}. \qquad (22)
\end{aligned}
$$

Observe that $K_j = \{k \in \mathbb{Z}^2 : |k_1|, |k_2| \le 2^{2j+1}\}$, so $\#|K_j| = (2^{2j+1}\cdot 2+1)^2$. Thus,

$$
\sum_{k\in I_j^C\cap K_j} |\alpha_{j,k}| \le \sum_{k\in I_j^C\cap K_j} \sigma_j \le \sum_{k\in K_j} \sigma_j = \sigma_j \sum_{k\in K_j} 1 = \sigma_j \#|K_j| = \sigma_j (2^{2j+1}\cdot 2+1)^2 \le 2^6
$$

since $\sigma_j \le 2^{-4j}$. Using (22) and the inequality above, we have

$$
\begin{aligned}
\|\mathbb{1}_{I_j^c} F * \mathcal{T}_j\|_1 &= \sum_{k\in I_j^C} |\beta_{j,k}| \le \sum_{k\in I_j^C\cap K_j} |\alpha_{j,k}| + C_1 2^{-2(2N-3)j} + C_2 2^{2j} h_j \\
&\le 2^6 + C_N 2^{-2(2N-3)j} + C_2 2^{2j} h_j = o(2^j)
\end{aligned}
$$

where $h_j = o(2^{-j})$ and $N = 2$. $\square$

# 5 Inpainting using shearlets

Now we analyze the inpainting problem using shearlet system $\Psi = \{\Psi_\eta : \eta \in M\}$ introduced in section 1.2. To do this analysis, we need to study the coefficients $\langle \psi^{(\nu)}_{j,\ell,k}, \mathcal{T}_{m,j} \rangle$, $\nu = 1$ or $\nu = 2$, in four different cases:

(1) $\psi^{(1)}_{j,\ell,k}$ is horizontal and the curve for $S_m$ is vertical,

(2) $\psi^{(2)}_{j,\ell,k}$ is vertical and the curve for $S_m$ is vertical,

(3) $\psi^{(1)}_{j,\ell,k}$ is horizontal and the curve for $S_m$ is horizontal,

(4) $\psi^{(2)}_{j,\ell,k}$ is vertical and the curve for $S_m$ is horizontal.

Since cases (1) and (2) are analogous to cases (3) and (4), we need only to consider cases (1) and (2). We also remark that boundary shearlets have localization and regularity properties very similar to the shearlet functions $\psi^{(\nu)}_{j,\ell,k}$, $\nu = 1, 2$, hence the same argument holds for such elements. We can fix $m$ for the locally vertical curve $S_m$ since no horizontal curve need to be examined. Also, as in the previous Section 4, we will also denote $S_m$ by $S$ and $\mathcal{T}_{m,j}$ by $\mathcal{T}_j$ to simplify the notation in this section.

## 5.1 Proof of Theorem 2.7 ($\ell_1$ minimization)

Let $\Psi = \{\psi_\eta : \eta \in M\}$ be the shearlet system where

$$M = \{\eta = (j, \ell, k, \nu) : \, j \geq 0, |\ell| \leq 2^j, k \in \mathbb{Z}^2, \nu = 1, 2\}.$$

We can write $M = M^{(1)} \cup M^{(2)}$, where $M^{(i)} = \{\eta = (j, \ell, k, \nu) \in M : \nu = i\}$, for $i = 1, 2$, and, for each $i$, $M^{(i)} = \bigcup_{j \geq 0} M^{(i)}_j$, where $M^{(i)}_j = \{(j', \ell, k) \in M^{(i)} : \, j' = j\}$.

As in Section 4, for each $j \in \mathbb{Z}$, we denote as $S_{s,j}$ the set of *indices of the cluster of significant shearlet coefficients* (at scale $j$). The explicit definition of this set will be given below, in the proof of Proposition 5.1. Corresponding to this set, we define the *shearlet approximation error* at the level $j$ as $\delta_j^s = \sum_{\eta \in S_{s,j}^c} |\langle \mathcal{T}_j, \psi_\eta \rangle|$ and the *cluster coherence* as

$$\mu_c(S_{s,j}, P_{\mathcal{M}_{h_j}}\Psi; \Psi) = \max_{\eta'} \sum_{\eta \in S_{s,j}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}, \psi_\eta \rangle|.$$

It will be convenient to write $S_{s,j} = S_{s,j,1} \bigcup S_{s,j,2} \subset M$, where we set $S_{s,j,2} = \emptyset$ and $S_{s,j,1} \subset M_j^{(1)}$ will be determined below. Since $S_{s,j,2} = \emptyset$, we have

$$\begin{aligned} \max_{\eta'} \sum_{\eta \in S_{s,j}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}, \psi_\eta \rangle| &\leq \max_{\eta'} \sum_{\eta \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}^{(1)}, \psi_\eta^{(1)} \rangle| \\ &+ \max_{\eta'} \sum_{\eta \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}^{(2)}, \psi_\eta^{(1)} \rangle|. \end{aligned}$$

Just like in the wavelet case, from Lemma 3.10 we have

$$\|R_j^\ell - \mathcal{T}_j\|_2 \leq \frac{2\delta_j^s}{1 - 2\mu_c(S_{s,j}, P_{\mathcal{M}_{h_j}}\Psi; \Psi)}. \tag{23}$$

Let $\beta_{j,\ell,k}^{(\nu)} = \langle \widehat{\psi_{j,\ell,k}^{(\nu)}}, \widehat{\mathcal{T}}_j \rangle$. To prove Theorem 2.7, we need (23) and to prove the propositions stated below where $S_{s,j,1}$ is also constructed. We present the proofs of the next propositions in the following subsections.

**Proposition 5.1.** *For any $j \in \mathbb{Z}$,*

$$\delta_j^s = \sum_{(\ell,k) \in M_j^{(1)} \setminus S_{s,j,1}} |\beta_{j,\ell,k}^{(1)}| + \sum_{(\ell,k) \in M_j^{(2)}} |\beta_{j,\ell,k}^{(2)}| = o(2^j) = o(\|\mathcal{T}_j\|_2). \tag{24}$$

**Proposition 5.2.** *Assume that $h_j = o(2^{-j})$. Then*

$$\max_{\eta'} \sum_{\eta \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}^{(1)}, \psi_\eta^{(1)} \rangle| \to 0 \ \textit{as}\ j \to \infty; \tag{25}$$

$$\max_{\eta'} \sum_{\eta \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi_{\eta'}^{(2)}, \psi_\eta^{(1)} \rangle| \to 0 \ \textit{as}\ j \to \infty. \tag{26}$$

#### 5.1.1 Proof of Propositions 5.1

We recall that $\overline{\widehat{\mathcal{T}}}(\xi) = \int_a^b e^{2\pi i\xi\cdot(f(u),u)} g(u)du$ where $[a,b] \subset [-1,1]$ and $|f(u)| \leq 1$. Using Plancherel Theorem, we have

$$\begin{aligned}
\beta^{(1)}_{j,\ell,k} &= \langle \widehat{\psi}^{(1)}_{j,\ell,k}, \widehat{\mathcal{T}}_j \rangle \\
&= \int_{\mathbb{R}^2} \widehat{\psi}^{(1)}_{j,\ell,k} \overline{\widehat{\mathcal{T}}}(\xi) d\xi \\
&= 2^{-3j/2} \int_a^b \left( \int_{\mathbb{R}^2} |W(2^{-2j}\xi)|^2 V\Big(2^j \frac{\xi_2}{\xi_1} - \ell\Big) e^{2\pi i\xi\cdot(A_1^{-j}B_1^{-\ell}k + (f(u),u))} d\xi \right) g(u)du. \qquad (27)
\end{aligned}$$

We make a change of variable $\eta = \xi A_1^{-j} B_1^{-\ell}$; therefore,

$$\begin{aligned}
\xi \cdot (A_1^{-j}B_1^{-\ell}k + (f(u),u)) &= \eta \cdot (k + B_1^{\ell}A_1^{j}(f(u),u)) \\
&= \eta \cdot (k_1 + 2^{2j}f(u) + 2^j\ell u, k_2 + 2^j u).
\end{aligned}$$

Let $x = (k_1 + 2^{2j}f(u) + 2^j\ell u, k_2 + 2^j u)$. So we have:

$$\beta^{(1)}_{j,\ell,k} = 2^{3j/2} \int_a^b \left( \int_{\mathbb{R}^2} \left|W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\right|^2 V(\eta_2/\eta_1) e^{2\pi\eta\cdot(k_1+2^{2j}f(u)+2^j\ell u, k_2+2^j u)} d\eta \right) g(u)du.$$

Remember that $\operatorname{supp} W \subset [-1/2,1/2]^2 \backslash [-1/16,1/16]^2$ and $\operatorname{supp} V \subset [-1,1]$. So, $1/16 \leq |\eta_1| \leq 1/2$ and $|\eta_2| \leq 1/2$. Therefore, the function $G(\eta) = \left|W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\right|^2 V(\eta_2/\eta_1)$ is still compactly supported and $C_o^\infty$. Additionally, observe $2^{-j}|\ell| \leq 1$, so using Lemma A.1 we have for each $N \in \mathbb{N}$

$$\begin{aligned}
|\beta^{(1)}_{j,\ell,k}| &= 2^{3j/2} \left| \int_a^b \left( \int_{\mathbb{R}^2} L^N(G(\eta)) L^{-N}\left(e^{2\pi\eta\cdot(k_1+2^{2j}f(u)+2^j\ell u, k_2+2^j u)}\right) d\eta \right) g(u)du \right| \\
&\leq 2^{3j/2} \int_a^b (1 + (k_1 + 2^{2j}f(u) + 2^j\ell u)^2)^{-N} (1 + (k_2 + 2^j u)^2)^{-N} \int_{\mathbb{R}^2} \left| L^N(G(\eta)) \right| d\eta |g(u)| du \\
&\leq 2^{3j/2} C_N \int_a^b (1 + (k_1 + 2^{2j}f(u) + 2^j\ell u)^2)^{-N} (1 + (k_2 + 2^j u)^2)^{-N} |g(u)| du. \qquad (28)
\end{aligned}$$

Similarly, for each $N \in \mathbb{N}$ there is a constant $C_N$ such that

$$|\beta^{(2)}_{j,\ell,k}| \leq 2^{3j/2} C_N \int_a^b (1+(k_1+2^j u)^2)^{-N}(1+(k_2+2^{2j}f(u)+2^j\ell u)^2)^{-N}|g(u)|du.$$

For each $j \geq 0$ in $\mathbb{Z}$, we define the set

$$K^{(1)}_j = \{(j,\ell,k) \in M^{(1)}_j : |k_1| \leq 3 \cdot 2^{2j},\, |k_2| \leq 2 \cdot 2^j\}.$$

We observe that, if $|k_2| \geq 2 \cdot 2^j$, then $|k_2 + 2^j u| \geq 2^j$ for all $u \in [a,b]$. Also if $|k_1| \geq 3 \cdot 2^{2j}$, and remembering $|\ell| \leq 2^j$, then $|k_1 + 2^{2j} f(u) + 2^j \ell u| \geq 2^{2j}$ for all $u \in [a,b]$. Then, using inequality (28) we see

$$\begin{aligned}
\sum_{(\ell,k)\in M^{(1)}_j \setminus K^{(1)}_j} |\beta^{(1)}_{j,\ell,k}| &\leq 2^{3j/2} C_N \int_a^b \sum_{|\ell|\leq 2^j} \sum_{|k_1|\geq 3\cdot 2^{2j} \text{ or } |k_2|\geq 2\cdot 2^j} \\
&\quad (1+(k_1+2^{2j}f(u)+2^j\ell u)^2)^{-N}(1+(k_2+2^j u)^2)^{-N}|g(u)|du \\
&\leq 2^{3j/2} C_N \int_a^b \sum_{|\ell|\leq 2^j} \Bigg( \\
&\quad \sum_{|k_1|\geq 3\cdot 2^{2j},\, k_2\in\mathbb{Z}} (1+(k_1+2^{2j}f(u)+2^j\ell u)^2)^{-N}(1+(k_2+2^j u)^2)^{-N} \\
&+ \sum_{k_1\in\mathbb{Z},\, |k_2|\geq 2\cdot 2^j} (1+(k_1+2^{2j}f(u)+2^j\ell u)^2)^{-N}(1+(k_2+2^j u)^2)^{-N} \Bigg) \\
&\times |g(u)|du \\
&\leq C_N\, 2^{\frac{5}{2}j}\, 2^{-(N-1)2j} = C_N\, 2^{\frac{9}{2}j}\, 2^{-2Nj}. \qquad (29)
\end{aligned}$$

Similarly $K^{(2)}_j = \{(j,\ell,k) \in M^{(2)}_j : |k_1| \leq 2 \cdot 2^j,\, |k_2| \leq 3 \cdot 2^{2j}\}$ and, using a very similar argument on $\beta^{(2)}_{j,\ell,k}$, we have that, for any $N \in \mathbb{N}$, there is a constant $C_N$ such that

$$\sum_{(\ell,k)\in M^{(2)}_j \setminus K^{(2)}_j} |\beta^{(2)}_{j,\ell,k}| \leq C_N\, 2^{\frac{7}{2}j}\, 2^{-2Nj}. \qquad (30)$$

Also, we observe that the set $K^{(2)}_j$ contains $O(2^{4j})$ elements. Then, using Lemma 3.1 we see that for each $N \in \mathbb{N}$ there is a constant $C_N$ such that

$$\begin{aligned} \sum_{(\ell,k)\in K_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| &\leq \sum_{(\ell,k)\in K_j^{(2)}} C_N\, 2^{\frac{5}{2}j}\;\; 2^{-2Nj} \\ &\leq C_N\, 2^{4j} 2^{\frac{5}{2}j} 2^{-2Nj} \\ &= C_N\, 2^{\frac{13}{2}j} 2^{-2Nj}. \end{aligned} \tag{31}$$

Using (30) and (31), we have

$$\begin{aligned} \sum_{(k,\ell)\in M_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| &= \Big( \sum_{(\ell,k)\in K_j^{(2)}} + \sum_{(\ell,k)\in M_j^{(2)}\setminus K_j^{(2)}} \Big) |\beta^{(2)}_{j,\ell,k}| \\ &\leq C_N \left( 2^{\frac{7}{2}j}\, 2^{-2Nj} + 2^{\frac{13}{2}j} 2^{-2Nj} \right) \\ &\leq C_N\; 2^{\frac{13}{2}j}\, 2^{-2Nj}. \end{aligned}$$

Choosing $N$ large enough in the above expression, we have

$$\sum_{(\ell,k)\in M_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| = o(2^j), \;\text{ as }\; j \to \infty. \tag{32}$$

We now need to prove $\sum_{(\ell,k)\in S^c_{s,j,1}} |\beta^{(1)}_{j,\ell,k}| = o(2^j)$ and define $S_{s,j,1}$. Now in the integral (27), we make another change of variable $\rho\Theta(\theta) = \rho(\cos(\theta), \sin(\theta)) = 2^{-2j}\xi$ and obtain

$$\begin{aligned} \beta^{(1)}_{j,\ell,k} &= 2^{-3j/2} \int_a^b \left( \int_{\mathbb{R}^2} |W(2^{-2j}\xi)|^2 V\Big(2^j \frac{\xi_2}{\xi_1} - \ell\Big) e^{2\pi i \xi\cdot(A_1^{-j}B_1^{-\ell}k + (f(u),u))} d\xi \right) g(u) du \\ &= 2^{\frac{5}{2}j} \int_a^b \int_0^\infty \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} |W(\rho\Theta(\theta))|^2\, V(2^j \tan\theta - \ell) \\ &\times\; e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot\left(A_{(1)}^{-j}B_{(1)}^{-\ell}k + (f(u),u)\right)} \rho\, d\rho\, d\theta g(u)\, du \\ &= 2^{\frac{5}{2}j} \int_0^\infty \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} |W(\rho\Theta(\theta))|^2\, V(2^j \tan\theta - \ell) e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot A_{(1)}^{-j}B_{(1)}^{-\ell}k} \\ &\times\; \left( \int_a^b e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\, du \right)\, d\theta\, \rho\, d\rho. \end{aligned}$$

As in the proof of Lemma 3.4, by a suitable translation and rotation of the curve segment $S$, we can assume that $f(0) = f'(0) = 0$. Also we may assume that $f''(x) > 0$ so that $f'(x)$ is strictly increasing (the same argument for the case of $f'(x)$ being strictly decreasing). We define

$$\phi(u,\theta) = 2\pi\Theta(\theta)\cdot(f(u),u) = 2\pi(\cos(\theta)f(u) + \sin(\theta)u) = 2\pi\cos(\theta)(f(u) + \tan\theta u)$$

and observe that $\phi'(u) = 2\pi\cos(\theta)(f'(u) + \tan(\theta))$. Again, by a change of parameter, we may assume $a = -\epsilon$ and $b = \epsilon$ . Since $g \in C_0^\infty(-\epsilon,\epsilon)$, one can find $0 < \epsilon_0 < \epsilon$ such that $\operatorname{supp}(g) \subset [-\epsilon_0, \epsilon_0]$. Let $\delta_0 = \frac{1}{2}(\epsilon - \epsilon_0)$ and $\theta_1 = \tan^{-1}(-f'(-(\epsilon_0+\delta_0)))$, $\theta_0 = |\tan^{-1}(-f'(\epsilon_0+\delta_0))|$ so that $\tan(\theta_1) = -f'(-(\epsilon_0+\delta_0))$ and $\tan(-\theta_0) = -f'(\epsilon_0+\delta_0)$. Since $\tan\theta$ is increasing on $[-\frac{\pi}{4},\frac{\pi}{4}]$ with $\tan 0 = 0$ and $f'(u)$ is increasing on $[-\epsilon,\epsilon]$ with $f'(0) = 0$, we see that the interval $[-\theta_0,\theta_1]$ matches the interval $[-(\epsilon_0+\delta_0), \epsilon_0+\delta_0]$. That is, the map

$$\tan^{-1}\circ(-f') : [-(\epsilon_0+\delta_0), \epsilon_0+\delta_0] \to [-\theta_0,\theta_1]$$

is onto and strictly decreasing; therefore bijective. So, for $\theta \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$ or $\theta - \pi \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$ and $|u| \le \epsilon_0$, we have $f'(u) + \tan\theta \ne 0$.

It follows that there exists a constant $c > 0$ such that $|\phi'_u(u,\theta)| \ge c$ for all $\theta \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$ or $\theta - \pi \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$ and $|u| \le \epsilon_0$. So, we have

$$\begin{aligned}\int_{-\epsilon}^{\epsilon} e^{i2^{2j}\phi(u,\theta)} g(u)\,du &= \left( g(u)\frac{e^{i2^{2j}\phi(u,\theta)}}{i2^{2j}\phi'_u(u,\theta)} \right)\Bigg|_{-\epsilon}^{\epsilon} - \int_{-\epsilon}^{\epsilon} g'(u)\frac{e^{i2^{2j}\phi(u,\theta)}}{i2^{2j}\phi'_u(u,\theta)}du \\ &= -2^{-2j}\int_{-\epsilon}^{\epsilon} g'(u)\frac{e^{i2^{2j}\phi(u,\theta)}}{i\phi'_u(u,\theta)}du.\end{aligned}$$

Thus integrating by parts $N$ times gives that for all $\theta \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$ or $\theta - \pi \in [-\frac{\pi}{4},\frac{\pi}{4}] \setminus (-\theta_0,\theta_1)$, we have

$$\left| \int_{-\epsilon}^{\epsilon} e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du \right| \le C_N 2^{-2Nj}.$$

Also as in the proof of lemma 3.4, for $\frac{\pi}{4} \le |\theta| \le \frac{\pi}{2}$ or $\frac{\pi}{4} \le |\theta - \pi| \le \frac{\pi}{2}$ and integrating by parts $N$

times, we have

$$\left|\int_{-\epsilon}^{\epsilon} e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du\right| \le C_N 2^{-2Nj}.$$

So, we see that

$$\begin{aligned}
|\beta^{(1)}_{j,\ell,k}| &\le 2^{\frac{5}{2}j}\bigg|\int_0^\infty\left[\int_{-\theta_0}^{\theta_1}+\int_{\pi-\theta_0}^{\pi+\theta_1}+\right]|W(\rho\Theta(\theta))|^2\, V(2^j\tan\theta-\ell)e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot A_{(1)}^{-j}B_{(1)}^{-\ell}k}\\
&\times \left(\int_a^b e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du\right)\, d\theta\,\rho\,d\rho\bigg| + C_N 2^{-2Nj}.
\end{aligned}$$

So, to find a bound for $|\beta^{(1)}_{j,\ell,k}|$, we may write $\beta^{(1)}_{j,\ell,k}$ as

$$\begin{aligned}
\beta^{(1)}_{j,\ell,k} &= 2^{\frac{5}{2}j}\int_0^\infty\left[\int_{-\theta_0}^{\theta_1}+\int_{\pi-\theta_0}^{\pi+\theta_1}+\right]|W(\rho\Theta(\theta))|^2\, V(2^j\tan\theta-\ell)e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot A_{(1)}^{-j}B_{(1)}^{-\ell}k}\\
&\times \left(\int_a^b e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du\right)\, d\theta\,\rho\,d\rho.
\end{aligned}$$

We also remark that the case for $\theta \in [\pi-\theta_0, \pi+\theta_1]$ is identical for the case $\theta \in [-\theta_0, \theta_1]$; therefore, we may again rewrite $\beta^{(1)}_{j,\ell,k}$ as

$$\begin{aligned}
\beta^{(1)}_{j,\ell,k} &= 2^{\frac{5}{2}j}\int_0^\infty\int_{-\theta_0}^{\theta_1}|W(\rho\Theta(\theta))|^2\, V(2^j\tan\theta-\ell)e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot A_{(1)}^{-j}B_{(1)}^{-\ell}k}\\
&\times \left(\int_a^b e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du\right)\, d\theta\,\rho\,d\rho.
\end{aligned}$$

From the choice of $\theta_0$ and $\theta_1$, we have $\tan^{-1}\circ(-f')$ is bijective. So, for each $\theta \in [-\theta_0, \theta_1]$, there is a unique $u_\theta \in [-(\epsilon_0+\delta_0), (\epsilon_0+\delta_0)]$ such that $\phi'_u(u_\theta, \theta) = 2\pi\cos(\theta)(f'(u_\theta)+\tan(\theta)) = 0$. Now as in the proof of Lemma 3.4, we apply Lemma 3.2 to get

$$\int_{-\epsilon}^{\epsilon} e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot(f(u),u)} g(u)\,du = 2^{-j}\rho^{-\frac{1}{2}}\left(a(u_\theta)\,e^{2\pi i 2^{2j}\rho\phi(u_\theta)} + O(\rho^{-\frac{1}{2}})\right)$$

for $\theta \in [-\theta_0, \theta_1]$ where $a(u_\theta) = \left(\frac{2\pi i}{\phi''_{u^2}(u_\theta,\theta)}\right)^{\frac{1}{2}} g(u_\theta)$.

Now, omitting high order terms in the above expression, we may write $\beta^{(1)}_{j,\ell,k}$ as

$$\begin{aligned}\beta^{(1)}_{j,\ell,k} &= 2^{\frac{3}{2}j}\int_0^\infty\int_{|\theta|\le\theta_0}|W(\rho\Theta(\theta))|^2\,V(2^j\tan\theta-\ell)e^{2\pi i 2^{2j}\rho\Theta(\theta)\cdot A_{(1)}^{-j}B_{(1)}^{-\ell}k}\\ &\quad\times\left(a(u_\theta)\,e^{2\pi i 2^{2j}\rho\phi(u_\theta)}\right)\,d\theta\,\rho^{\frac{1}{2}}\,d\rho.\end{aligned}$$

Remember supp $(V)\in[-1,1]$, so for given $\ell$ and $\theta\in[-\theta_0,\theta_1]$, we must have $|2^j\tan(\theta)-\ell|\le 1$. This is possible when $\tan(\theta)\sim 2^{-j}\ell$. This means that $|2^{-j}\ell|$ needs to be small because $\theta\in[-\theta_0,\theta_1]$. Also, remember $|\ell|\le 2^j$, so we make another change of variable $t=2^j-\ell$ and have $|t|\le 1$. So, we see

$$\tan(\theta(t)) = 2^{-j}(t+\ell)$$

and

$$u_\theta(t) = (f')^{-1}(-2^{-j}(t+\ell)).$$

Observe that $u_\theta(t)$ it is well defined because for large valued of $2^{-j}(t+\ell)=\tan(\theta)$ we have $\theta\notin[-\theta_0,\theta_1]$ which correspond to the values where $\beta^{(1)}_{j,\ell,k}$ is neglected. So, using the last change of variable we have,

$$\beta^{(1)}_{j,\ell,k} = 2^{\frac{3}{2}j}\int_0^\infty\int_{-1}^{1}|W(\rho\Theta(\theta(t))|^2\,V(t)\,a(u_{\theta(t)})\,\frac{e^{2\pi i\rho\,G(t)\cos\theta(t)}}{1+2^{-2j}(t+\ell)^2}\,\rho^{\frac{1}{2}}dt\,d\rho, \tag{33}$$

where $G:[-1,1]\mapsto\mathbb{R}$ is given by

$$G(t) = k_1+tk_2+2^{2j}f\left((f')^{-1}(-2^{-j}(t+\ell)\right)+2^j(t+\ell)\,(f')^{-1}(-2^{-j}(t+\ell)).$$

We observe that $G$ is continuous because $f$ and $(f')^{-1}$ are, also and compactly supported because $|2^{-j}(t+\ell)|\le 2^{-j}(|t|+|\ell|)\le 2^{-j}+1\le 2$. Therefore, we there is $t_{k,\ell}\in[-1,1]$ such that

$$|G(t_{k,\ell})| = \inf_{|t|\le 1}|G(t)|. \tag{34}$$

For $j > 0$ fixed, we define the set

$$S_{s,j,1} = \{(j,\ell,k) \in M_j^{(1)} : |k_1| \le 3 \cdot 2^{2j},\, |k_2| \le 2 \cdot 2^j,\, |G(t_{k,\ell})| \le 2^{\Delta_0 j}\} \tag{35}$$

where $\Delta_0 > 0$. Also remember for $j > 0$ fixed, we defined the sets $K_j^{(1)} = \{(j,\ell,k) \in M_j^{(1)} : |k_1| \le 3 \cdot 2^{2j},\, |k_2| \le 2 \cdot 2^j\}$. Similarly we define

$$Q_j^{(1)} = \{(j,\ell,k) \in M_j^{(1)} : |G(t_{k,\ell})| \le 2^{\Delta_0 j}\}$$

and observe that $S_{s,j,1} = K_j^{(1)} \cap Q_j^{(1)}$. Thus

$$M_j^{(1)} \setminus S_{s,j,1} = (M_j^{(1)} \setminus K_j^{(1)}) \cup (M_j^{(1)} \setminus Q_j^{(1)}) = (M_j^{(1)} \setminus K_j^{(1)}) \cup \Big((M_j^{(1)} \setminus Q_j^{(1)}) \cap K_j^{(1)}\Big).$$

Therefore, we may write the first sum in (24) as

$$\sum_{(\ell,k)\in M_j^{(1)}\setminus S_{s,j,1}} |\beta_{j,\ell,k}^{(1)}| = \sum_{(\ell,k)\in M_j^{(1)}\setminus K_j^{(1)}} |\beta_{j,\ell,k}^{(1)}| + \sum_{(\ell,k)\in (M_j^{(1)}\setminus Q_j^{(1)})\cap K_j} |\beta_{j,\ell,k}^{(1)}|. \tag{36}$$

From inequality (29) we have that, for every $N \in \mathbb{N}$, there is a constant $C_N > 0$ such that $\sum_{(\ell,k)\in (M_j^{(1)}\setminus K_j^{(1)})} |\beta_{j,\ell,k}^{(1)}| \le C_N\, 2^{\frac{9}{2}j}\, 2^{-2Nj}$. Therefore, choosing $N$ large enough in the last expression, we have

$$\sum_{(\ell,k)\in (M_j^{(1)}\setminus K_j^{(1)})} |\beta_{j,\ell,k}^{(1)}| = o(2^j). \tag{37}$$

For the second sum in (36), we observe that, for $(\ell,k) \in (M_j^{(1)} \setminus Q_j^{(1)}) \cap K_j^{(1)}$, we have $|G(t)| \ge 2^{\Delta_0 j}$ for all $t \in [-1,1]$. So, using (33) we see

$$\begin{aligned}
\beta_{j,\ell,k}^{(1)} &= 2^{\frac{3}{2}j} \int_0^\infty \int_{-1}^1 |W(\rho\Theta(\theta(t))|^2\, V(t)\, a(u_{\theta(t)})\, \frac{e^{2\pi i\rho\, G(t)\cos\theta(t)}}{1+2^{-2j}(t+\ell)^2}\, \rho^{\frac{1}{2}}\, dt\, d\rho \\
&= 2^{\frac{3}{2}j} \int_{-1}^1 V(t)\, a(u_{\theta(t)})\, \frac{1}{1+2^{-2j}(t+\ell)^2} \int_0^\infty |W(\rho\Theta(\theta(t))|^2\, \rho^{\frac{1}{2}} e^{2\pi i\rho\, G(t)\cos\theta(t)} d\rho\, dt.
\end{aligned}$$

Since $\mathrm{supp}\,(W) \in [-1/2,1/2]^2 \setminus [-1/16,1/16]^2$, we integrate by parts and have

$$\begin{aligned}\int_0^\infty |W(\rho\Theta(\theta(t))|^2\, \rho^{\frac{1}{2}} e^{2\pi i \rho\, G(t)\cos\theta(t)} d\rho &= \left( |W(\rho\Theta(\theta(t))|^2 \rho^{\frac{1}{2}} \frac{e^{2\pi i\rho\, G(t)\cos\theta(t)}}{2\pi i G(t)\cos\theta(t)} \right)\Big|_0^\infty \\ &- \frac{1}{2\pi i G(t)\cos\theta(t)} \\ &\times \int_0^\infty \frac{\partial}{\partial\rho} \left( |W(\rho\Theta(\theta(t))|^2\, \rho^{\frac{1}{2}} \right) e^{2\pi i\rho\, G(t)\cos\theta(t)} d\rho\end{aligned}$$

By repeated integration by parts with respect to the variable $\rho$ in the integral of $\beta^{(1)}_{j,\ell,k}$, we have that, for any $N \in \mathbb{N}$, there is a constant $C_N$ such that

$$|\beta^{(1)}_{j,\ell,k}| \le C_N\, 2^{\frac{3}{2}j} \int_{-1}^{1} |V(t)|\, \frac{1}{|G(t)\cos\theta(t)|^N} \frac{dt}{1+2^{-2j}(t+\ell)^2}$$

Hence, for $(\ell,k) \in (M_j^{(1)} \setminus Q_j^{(1)}) \cap K_j^{(1)}$ and any $N \in \mathbb{N}$, there is a constant $C_N$ such that

$$|\beta^{(1)}_{j,\ell,k}| \le C_N\, 2^{\frac{3}{2}j}\, 2^{-N\Delta_o j}.$$

Therefore, observing that the cardinality of $K_j^{(1)}$ is of order $2^{4j}$, we have that

$$\sum_{(\ell,k)\in(M_j^{(1)}\setminus Q_j)\cap K_j} |\beta^{(1)}_{j,\ell,k}| \le C_N\, 2^{4j}\, 2^{\frac{3}{2}j}\, 2^{-N\Delta_o j}.$$

If we choose $N$ large enough, we have that $N\Delta_0 > \frac{11}{2}$ so that the sum in the last expression is $o(2^j)$. Combining this last estimate with (37) in (36), and then using the estimate (32), we have that

$$\delta_j^s = \sum_{(\ell,k)\in M_j^{(1)}\setminus S_{s,j,1}} |\beta^{(1)}_{j,\ell,k}| + \sum_{(\ell,k)\in M_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| = o(2^j). \quad \square$$

### 5.1.2 Proof of Proposition 5.2

In order to prove Proposition 5.2, we need the following result which will be proved later.

**Lemma 5.3.** *For $j > 0$ fixed and $k \in \mathbb{Z}^2$, $\ell \in \mathbb{Z}$, let $t_{k,\ell}$ be defined by equation (34) in the proof of Proposition 5.1. Set*

$$G_{k,\ell} = k_1 + t_{k,\ell}k_2 + 2^{2j} f[(f')^{-1}(-2^{-j}(t_{k,\ell}+\ell)] + 2^j (t_{k,\ell}+\ell)\,(f')^{-1}(-2^{-j}(t_{k,\ell}+\ell))$$

*and $Q_k = \{|\ell| \le 2^j : G_{k,\ell} \le 2^{\Delta_0 j}\}$. Then for each fixed $k$, the cardinality of the set $Q_k$ satisfies the inequality $\#(Q_k) \le C\, 2^{\frac{1}{2}\Delta_0 j}$, where the constant $C$ is independent of $j, k$.*

*Proof of Proposition 5.2*

We begin proving (25) using the definition of $S_{s,j,1}$ given in (35). As mentioned in Remark 2.3, for a technical reason, we use $P_{\mathcal{M}_{h_j}} = h_j^{\Delta_0} \mathbb{1}_{|x_2| \le h_j}$. This does not affect the $P_{\mathcal{M}_{h_j}}$ significantly because for small values of $h_j$ we have $h_j^{\Delta_0} \to 1$ as $\Delta_0 \to 0$.

Like in the proof of Proposition 4.2, we use Plancherel Theorem and Lemma A.2. So, we have

$$h_j^{\Delta_0} \widehat{\mathbb{1}_{\mathcal{M}_{h_j}}} * \widehat{\psi^{(1)}_{j,\ell,k}} = 2h_j^{1+\Delta_0} \int_{\widehat{\mathbb{R}}} \operatorname{sinc}(2\pi h_j \tau_2) \widehat{\psi^{(1)}_{j,\ell,k}}\big((\xi_1,\xi_2) - (0,\tau_2)\big)\, d\tau_2$$

$$\begin{aligned}
& \langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle \\
= \; & \langle h_j^{\Delta_0} \widehat{\mathbb{1}_{\mathcal{M}_{h_j}}} * \widehat{\psi^{(1)}_{j,\ell,k}}, \widehat{\psi^{(1)}_{j,\ell',k'}} \rangle \\
= \; & 2h_j^{1+\Delta_0} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} \operatorname{sinc}(2\pi h_j \tau_2)\, \widehat{\psi^{(1)}_{j,\ell,k}}\big((\xi_1,\xi_2) - (0,\tau_2)\big)\, d\tau_2\, \overline{\widehat{\psi^{(1)}_{j,\ell',k'}}}(\xi)\, d\xi \\
= \; & 2h_j^{1+\Delta_0} 2^{-3j} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} W(2^{-2j}(\xi_1, \xi_2 - \tau_2))\, V(2^j \frac{\xi_2 - \tau_2}{\xi_1} - \ell) \operatorname{sinc}(2\pi h_j \tau_2) \\
\times \; & e^{-2\pi i (0,\tau_2)\cdot A_{(1)}^{-j} B_{(1)}^{-\ell} k} d\tau_2\, \overline{W}(2^{-2j}\xi)\, \overline{V}(2^j \frac{\xi_2}{\xi_1} - \ell')\, e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} (k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k')} d\xi.
\end{aligned}$$

We now make a change of variable $\eta = \xi A_{(1)}^{-j} B_{(1)}^{-\ell}$ so that $\xi = \eta B_1^\ell A_1^j = (2^{2j}\eta_1, 2^j(\ell\eta_1 + \eta_2))$ and $d\xi_1 = 2^{2j} d\eta_1$ and $d\xi_2 = 2^j d\eta_2$. Thus,

$$\langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle$$

$$
\begin{aligned}
&= 2h_j^{1+\Delta_0} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2) - 2^{-2j}\tau_2) V(\frac{\eta_2}{\eta_1} - \frac{2^{-j}\tau_2}{\eta_1}) e^{-2\pi i 2^{-j}\tau_2 k_2} \\
&\times \ \mathrm{sinc}(2\pi\tau_2 2^{-j})\, d\tau_2\, \overline{W}(\eta_1, 2^{-j}(\ell'\eta_1+\eta_2))\, \overline{V}(\frac{\eta_2}{\eta_1})\, e^{2\pi i \eta (k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k')}\, d\eta_1 d\eta_2.
\end{aligned}
$$

We now let $\gamma = 2^{-j}\tau_2$ and have

$$
\begin{aligned}
&\langle P_{\mathcal{M}_{h_j}} \psi_{j,\ell,k}^{(1)}, \psi_{j,\ell',k'}^{(1)} \rangle \\
&= 2h_j^{1+\Delta_0} 2^j \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} W\left(\eta_1, 2^{-j}(\ell\eta_1+\eta_2) - 2^{-j}\gamma\right) V\left(\frac{\eta_2}{\eta_1} - \frac{\gamma}{\eta_1} + \ell - \ell'\right) \\
&\times \ \overline{W}(\eta_1, 2^{-j}(\ell\eta_1+\eta_2)) \overline{V}\left(\frac{\eta_2}{\eta_1}\right) \\
&\times \ e^{-2\pi i \gamma k_2} \mathrm{sinc}(2\pi\gamma)\, d\gamma\, e^{2\pi i \eta (k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k')}\, d\eta.
\end{aligned} \tag{38}
$$

Let

$$
g_{j,\ell,\ell'}(\eta,\gamma) = W\left(\eta_1, 2^{-j}(\ell\eta_1+\eta_2) - 2^{-j}\gamma\right) V\left(\frac{\eta_2}{\eta_1} - \frac{\gamma}{\eta_1} + \ell - \ell'\right) \overline{W}(\eta_1, 2^{-j}(\ell\eta_1+\eta_2)) \overline{V}\left(\frac{\eta_2}{\eta_1}\right).
$$

We compute the following

$$
\begin{aligned}
\frac{\partial}{\partial\eta_1} W &= \partial_1 W + \partial_2 W 2^{-j}\ell \\
\frac{\partial^2}{\partial\eta_1^2} W &= \partial_1\partial_1 W + \partial_2\partial_1 W 2^{-j}\ell + \partial_1\partial_2 W 2^{-j}\ell + \partial_2\partial_2 W 2^{-j}\ell 2^{-j}\ell \\
\frac{\partial}{\partial\eta_2} W &= \partial_2 W 2^{-j} \\
\frac{\partial^2}{\partial\eta_2^2} W &= \partial_2\partial_2 W 2^{-j} 2^{-j} \\
\frac{\partial}{\partial\eta_1} V &= V' \frac{-(\eta_2-\gamma)}{\eta_1^2} \\
\frac{\partial^2}{\partial\eta_1^2} V &= V'' \frac{(\eta_2-\gamma)^2}{\eta_1^4} + V' \frac{2(\eta_2-\gamma)}{\eta_1^3} \\
\frac{\partial}{\partial\eta_2} V &= V' \frac{1}{\eta_1} \\
\frac{\partial^2}{\partial\eta_2^2} V &= V'' \frac{1}{\eta_1^2}
\end{aligned} \tag{39}
$$

And very similar for $\overline{W}$ and $\overline{V}$. From the computations in (39) and since $V$ and $W$ are compactly supported and smooth, it follows that there is a constant $C$ independent of $j, \ell, \ell'$ such that $|L(g_{j,\ell,\ell'}(\eta,\gamma))| \le C$ where $L$ is the operator defined in (13). Also, observe that $(k - B^{\ell}_{(1)} B^{-\ell'}_{(1)} k') = (k_1 - k_1' - (\ell-\ell')k_2', k_2 - k_2')$. Therefore, using the above observations, Lemma A.1 and (38) we have

$$
\begin{aligned}
& |\langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle| \\
=\ & 2h_j^{1+\Delta_0} 2^j \left| \int_{\widehat{\mathbb{R}}^2} L \left( \int_{\widehat{\mathbb{R}}} g_{j,\ell,\ell'}(\eta,\gamma) e^{-2\pi i \gamma k_2} \operatorname{sinc}(2\pi\gamma)\, d\gamma \right) L^{-1} \left( e^{2\pi i \eta (k - B^{\ell}_{(1)} B^{-\ell'}_{(1)} k')} \right) d\eta \right| \\
=\ & 2h_j^{1+\Delta_0} 2^j \left| \int_{\widehat{\mathbb{R}}^2} L \left( \int_{\widehat{\mathbb{R}}} g_{j,\ell,\ell'}(\eta,\gamma) e^{-2\pi i \gamma k_2} \operatorname{sinc}(2\pi\gamma)\, d\gamma \right) \right. \\
\times\ & \left. (1 + (k_1 - k_1' - (\ell-\ell')k_2')^2)^{-2} (1 + (k_2 - k_2')^2)^{-2} e^{2\pi i \eta (k - B^{\ell}_{(1)} B^{-\ell'}_{(1)} k')}\, d\eta \right| \\
\le\ & 2h_j^{1+\Delta_0} 2^j \left| \int_{\widehat{\mathbb{R}}^2} L \left( \int_{\widehat{\mathbb{R}}} g_{j,\ell,\ell'}(\eta,\gamma) e^{-2\pi i \gamma k_2} \operatorname{sinc}(2\pi\gamma)\, d\gamma \right) \right. \\
\times\ & \left. (1 + (k_1 - k_1' - (\ell-\ell')k_2')^2)^{-1} (1 + (k_2 - k_2')^2)^{-1}\, e^{2\pi i \eta (k - B^{\ell}_{(1)} B^{-\ell'}_{(1)} k')}\, d\eta \right| \\
\le\ & 2h_j^{1+\Delta_0} 2^j C (1 + (k_1 - k_1' - (\ell-\ell')k_2')^2)^{-1} (1 + (k_2 - k_2')^2)^{-1}
\end{aligned}
$$

and using the definition of $S_{s,j,1}$ given by (35) with Lemma 5.3 we have

$$
\begin{aligned}
& \sum_{(k,\ell) \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle| \\
& \le C\, h_j^{1+\Delta_0} 2^j \sum_{|k_1| \le 3 2^{2j}} \sum_{|k_2| \le 2 2^j} \sum_{\ell \in Q_k} \left( 1 + |(k_1 - k_1' - (\ell-\ell')k_2'|^2 \right)^{-1} \left( 1 + |k_2 - k_2'|^2 \right)^{-1} \\
& \le C\, h_j^{1+\Delta_0} 2^j\, 2^{\frac{1}{2}\Delta_0 j} \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \left( 1 + |k_1|^2 \right)^{-1} \left( 1 + |k_2|^2 \right)^{-1} \\
& \le C\, h_j^{1+\Delta_0} 2^j\, 2^{\frac{1}{2}\Delta_0 j}.
\end{aligned}
$$

Hence, since $h_j = o(2^{-j})$, it follows that

$$
\max_{\ell',k'} \sum_{\ell,k \in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle| \to 0 \ \text{ as } \ j \to \infty.
$$

This proves (25).

To prove (26), similarly to the computation above, we apply Plancherel Theorem and the Fourier transform of $\mathbb{1}_{\mathcal{M}_{h_j}}$ to write

$$
\begin{aligned}
&\langle P_{\mathcal{M}_{h_j}} \psi^{(2)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle \\
&= \langle \psi^{(2)}_{j,\ell,k}, P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell',k'} \rangle \\
&= \langle \widehat{\psi^{(2)}_{j,\ell,k}}, h_j^{\Delta_0} \widehat{\mathbb{1}_{\mathcal{M}_{h_j}}} * \widehat{\psi^{(1)}_{j,\ell',k'}} \rangle \\
&= 2h_j^{1+\Delta_0} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} \operatorname{sinc}(2\pi h_j \tau_2)\, \overline{\widehat{\psi^{(1)}_{j,\ell',k'}}} \big((\xi_1,\xi_2)-(0,\tau_2)\big)\, d\tau_2\, \widehat{\psi^{(2)}_{j,\ell,k}}(\xi)\, d\xi \\
&= 2h_j^{1+\Delta_0} 2^{-3j} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} \overline{W}(2^{-2j}(\xi_1,\xi_2-\tau_2))\, \overline{V}\left(2^j \frac{\xi_2-\tau_2}{\xi_1} - \ell\right) e^{2\pi i (0,\tau_2) A_{(1)}^{-j} B_{(1)}^{-\ell} k} \\
&\times \operatorname{sinc}(2\pi h_j \tau_2) d\tau_2\, W(2^{-2j}\xi)\, V\left(\frac{2^j \xi_1}{\xi_2} - \ell'\right) e^{-2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} (k - B_{(1)}^{\ell} A_{(1)}^{j} A_{(2)}^{-j} B_{(2)}^{-\ell'} k')} d\xi.
\end{aligned}
$$

We now apply a change of variable $\eta = \xi A_{(1)}^{-j} B_{(1)}^{-\ell} = (2^{-2j}\xi_1, -\ell 2^{-2j}\xi_1 + 2^{-j}\xi_2)$, so that $\xi = \eta B_{(1)}^{\ell} A_{(1)}^{j} = (2^{2j}\eta_1, 2^j(\ell\eta_1 + \eta_2))$, and let $\alpha = (\alpha_1,\alpha_2) = B_{(1)}^{\ell} A_{(1)}^{j} A_{(2)}^{-j} B_{(2)}^{-\ell'} k'$. Thus

$$
\begin{aligned}
&\langle P_{\mathcal{M}_{h_j}} \psi^{(2)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle \\
&= 2h_j^{1+\Delta_0} \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} \overline{W}(\eta_1, 2^{-j}(\ell\eta_1+\eta_2)) - 2^{-2j}\tau_2) \overline{V}\left(\frac{\eta_2 - 2^{-j}\tau_2}{\eta_1}\right) \operatorname{sinc}(2\pi h_j \tau_2) \\
&\times\ e^{2\pi i 2^{-j}\tau_2 k_2} d\tau_2\, W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V\left(\frac{2^{2j}\eta_1}{\ell\eta_1+\eta_2} - \ell'\right) e^{-2\pi i \eta (k-\alpha)}\, d\eta.
\end{aligned}
$$

Similar to the calculation above, letting $\gamma = 2^{-j}\tau_2$ and then applying Lemma A.1, where $L$ is the differential operator (13), we have that

$$
\langle P_{\mathcal{M}_{h_j}} \psi^{(2)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle = 2h_j^{1+\Delta_0} 2^j \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} \tilde{g}_{j,\ell,\ell'}(\eta,\gamma) \operatorname{sinc}(2\pi h_j 2^{-j}\gamma)\, e^{2\pi i \gamma k_2} d\gamma\, e^{-2\pi i \eta(k-\alpha)} d\eta \quad (40)
$$

where

$$
\tilde{g}_{j,\ell,\ell'}(\eta,\gamma) = \overline{W}(\eta_1, 2^{-j}(\ell\eta_1+\eta_2-\gamma)) \overline{V}(\tfrac{\eta_2-\gamma}{\eta_1}) W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{2^{2j}\eta_1}{\ell\eta_1+\eta_2} - \ell').
$$

We also compute

$$
\begin{aligned}
\frac{\partial}{\partial\eta_1}\overline{V} &= \overline{V}' \cdot \frac{2^{2j}\eta_2}{(\ell\eta_1+\eta_2)^2} \\
\frac{\partial^2}{\partial\eta_1^2}\overline{V} &= \overline{V}'' \cdot \frac{2^{4j}\eta_2^2}{(\ell\eta_1+\eta_2)^4} + \overline{V}' \cdot \frac{-2\cdot 2^{3j}\eta_2}{(\ell\eta_1+\eta_2)^3} \cdot \ell 2^{-j} \\
\frac{\partial}{\partial\eta_2}\overline{V} &= \overline{V}' \cdot \frac{-2^{2j}\eta_1}{(\ell\eta_1+\eta_2)^2} \\
\frac{\partial^2}{\partial\eta_2^2}\overline{V} &= \overline{V}'' \cdot \frac{2^{4j}\eta_1^2}{(\ell\eta_1+\eta_2)^4} + \overline{V}' \cdot \frac{-2\cdot 2^{3j}\eta_1}{(\ell\eta_1+\eta_2)^3} 2^{-j}.
\end{aligned}
$$

Since $\operatorname{supp} W \subset [-1/2,1/2]^2 \setminus [-1/16,1/16]^2$, we see that $1/16 \leq 2^{-j}|\ell\eta_1+\eta_2| \leq 1/2$, so $|\ell\eta_1+\eta_2| \sim 2^j$. Therefore, with the above computations and similar computations performed in (39) for $W$, $\overline{W}$ and $V$ and the fact they are smooth, it follows that there is a constant $C$ independent of $j,\ell,\ell'$ such that $|L(g_{j,\ell,\ell'}(\eta,\gamma))| \leq C$ where $L$ is the operator given in (13).

Thus applying Lemma A.1 in (40) we have

$$
\begin{aligned}
|\langle P_{\mathcal{M}_{h_j}}\psi^{(2)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'}\rangle| &= 2h_j^{1+\Delta_0}2^j \bigg| \int_{\widehat{\mathbb{R}}^2} L\bigg( \int_{\widehat{\mathbb{R}}} \tilde{g}_{j,\ell,\ell'}(\eta,\gamma)\,\mathrm{sinc}(2\pi h_j 2^{-j}\gamma) \\
&\times\ e^{2\pi i\gamma k_2} d\gamma \bigg) L^{-1}\left(e^{-2\pi i\eta(k-\alpha)}\right) d\eta \bigg| \\
&\leq\ C\, h_j^{1+\Delta_o}\, 2^j\, (1+(k_1-\alpha_1)^2)^{-1}(1+(k_2-\alpha_2)^2)^{-1}
\end{aligned}
$$

where the indices $\alpha_1,\alpha_2$ depend on $\ell$. Using the definition of $S_{s,j,1}$, given by (35), and next applying Lemma 5.3 to estimate the cardinality of $Q_k$, we have

$$
\begin{aligned}
&\sum_{(k,\ell)\in S_{s,j,1}} |\langle P_{M_{h_j}}\psi^{(2)}_{j,\ell',k'}, \psi^{(1)}_{j,\ell,k}\rangle| \\
&\leq C\, h_j^{1+\Delta_0}2^j \sum_{|k_1|\leq 3 2^{2j}} \sum_{|k_2|\leq 2 2^j} \sum_{\ell\in Q_k} (1+(k_1-\alpha_1)^2)^{-1}(1+(k_2-\alpha_2)^2)^{-1} \\
&\leq C\, h_j^{1+\Delta_0}2^j\, 2^{\frac{1}{2}\Delta_0 j} \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} \left(1+|k_1|^2\right)^{-1} \left(1+|k_2|^2\right)^{-1} \\
&\leq C\, h_j^{1+\Delta_0}2^j\, 2^{\frac{1}{2}\Delta_0 j}.
\end{aligned}
$$

Since $h_j = o(2^{-j})$, it follows that

$$\max_{\ell',k'} \sum_{\ell,k \in S_{s,j,1}} |\langle P_{M_{h_j}} \psi^{(2)}_{j,\ell',k'}, \psi^{(1)}_{j,\ell,k}\rangle| \to 0 \;\text{ as }\; j \to \infty. \qquad \square$$

We finally prove Lemma 5.3.

*Proof of Lemma 5.3.* Letting $y = (f')^{-1}(-2^j(t_{k,\ell} + \ell))$, we can write

$$G(t_{k,\ell}) = (k_1 + t_{k,\ell} k_2) + 2^{2j}\big(f(y) - f'(y)y\big).$$

Recalling that $f(0) = f'(0) = 0$, we have that the second order Taylor expansion of $f$ about 0 on $[-\epsilon, \epsilon]$ is $f(y) = f''(c)\frac{y^2}{2}$ where $c \in (-\epsilon, \epsilon)$ and $f'(y) = f''(c)\, y$. Since $f''(y) > k > 0$ on $[-\epsilon, \epsilon]$ , then

$$f(y) - f'(y)y = -\frac{1}{2} f''(c)\, y^2 \leq 0.$$

Neglecting the higher order terms, we have

$$|G(t_{k,\ell})| = |(k_1 + t_{k,\ell} k_2) + 2^{2j}\big(f(y) - f'(y)y\big)| \simeq |(k_1 + t_{k,\ell} k_2) - 2^{2j}\tfrac{1}{2} f''(c) y^2|. \tag{41}$$

We consider three cases below and recall that $|G(t_{k,\ell})| \leq 2^{\Delta_o j}$ by definition of $Q_k$.

**Case 1:** $k_1 + t_{k,y} k_2 < 0$. It follows that $| - 2^{2j}\frac{1}{2} f''(c) y^2| \leq 2^{\Delta_o j}$. This implies that

$$|t_{k,\ell} + \ell| = 2^j |f'(y)| \simeq 2^j f''(c)|y| \leq \sqrt{2}\sqrt{f''(c)} \cdot 2^{\Delta_0 j/2}$$

and

$$|\ell| \lesssim \sqrt{2}\sqrt{f''(c)} \cdot 2^{\frac{1}{2}\Delta_0 j} + |t_{k,\ell}| \leq \sqrt{2}\sqrt{f''(c)} \cdot 2^{\frac{1}{2}\Delta_0 j} + 1.$$

So, $\ell$ is contained on an interval of length less or equal than $2 \cdot (\sqrt{2}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1)$. And also remember that $\ell \in \mathbb{Z}$ so that the there are at most $2 \cdot (\sqrt{2}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1)$ of such elements $\ell$. Hence $\#|Q_{k_1,k_2}| \leq 2 \cdot (\sqrt{2}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1)$. So, there is a constant $C$ independent of $j, k_1, k_2$

such that $\#(Q_k) \leq C\, 2^{\frac{1}{2}\Delta_0 j}$.

**Case 2:** $0 \leq k_1 + t_{k,y} k_2 \leq 2^{\Delta_0 j+1}$. Then (41) implies that

$$2^{\Delta_o j} \gtrsim |(k_1 + t_{k,\ell}k_2) - 2^{2j}\frac{1}{2}f''(c)y^2| \geq |2^{2j}\frac{1}{2}f''(c)y^2| - |k_1 + t_{k,\ell}k_2|.$$

Therefore, in this case, we have that

$$|2^{2j}\frac{1}{2}f''(c)y^2| \lesssim |k_1 + t_{k,\ell}k_2| + 2^{\Delta_o j} \leq 2^{\Delta_o j+1} + 2^{\Delta_o j} = 3 \cdot 2^{\Delta_o j}.$$

Similar to case 1, it follows that

$$|t_{k,\ell} + \ell| = 2^j |f'(y)| \simeq 2^j f''(c)|y| \leq \sqrt{6}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2}$$

and

$$|\ell| \lesssim \sqrt{6}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + |t_{k,\ell}| \leq \sqrt{6}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1.$$

As in case 1, $\ell$ is contained on interval of length less or equal to $2 \cdot (\sqrt{6}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1)$. And also remember that $\ell \in \mathbb{Z}$ so that the there are at most $2 \cdot (\sqrt{6}\sqrt{f''(c)} \cdot 2^{\Delta_o j/2} + 1)$ of such elements $\ell$. So, there is a constant $C$ independent of $j, k_1, k_2$ such that $\#(Q_k) \leq C\, 2^{\frac{1}{2}\Delta_0 j}$.

**Case 3:** $k_1 + t_{k,y} k_2 \geq 2^{\Delta_0 j+1}$. Again (41) implies that

$$k_1 + t_{k,\ell}k_2 - 2^{\Delta_0 j} \lesssim 2^{2j}\,\frac{1}{2}f''(c)\,y^2 \lesssim k_1 + t_{k,\ell}k_2 + 2^{\Delta_0 j}$$

and, thus,

$$2^{-j}\frac{\sqrt{2}}{\sqrt{f''(c)}}\sqrt{k_1 + t_{k,\ell}k_2 - 2^{\Delta_o j}} \lesssim |y| \lesssim 2^{-j}\frac{\sqrt{2}}{\sqrt{f''(c)}}\sqrt{k_1 + t_{k,\ell}k_2 + 2^{\Delta_o j}}.$$

This shows that $|y|$ is contained in the interval

$$I_y = \left[2^{-j}\frac{\sqrt{2}}{\sqrt{f''(c)}}\sqrt{k_1 + t_{k,\ell}k_2 - 2^{\Delta_o j}}, 2^{-j}\frac{\sqrt{2}}{\sqrt{f''(c)}}\sqrt{k_1 + t_{k,\ell}k_2 + 2^{\Delta_o j}}\right],$$

whose length satisfies the inequality

$$|I_y| = \frac{\sqrt{2}\,2^{-j}}{\sqrt{f''(c)}}\left(\sqrt{k_1 + t_{k,\ell}k_2 + 2^{\Delta_o j}} - \sqrt{k_1 + t_{k,\ell}k_2 - 2^{\Delta_o j}}\right) \leq \frac{\sqrt{2}\,2^{-j+1}}{\sqrt{f''(c)}}\,2^{\frac{1}{2}\Delta_o}.$$

Let $m = |\ell + t_{k,\ell}|$ so that $m = 2^j|f'(y)| \simeq 2^j f''(c)|y|$. Since the map $x \mapsto f''(c)x$ is continuous, then the expression above maps the interval $I_y$ to some other interval $I_m$. For any $m_1, m_2 \in I_m$, we have that

$$|m_2 - m_1| \simeq 2^j f''(c)\,||y_2| - |y_1|| \leq 2\sqrt{2}\sqrt{f''(c)}\,2^{\frac{1}{2}\Delta_o},$$

that is, the length of $I_m$ satisfies $|I_m| \leq 2\sqrt{2}\sqrt{f''(c)}\,2^{\frac{1}{2}\Delta_o}$. From $|\ell + t_{k,\ell}| = m \in I_m$, we have $|\ell| \in I_m \pm t_{k,\ell}$. Since $|t_{k,\ell}| \leq 1$, as in Cases 1 and 2, there at most $2\sqrt{2}\sqrt{f''(c)}\,2^{\frac{1}{2}\Delta_o} + 2$ of those $\ell$ elements in $I_m$. So, there is a constant $C$ independent of $j, k_1, k_2$ such that $\#(Q_k) \leq C\,2^{\frac{1}{2}\Delta_0 j}$. □

## 5.2 Proof of Theorem 2.8 (Thresholding)

For $\nu = 1, 2$, let $\gamma^{(\nu)}_{j,\ell,k} = \langle \psi^{(\nu)}_{j,\ell,k}, P_{\mathcal{M}_{h_j}}\mathcal{T}_j\rangle$ and $\beta^{(\nu)}_{j,\ell,k} = \langle \psi^{(\nu)}_{j,\ell,k}, \mathcal{T}_j\rangle$

Since $\mathcal{T}_j$ is related to a local vertical curve, as for the $\ell_1$ minimization case, we need only consider the case $\nu = 1$. In the following, we simply denote $\gamma^{(1)}_{j,\ell,k}$ as $\gamma_{j,\ell,k}$, $\beta^{(1)}_{j,\ell,k}$ as $\beta_{j,\ell,k}$ and set $\alpha_{j,\ell,k} = \beta_{j,\ell,k} - \gamma_{j,\ell,k}$.

For any $j \geq 0$ and any $0 \leq \sigma_j \leq 2^{-4j}$, we let $I_j = \{(\ell, k) : \ |\alpha_{j,\ell,k}| \geq \sigma_j\}$ and $\delta^s_j = \sum_{k \in I^c_j} |\beta_{j,\ell,k}|$,.

We recall that $R^\tau_j = F[\mathbb{1}_{I_j} F^*\mathcal{T}_j]$ and observe that $\|\mathbb{1}_{I_j}\Psi^* P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 = \sum_{(\ell,k)\in I_j} |\gamma_{j,k}|$. Lemma 3.13 then implies the following estimate.

**Proposition 5.4.** *For any $j \in \mathbb{Z}$, let $R^\tau_j$, $I_j$ and $\delta^s_j$ be defined as above. Then there is a constant*

*$C$ independent of $j$ and $\mathcal{T}$ such that*

$$\|R_j^\tau - \mathcal{T}_j\|_2 \leq C(\delta_j^s + \|\mathbb{1}_{I_j} F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1).$$

A simple observation shows that, for any $j \in \mathbb{Z}$,

$$\|\mathbb{1}_{I_j} F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1 \leq \|F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1 = \sum_{(\ell,k)\in M_j} |\gamma_{j,\ell,k}|.$$

So, Theorem 2.8 follows from Proposition 5.4 and the following result.

**Proposition 5.5.** *Let $j \geq 0$. For any $0 \leq \sigma_j \leq 2^{-4j}$ and $h_j = o(2^{-\frac{3}{4}j})$, we have*

$$\sum_{(\ell,k)\in M_j} |\gamma_{j,\ell,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2) \tag{42}$$

$$\sum_{(\ell,k)\in I_j^c} |\beta_{j,\ell,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2), \quad \text{as } j \to \infty \tag{43}$$

### 5.2.1 Proof of Proposition 5.5

Making a change of variable $\eta = \xi A_{(1)}^{-j} B_{(1)}^{-\ell}$ and using the expression for $P_{\mathcal{M}_{h_j}}$ computed in Proposition 4.4, we have

$$\begin{aligned}
\gamma_{j,\ell,k} &= \langle \hat{\psi}_{j,\ell,k}^{(\nu)}, \widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j} \rangle \\
&= 2^{-\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi)\, V\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{-\ell} k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(\xi)}\, d\xi \\
&= 2^{\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i \eta k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(\eta\, B_{(1)}^{\ell} A_{(1)}^{j})}\, d\eta \\
&= 2^{\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i \eta k} \\
&\times \int_{\mathbb{R}^2} \left( \int_a^b \mathbb{1}_{\mathcal{M}_{h_j}} \big(x + (f(u), u)\big)\, e^{2\pi i \eta\, B_{(1)}^{\ell} A_{(1)}^{j}\, (x + (f(u),u))} g(u) du \right) 2^{4j} \overline{\check{W}}(2^{2j}x) dx\, d\eta \\
&= 2^{\frac{11}{2}j} \int_{\mathbb{R}^2} \int_a^b \left( \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i \eta \left(k + B_{(1)}^{\ell} A_{(1)}^{j}\, (x + (f(u),u))\right)}\, d\eta \right) \\
&\times\ \mathbb{1}_{\mathcal{M}_{h_j}} \big(x + (f(u), u)\big)\, g(u)\, du\, \overline{\check{W}}(2^{2j}x) dx.
\end{aligned}$$

In a similar way as in Proposition 4.4 we write $\gamma_{j,\ell,k} = \gamma^{(1)}_{j,\ell,k} + \gamma^{(2)}_{j,\ell,k}$ where

$$\begin{aligned}\gamma^{(1)}_{j,\ell,k} &= 2^{\frac{11}{2}j} \int_{B_{\Delta_0}} \int_a^b \left( \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i\eta\left(k + B^{\ell}_{(1)} A^{j}_{(1)}\,(x + (f(u),u))\right)}\, d\eta \right) \\ &\times\ \mathbb{1}_{\mathcal{M}_{h_j}}\big(x + (f(u),u)\big)\, g(u)\, du\, \overline{\check{W}}(2^{2j}x) dx, \end{aligned} \tag{44}$$

$$\begin{aligned}\gamma^{(2)}_{j,\ell,k} &= 2^{\frac{11}{2}j} \int_{\mathbb{R}\setminus B_{\Delta_0}} \int_a^b \left( \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i\eta\left(k + B^{\ell}_{(1)} A^{j}_{(1)}\,(x + (f(u),u))\right)}\, d\eta \right) \\ &\times\ \mathbb{1}_{\mathcal{M}_{h_j}}\big(x + (f(u),u)\big)\, g(u)\, du\, \overline{\check{W}}(2^{2j}x) dx. \end{aligned} \tag{45}$$

and $B_{\Delta_0} = \{x \in \mathbb{R}^2 : |x| \le 2^{-(2-\Delta_0)j}\}$, with any $\Delta_0 > 0$.

For $\gamma^{(2)}_{j,\ell,k}$, we proceed as in the wavelet case. Notice that

$$k + B_1^{\ell} A_1^{j}(x + (f(u),u)) = \Big(k_1 + 2^{2j}(x_1 + f(u)) + 2^j \ell(x_2 + u), k_2 + 2^j(x_2 + u)\Big).$$

Now we apply Lemmas A.1 and A.3 in (45) and remember that $\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}(x + (f(u),u))|g(u)|du \le \int_a^b |g(u)|du \le C$. So, we have

$$\begin{aligned}\sum_{(\ell,k)\in M_j} |\gamma^{(2)}_{j,\ell,k}| &\le \sum_{|\ell|\le 2^j} \sum_{k\in\mathbb{Z}^2} 2^{\frac{11}{2}j} \int_{\mathbb{R}\setminus B_{\Delta_0}} \int_a^b C \\ &\times \left(1 + \Big(k_1 + 2^{2j}(x_1 + f(u)) + 2^j\ell(x_2 + u)\Big)^2\right)^{-1} \left(1 + \Big(k_2 + 2^j(x_2 + u)\Big)^2\right)^{-1} \\ &\times |g(u)|du|\overline{\check{W}}(2^{2j}x)|dx \\ &\le C 2^{\frac{11}{2}j} \sum_{|\ell|\le 2^j} \int_{|x|>2^{-(2-\Delta_0)j}} |\overline{\check{W}}(2^{2j}x)|dx \\ &\le C 2^{\frac{11}{2}j} \sum_{|\ell|\le 2^j} \int_{|x|>2^{\Delta_0 j}} C C_N (1 + |x|^2)^{-N} dx \\ &\le C 2^{\frac{5}{2}j} C_N \frac{\pi}{N-1} (1 + 2^{2\Delta_o j})^{-N+1}\end{aligned}$$

where $C$ is an independent constant. So, for a suitable value of $N$, we have $\sum_{(\ell,k)\in M_j} |\gamma^{(2)}_{j,\ell,k}| \le o(2^j)$.

Now we work on (44). Like in the proof of Lemma 3.4, we may assume that $f(0) = f'(0) = 0$, and that $a = -\epsilon, b = \epsilon$. So $f(u) \simeq f''(c)u^2/2$, where $c \in [-\epsilon, \epsilon]$. Since $f \in C^\infty[a,b]$, there is $M > 0$ such that $|f''(u)| \le \frac{1}{2}M$ for all $u \in [-\epsilon,\epsilon]$ and $|x| \le 2^{-(2-\Delta_0)j}$, we have $|f'(u)| \le M\, h_j = o(2^{-\frac{3}{4}j}) \le \frac{1}{3}\, 2^{-\frac{3}{4}j}$ for all large $j$ and all $u \in [-h_j - x_2, h_j - x_2] \subset [-\epsilon,\epsilon]$.

We consider first the case $|\ell| \le 2^{j/4}$. Applying Lemma A.1 to (44) we have

$$
\begin{aligned}
\gamma^{(1)}_{j,\ell,k} &= 2^{\frac{11}{2}j} \int_{B_{\Delta_0}} \int_a^b \int_{\widehat{\mathbb{R}}^2} L\left(W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1})\right) \\
&\times\; L^{-1}\!\left(e^{2\pi i\eta\left(k+B_1^\ell A_1^j(x+(f(u),u))\right)}\right) d\eta\, \mathbb{1}_{h_j}(x+(f(u),u))\, g(u)\, du\, \overline{\check{W}}(2^{2j}x)\, dx \\
&= 2^{\frac{11}{2}j} \int_{B_{\Delta_0}} \int_a^b \int_{\widehat{\mathbb{R}}^2} L\left(W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1})\right) \\
&\times \left(1+\left(k_1+2^{2j}(x_1+f(u))+2^j\ell(x_2+u)\right)^2\right)^{-1} \left(1+\left(k_2+2^j(x_2+u)\right)^2\right)^{-1} d\eta \\
&\times\;\; \mathbb{1}_{h_j}(x+(f(u),u))\, g(u)\, du\, \overline{\check{W}}(2^{2j}x)\, dx
\end{aligned}
$$

We recall from the proof of Proposition 4.4 that there is $C$ independent of $j$ and $x$ such that $\int_a^b \mathbb{1}_{h_j}(x+(f(u),u))\,|g(u)|\,du \le C\,h_j$. So, we have

$$
\begin{aligned}
&\sum_{k\in\mathbb{Z}^2} |\gamma^{(1)}_{j,\ell,k}| \\
&\le C\,2^{\frac{11}{2}j} \int_{\mathbb{R}^2} \int_a^b \sum_{k\in\mathbb{Z}^2} \left(1+(k_1+2^{2j}(x_1+f(u))+2^j\ell(x_2+u))^2\right)^{-1} \\
&\times \left(1+(k_2+2^j(x_2+u))^2\right)^{-1} \mathbb{1}_{h_j}(x+(f(u),u))\, du\, |\check{W}(2^{2j}x)|\, dx \\
&\le C\,2^{\frac{11}{2}j} h_j \int_{\mathbb{R}^2} |\check{W}(2^{2j}x)|\, dx \\
&\le C\,2^{\frac{3}{2}j} h_j.
\end{aligned}
$$

Therefore, using the assumption that $h_j = o(2^{-\frac{3}{4}j})$, we conclude that

$$\sum_{|\ell|\le 2^{j/4}} \sum_{k\in\mathbb{Z}^2} |\gamma^{(1)}_{j,\ell,k}| \le C\, 2^{\frac{1}{4}j} 2^{\frac{3}{2}j} h_j = C\, 2^{\frac{7}{4}j} h_j = o(2^j),$$

We now consider the case $2^{j/4} < |\ell| \le 2^j$. For fixed $\eta$ and $|x| \le 2^{-(2-\Delta_0)j}$, let

$$\begin{aligned} \phi(\eta, x, u) &= \eta\, B_1^\ell A_1^j\, (x + (f(u), u)) = \\ &= 2^{2j}\eta_1(x_1 + f(u)) + 2^j \ell \eta_1 (x_2 + u) + 2^j \eta_2 (x_2 + u). \end{aligned}$$

Then $\phi'_u(\eta, x, u) = \eta_1(2^{2j} f'(u) + 2^j \ell) + 2^j \eta_2 = 2^j \eta_1 \left(2^j f'(u) + \ell + \frac{\eta_2}{\eta_1}\right)$, where $\phi'_u = \frac{\partial}{\partial u}\phi$. From the assumptions of the support of $V$ and $W$ and from (44), it follows that $\frac{1}{16} \le |\eta_1| \le \frac{1}{2}$ and $|\frac{\eta_2}{\eta_1}| \le 1 \le \frac{1}{6} 2^{j/4}$ (for $j \ge 11$). Hence, for all $2^{\frac{1}{4}j} \le |\ell| \le 2^j$ and all $u \in [-h_j - x_2, h_j - x_2]$, there is uniform positive constant independent of $j, \ell$ such that

$$\begin{aligned} |\phi'_u(\eta, x, u)| &\ge |\eta_1| 2^j \left( |\ell| - 2^j |f'(u)| - |\frac{\eta_2}{\eta_1}| \right) \\ &\ge |\eta_1| 2^j \left( |\ell| - \frac{1}{3} 2^{j/4} - \frac{1}{6} 2^{j/4} \right) \\ &\ge C\, 2^j |\ell|. \end{aligned}$$

Hence

$$\frac{1}{|\phi'_u(\eta, x, u)|} \le C 2^{-j} |\ell|^{-1}. \tag{46}$$

So, for fixed $|x| \le 2^{-(2-\Delta_0)j}$ and $\eta$, we define

$$U(\eta, x) = \int_{-\epsilon}^{\epsilon} e^{2\pi i \phi(\eta, x, u)} \mathbb{1}_{h_j}(x + (f(u), u))\, g(u)\, du$$

and observe that $\mathbb{1}_{h_j}(x+(f(u),u)) = 1$ if and only if $|x_2+u_2| \le h_j$ or $-h_j - x_2 \le u \le h_j - x_2$. So,

$$\begin{aligned}
U(\eta,x) &= \int_{-h_j-x_2}^{h_j-x_2} e^{2\pi i\phi(\eta,x,u)}\, g(u)\, du \\
&= \frac{1}{2\pi i} e^{2\pi i\phi(\eta,x,h_j-x_2)} \frac{1}{\phi'_u(\eta,x,u)}\, g(u)\Big|_{-h_j-x_2}^{h_j-x_2} \\
&- \frac{1}{2\pi i}\int_{-h_j-x_2}^{h_j-x_2} (e^{2\pi i\phi(\eta,x,u)})'_u \,\frac{1}{\phi'_u(\eta,x,u)}\, g(u)\, du \\
&= U_1(\eta,x) + U_2(\eta,x) + U_3(\eta,x),
\end{aligned}$$

where

$$\begin{aligned}
U_1(\eta,x) &= \frac{1}{2\pi i} e^{2\pi i\phi(\eta,x,h_j-x_2)} \frac{1}{\phi'_u(\eta,x,h_j-x_2)}\, g(h_j-x_2) \\
U_2(\eta,x) &= -\frac{1}{2\pi i} e^{2\pi i\phi(\eta,x,-h_j-x_2)} \frac{1}{\phi'_u(\eta,x,-h_j-x_2)}\, g(-h_j-x_2) \\
U_3(\eta,x) &= -\frac{1}{2\pi i}\int_{-h_j-x_2}^{h_j-x_2} e^{2\pi i\phi(\eta,x,u)} \left(\frac{1}{\phi'_u(\eta,x,u)} g(u)\right)\Big|'_u \, du.
\end{aligned}$$

Correspondingly, we may write (44) as $\gamma^{(1)}_{j,\ell,k} = \gamma^{(1,1)}_{j,\ell,k} + \gamma^{(1,2)}_{j,\ell,k} + \gamma^{(1,3)}_{j,\ell,k}$, where, for $m = 1,2,3$,

$$\gamma^{(1,m)}_{j,\ell,k} = 2^{\frac{11}{2}j} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1})\, e^{-2\pi i\eta k} \int_{B_{\Delta_0}} U_m(\eta,x)\, \overline{\check{W}(2^{2j}x)}\, dx\, d\eta.$$

We first examine $\gamma^{(1,1)}_{j,\ell,k}$. Using Lemma A.1, where $L$ is given by (13), we have that

$$\begin{aligned}
\gamma^{(1,1)}_{j,\ell,k} &= \frac{2^{\frac{11}{2}j}}{2\pi i} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V\left(\tfrac{\eta_2}{\eta_1}\right)\, e^{-2\pi i\eta k} \\
&\times \int_{B_{\Delta_0}} e^{2\pi i\phi(\eta,x,h_j-x_2)} \frac{1}{\phi'_u(\eta,x,h_j-x_2)}\, g(h_j-x_2)\, \overline{\check{W}(2^{2j}(x))}\, dx\, d\eta \\
&= \frac{2^{\frac{11}{2}j}}{2\pi i} \int_{B_{\Delta_0}} \int_{\widehat{\mathbb{R}}^2} L\left( W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V\left(\tfrac{\eta_2}{\eta_1}\right) \frac{1}{\phi'_u(\eta,x,h_j-x_2)}\right) \\
&\times L^{-1}\left( e^{-2\pi i\eta\cdot\left(k-\left(2^{2j}(x_1+f(h_j-x_2))+\ell 2^j h_j, 2^j h_j\right)\right)}\right) d\eta\, g(h_j-x_2)\, \overline{\check{W}(2^{2j}x)}\, dx \\
&= \frac{2^{\frac{11}{2}j}}{2\pi i} \int_{B_{\Delta_0}} \int_{\widehat{\mathbb{R}}^2} L\left( W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V\left(\tfrac{\eta_2}{\eta_1}\right) \frac{1}{\phi'_u(\eta,x,h_j-x_2)}\right)
\end{aligned}$$

$$\times \left(1+\left(k_1+2^{2j}(x_1+f(h_j-x_2))+2^j\ell h_j\right)^2\right)^{-1}\left(1+\left(k_2+2^jh_j\right)^2\right)^{-1}$$
$$\times \quad e^{-2\pi i\eta\cdot\left(k-\left(2^{2j}(x_1+f(h_j-x_2))+\ell 2^j h_j, 2^j h_j\right)\right)}\,d\eta\, g(h_j-x_2)\overline{\check{W}(2^{2j}x)}\,dx$$

Using inequality (46) and the fact that $2^{\frac{1}{4}j}\le|\ell|\le 2^j$, $\frac{1}{16}\le|\eta_1|\le\frac{1}{2}$, and $|f'(h_j-x_2)|<\frac{1}{3}2^{-\frac{3}{4}j}$, a direct computation shows that there is a uniform constant $C$, independent of $j,\ell$, such that

$$\begin{aligned}
\left|\left(\frac{1}{\phi'_u(\eta,x,h_j-x_2)}\right)'_{\eta_1}\right| &= \left|-\frac{1}{\left(\phi'_u(\eta,x,h_j-x_2)\right)^2}2^j(2^jf'(h_j-x_2)+\ell)\right| \\
&\le C\,\frac{2^j|2^jf'(h_j-x_2)+\ell|}{(|\ell\eta_1|2^j)^2} \\
&\le C\,2^{-j}|\ell|^{-1}
\end{aligned}$$

$$\begin{aligned}
\left|\left(\frac{1}{\phi'_u(\eta,x,h_j-x_2)}\right)'_{\eta_2}\right| &= \left|-\frac{1}{\left(\phi'_u(\eta,x,h_j-x_2)\right)^2}2^j\right| \\
&\le C\,\frac{2^j}{(|\ell||\eta_1|2^j)^2} \\
&\le C\;2^{-j}|\ell|^{-2} \\
&\le C\;2^{-j}|\ell|^{-1}
\end{aligned}$$

The same estimates hold for mixed derivatives. Thus, using these estimates, we have that

$$\begin{aligned}
\sum_{2^{\frac{1}{4}j}\le|\ell|\le 2^j}\sum_{k\in\mathbb{Z}^2}|\gamma^{(1,1)}_{j,\ell,k}| &= \sum_{2^{\frac{1}{4}j}\le|\ell|\le 2^j}\sum_{k\in\mathbb{Z}^2}\frac{2^{\frac{11}{2}j}}{2\pi} \\
&\times\quad\left|\int_{B_{\Delta_0}}\int_{\widehat{\mathbb{R}}^2} L\left(W(\eta_1,2^{-j}(\ell\eta_1+\eta_2))\,V\left(\tfrac{\eta_2}{\eta_1}\right)\frac{1}{\phi'_u(\eta,x,h_j-x_2)}\right)\right. \\
&\times\left(1+\left(k_1+2^{2j}(x_1+f(h_j-x_2))+2^j\ell h_j\right)^2\right)^{-1} \\
&\times\quad\left(1+\left(k_2+2^jh_j\right)^2\right)^{-1} \\
&\times\quad\left. e^{-2\pi i\eta\cdot\left(k-\left(2^{2j}(x_1+f(h_j-x_2))+\ell 2^j h_j, 2^j h_j\right)\right)}\,d\eta\, g(h_j-x_2)\overline{\check{W}(2^{2j}x)}\,dx\right|
\end{aligned}$$

$$
\begin{aligned}
&\le \; 2^{\frac{11}{2}j}\, C \int_{B_{\Delta_0}} \sum_{2^{\frac{1}{4}j} \le |\ell| \le 2^j} |\ell|^{-1}\, 2^{-j} \\
&\times \; \sum_{k\in\mathbb{Z}^2} \left(1+\left(k_1+2^{2j}(x_1+f(h_j-x_2))+2^j\ell h_j\right)^2\right)^{-1} \\
&\times \; \left(1+\left(k_2+2^j h_j\right)^2\right)^{-1} \\
&\times \; |g(h_j-x_2)| \left|\overline{\check{W}(2^{2j}x)}\right| \, dx \\
&\le C\, 2^{\frac{1}{2}j} \sum_{2^{\frac{1}{4}j} \le |\ell| \le 2^j} |\ell|^{-1} \int_{|x|\le 2^{\Delta_0 j}} |\check{W}(x)|\, dx
\end{aligned}
$$

Observe that,

$$
\sum_{2^{j/4}\le|\ell|\le 2^j} |\ell|^{-1} = 2 \sum_{2^{j/4}\le \ell \le 2^j} \ell^{-1} \le 2\int_{2^{j/4}}^{2^j} \ell^{-1} d\ell = 2\ln(2^j) - 2\ln(2^{j/4}) = 2\ln(2^{3j/4}) = \ln(2)3j/2.
$$

$$
\sum_{2^{\frac{1}{4}j}\le|\ell|\le 2^j} \sum_{k\in\mathbb{Z}^2} |\gamma^{(1,1)}_{j,\ell,k}| \le C\, 2^{\frac{1}{2}j} \sum_{2^{\frac{1}{4}j}\le|\ell|\le 2^j} |\ell|^{-1} \int_{|x|\le 2^{\Delta_0 j}} |\check{W}(x)|\, dx \le C\, j\, 2^{\frac{1}{2}j} = o(2^j).
$$

And very similar argument shows that $\sum_{2^{\frac{1}{4}j}\le|\ell|\le 2^j} \sum_{k\in\mathbb{Z}^2} |\gamma^{(1,2)}_{j,\ell,k}| = o(2^j)$.

Finally, for the analysis of $\gamma^{(1,3)}_{j,\ell,k}$, we apply again Lemma A.1 as above. So, we have that

$$
\begin{aligned}
\gamma^{(1,3)}_{j,\ell,k} = \frac{2^{\frac{11}{2}j}}{2\pi i} \int_{B_{\Delta_0}} \int_{-h_j-x_2}^{h_j-x_2} \int_{\widehat{\mathbb{R}}^2} L\left( W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1}) \left(\frac{g(u)}{\phi'_u(\eta,x,u)}\right)'_u \right) \\
\times L^{-1}\left( e^{-2\pi i\eta\cdot\left(k-\left(2^{2j}(x_1+f(u))+\ell 2^j(x_2+u),\, 2^j(x_2+u)\right)\right)} \right) d\eta\, du\, \overline{\check{W}(2^{2j}x)}\, dx. \qquad (47)
\end{aligned}
$$

We observe that

$$
\left[\frac{g(u)}{\phi'_u(\eta,x,u)}\right]'_u = -\frac{\phi''_{u^2}(\eta,x,u)\, g(u)}{(\phi'_u(\eta,x,u))^2} + \frac{g'(u)}{\phi'_u(\eta,x,u)}. \qquad (48)
$$

As we did before, from the assumptions of the support of $V$ and $W$ in the integral (47), we have that $\frac{1}{16} \le |\eta_1| \le \frac{1}{2}$ and $|\frac{\eta_2}{\eta_1}| \le 1 \le \frac{1}{6}2^{j/4}$. Also, recall that $2^{2j}|f''(u)| \le \frac{M}{2}2^{2j}$ for some constant $M > 0$. Hence, for all $2^{\frac{1}{4}j} \le |\ell| \le 2^j$ and all $u \in [-h_j - x_2, h_j - x_2]$, there is uniform positive constant $C$ independent of $j,\ell$ such that

$$\left|\frac{\phi''_{u^2}(\eta,x,u)\,g(u)}{(\phi'_u(\eta,x,u))^2}\right| = \frac{|\eta_1|2^{2j}|f''(u)|\,|g(u)|}{\left(\eta_1\,2^j(2^jf'(u)+2^j+\frac{\eta_2}{\eta_1})\right)^2} \le C\,|\ell|^{-2}.$$

Also, from (46) we have that

$$\left|\frac{g'(u)}{\phi'_u(\eta,x,u)}\right| \le C\,2^{-j}|\ell|^{-1}.$$

Thus, applying these observations in (48), we conclude that, for all $2^{\frac{1}{4}j} \le |\ell| \le 2^j$ and all $u \in [-h_j-x_2, h_j-x_2]$, there is a uniform positive constant $C$ independent of $j,\ell$ such that

$$\left|\left[\frac{g(u)}{\phi'_u(\eta,x,u)}\right]'_u\right| \le C\,(|\ell|^{-2}+2^{-j}|\ell|^{-1}) \le C\,|\ell|^{-1}$$

and, so, that

$$\int_{\widehat{\mathbb{R}}}\left|L\left(W(\eta_1,2^{-j}(\ell\eta_1+\eta_2))\,V(\tfrac{\eta_2}{\eta_1})\left(\frac{g(u)}{\phi'_u(\eta,x,u)}\right)'_u\right)\right|d\eta \le C\,|\ell|^{-1}.$$

Using this estimate in (47), we have that

$$\begin{aligned}
\sum_{k\in\mathbb{Z}^2}|\gamma^{(1,3)}_{j,\ell,k}| \;&\le\; \frac{2^{\frac{11}{2}j}}{2\pi}\int_{B_{\Delta_0}}\int_{-h_j-x_2}^{h_j-x_2}\int_{\widehat{\mathbb{R}}^2}\left|L\left(W(\eta_1,2^{-j}(\ell\eta_1+\eta_2))\,V(\tfrac{\eta_2}{\eta_1})\left(\frac{g(u)}{\phi'_u(\eta,x,u)}\right)'_u\right)\right| \\
&\times\; \sum_{k\in\mathbb{Z}^2}\left(1+\left(k_1-x_1-f(u))-\ell 2^j(x_2+u)\right)^2\right)^{-1} \\
&\times\; \left(1+\left(k_2-2^j(x_2+u)\right)^2\right)^{-1} d\eta\,du\,\overline{\check{W}(2^{2j}x)}\,dx \\
&\le\; 2^{\frac{11}{2}j}C|\ell|^{-1}\int_{B_{\Delta_0}}\int_{-h_j-x_2}^{h_j-x_2}du\,\overline{\check{W}(2^{2j}x)}\,dx \\
&\le\; 2^{\frac{3}{2}j}C|\ell|^{-1}h_j\int_{B_{\Delta_0}}\overline{\check{W}(x)}\,dx \\
&\le\; 2^{\frac{3}{2}j}C|\ell|^{-1}h_j
\end{aligned}$$

Thus,

$$\sum_{2^{\frac{1}{4}j} \le |\ell| \le 2^j} \sum_{k \in \mathbb{Z}^2} |\gamma^{(1,3)}_{j,\ell,k}| \le C\, 2^{\frac{3}{2}j}\, h_j \sum_{2^{\frac{1}{4}j} \le |\ell| \le 2^j} |\ell|^{-1} \le C\, j\, h_j\, 2^{\frac{3}{2}j} = o(2^j).$$

To estimate the terms $\beta_{j,\ell,k}$, we start from the inequality (28) derived above.

$$|\beta_{j,\ell,k}| \le 2^{3j/2} C_N \int_a^b (1 + (k_1 + 2^{2j} f(u) + 2^j \ell u)^2)^{-N} (1 + (k_2 + 2^j u)^2)^{-N} |g(u)| du.$$

In the above inequality, we have $|f(u)| \le 1$ for all $u \in [a,b] \subset [-\epsilon, \epsilon]$, with $\epsilon$ small. For each $j, \ell$, we set

$$K_{j,\ell} = \{k \in \mathbb{Z}^2 : |k_1| \le 2^{2j+2},\, |k_2| \le 2^{j+1}\}$$

and

$$G_{j,\ell} = \{k \in \mathbb{Z}^2 : (\ell, k) \in I_j^c\}.$$

It follows from the definition that, if $k \in K^c_{\ell,j}$, then either $|k_1| > 2^{2j+2}$ or $|k_2| \le 2^{j+1}$. So we have that either $|k_1 - 2^{2j} f(u) - \ell 2^j| \ge 2^{2j}$ or $|k_2 - 2^j u| \ge 2^j$ for all $|\ell| \le 2^j$ (with $|f(u)| \le 1$, $|u| \le \epsilon$). Therefore, performing similar computations as in Proof of Propositions 5.1, it follows from (28) that, for any $N \in \mathbb{N}$, there is a constant $C_N$ such that

$$\sum_{k \in K^c_{j,\ell}} |\beta_{j,\ell,k}| \le C_N\, 2^{\frac{3}{2}j}\, 2^{-(2N-1)j}.$$

Setting $N = 2$ in the last expression, we have that

$$\sum_{k \in K^c_{j,\ell}} |\beta_{j,\ell,k}| \le C\, 2^{\frac{3}{2}j}\, 2^{-3j} = C\, 2^{-\frac{3}{2}j}. \tag{49}$$

We can write

$$\sum_{(\ell,k) \in I_j^c} |\beta_{j,\ell,k}|$$

$$
\begin{aligned}
&\leq \sum_{(\ell,k)\in I_j^c} |\alpha_{j,\ell,k}| + \sum_{(\ell,k)\in I_j^c} |\gamma_{j,\ell,k}| \\
&\leq \sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}} |\alpha_{j,\ell,k}| + \sum_{|\ell|\leq 2^j} \sum_{k\in \mathbb{Z}^2} |\gamma_{j,\ell,k}| \\
&\leq \sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}} |\alpha_{j,\ell,k}| + \sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}^c} |\alpha_{j,\ell,k}| + \sum_{|\ell|\leq 2^j} \sum_{k\in \mathbb{Z}^2} |\gamma_{j,\ell,k}| \\
&\leq \sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}} |\alpha_{j,\ell,k}| + \sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}^c} |\beta_{j,\ell,k}| + 2\sum_{|\ell|\leq 2^j} \sum_{k\in \mathbb{Z}^2} |\gamma_{j,\ell,k}|.
\end{aligned}
$$

Since $k \in G_{j,\ell}$ means $(\ell, k) \in I_j^c$ and since $\#(K_{j,\ell}) = O(2^{3j})$, it follows that

$$
\sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}} |\alpha_{j,\ell,k}| \leq C 2^{3j}\, 2^{-4j}
$$

and, hence,

$$
\sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}} |\alpha_{j,\ell,k}| \leq C 2^j\, 2^{3j}\, 2^{-4j} = C = o(2^j). \tag{50}
$$

Since $G_{j,\ell} \bigcap K_{j,\ell}^c \subset K_{j,\ell}^c$, the estimate (49) gives that

$$
\sum_{|\ell|\leq 2^j} \sum_{k\in G_{j,\ell}\bigcap K_{j,\ell}^c} |\beta_{j,\ell,k}| \leq \sum_{|\ell|\leq 2^j} C\, 2^{-\frac{3}{2}j} \leq C 2^{-\frac{1}{2}j} = o(2^j). \tag{51}
$$

Finally, since $\sum_{(\ell,k)\in M_j} |\gamma_{j,\ell,k}| = o(2^j)$ by (42), combining this estimate with (50) and (51), we have proved (43). $\square$

# 6 Introduction to framelets

In this section we study a signal representation using a special type of function systems called *framelets* that are formed by convolving bases (or possibly frames). The idea of framelets, originally introduced in [64], is to capture both global and local properties of a signal.

Let $f = [f_0, f_1, \ldots, f_{N-1}]^T \in \mathbb{R}^N$ and $\{\phi_i\} \subset \mathbb{R}^N$ and $\{v_j\} \subset \mathbb{R}^\ell$ be orthonormal bases. From [64] we have the following expansion

$$f = \frac{1}{\ell} \sum_{i=1}^{N} \sum_{j=1}^{\ell} \langle f, \phi_i * v_j \rangle \, \phi_i * v_j,$$

where the convolution is assumed to be circular convolution.

For some $d < N$, let $F_m = [f_m, f_{m+1}, \ldots, f_{m+d-1}] \in \mathbb{R}^d$ and $F = [F_0, F_1, \ldots, F_N]^T \in \mathbb{R}^{N \times d}$. $F$ is known as the *Hankel Matrix*. Also, from [64] we have

$$F = \sum_{i=1}^{N} \sum_{j=1}^{\ell} C_{i,j} \phi_i v_j^T$$

where $C_{i,j} = tr(F(\phi_i v_j^T)^T) = \langle f, \phi_i * v_j \rangle$. One typical example of the former decomposition of $F$ is the classical *Singular Value Decomposition* (SVD) $F = U \Sigma V^T = \sum_{i=1}^{rank(F)} \sigma_i \phi_i v_i^T$ where $U = [\phi_1 \cdots \phi_{rank(F)}]$ and $V = [v_1 \ldots v_{rank(F)}]$ are orthogonal matrices and $\Sigma = diag(\sigma_1, \cdots, \sigma_{rank(F)})$ is a diagonal matrix containing the singular values of $F$. As shown in [64], the matrix $C$ concentrates most of its energy (non-zero coefficients) on its upper left corner, so that the representation of $F$ and therefore of $f$ with $\{\phi_i * v_j\}$ is sparse.

## 6.1 2D framelet

For a 2D signal we have from [64] the following result whose proof is presented to contrast the results we present after.

**Proposition 6.1.** *Let $f \in \mathbb{R}^{N \times N}$, let $\{v_{\ell_1,\ell_2} : 1 \leq \ell_1, \ell_2 \leq \ell\} \subset \mathbb{R}^{\ell \times \ell}$ be an orthonormal basis supported inside the square sub-lattice $\{(\ell_1, \ell_2) : 1 \leq \ell_1, \ell_2 \leq \ell\}$ and let $\{\phi_{j_1,j_2}\} \subset \mathbb{R}^{N \times N}$ be an* *orthonormal global basis. We also assume $f$ is periodic inside a lattice (torus). Then, we can write $f$ as*

$$f = \frac{1}{\ell^2} \sum_{j_1,j_2=1}^{N} \sum_{\ell_1,\ell_2=1}^{\ell} \langle f, \phi_{j_1,j_2} * v_{\ell_1,\ell_2} \rangle \phi_{j_1,j_2} * v_{\ell_1,\ell_2}.$$

For completeness we introduce the proof given in [64] here.

**Proof**. Since $\{\phi_{j_1,j_2}\} \subset \mathbb{R}^{N\times N}$ is an orthonormal basis, we have:

$$f = \sum_{j_1,j_2=1}^{N} \langle f, \phi_{j_1,j_2} \rangle \phi_{j_1,j_2} = \sum_{j_1,j_2=1}^{N} tr(f\phi_{j_1,j_2}^T)\phi_{j_1,j_2}.$$

$$\begin{aligned} tr(f\phi_{j_1,j_2}^T) &= tr\Bigg(\sum_{m=1}^{N} f[n,m]\phi_{j_1,j_2}[r,m]\Bigg) \\ &= \sum_{n,m=1}^{N} f[n,m]\phi_{j_1,j_2}[n,m] \\ &= \sum_{s_1,s_2=1}^{N} f[s_1,s_2]\phi_{j_1,j_2}[s_1,s_2]. \end{aligned}$$

Therefore,

$$f = \sum_{j_1,j_2=1}^{N} tr(f\phi_{j_1,j_2}^T)\phi_{j_1,j_2} = \sum_{j_1,j_2=1}^{N} \Bigg( \sum_{s_1,s_2=1}^{N} f[s_1,s_2]\phi_{j_1,j_2}[s_1,s_2] \Bigg) \phi_{j_1,j_2}.$$

$$f[i_1,i_2] = \sum_{j_1,j_2=1}^{N} \Bigg( \sum_{s_1,s_2=1}^{N} f[s_1,s_2]\phi_{j_1,j_2}[s_1,s_2] \Bigg) \phi_{j_1,j_2}(i_1,i_2)$$

By translation we have:

$$f[i_1+\ell_1, i_2+\ell_2] = \sum_{j_1,j_2=1}^{N} \Bigg( \sum_{s_1,s_2=1}^{N} f[s_1+\ell_1, s_2+\ell_2]\phi_{j_1,j_2}[s_1,s_2] \Bigg) \phi_{j_1,j_2}[i_1,i_2].$$

Even though $\{v_{\ell_1,\ell_2}\} \subset \mathbb{R}^{\ell\times\ell}$, we can pad the elements to 0 and use these functions on $\mathbb{R}^{N\times N}$. Then we can write

$$tr(fv_{n_1,n_2}^T) = \left(\sum_{n=1}^{N} f[m,n]v_{n_1,n_2}[r,n]\right) = \sum_{k_1,k_2=1}^{N} f[k_1,k_2]v_{n_1,n_2}[k_1,k_2]$$

and we can write

$$\begin{aligned} P_{\mathbb{R}^{\ell\times\ell}}f &= \sum_{n_1,n_2=1}^{\ell} tr(fv_{n_1,n_2}^T)v_{n_1,n_2} \\ &= \sum_{n_1,n_2=1}^{\ell}\left(\sum_{k_1,k_2=1}^{N} f[k_1,k_2]v_{n_1,n_2}[k_1,k_2]\right)v_{n_1,n_2} \end{aligned}$$

Hence, for translations in $\mathbb{R}^{N\times N}$, we can write elements of $f$ as

$$f[s_1+\ell_1, s_2+\ell_2] = \sum_{n_1,n_2=1}^{\ell}\left(\sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\right)v_{n_1,n_2}[\ell_1,\ell_2].$$

Therefore, combining these observations, we have:

$$\begin{aligned} & f[i_1+\ell_1, i_2+\ell_2] = \sum_{j_1,j_2=1}^{N}\left(\sum_{s_1,s_2=1}^{N} f[s_1+\ell_1,s_2+\ell_2]\phi_{j_1,j_2}[s_1,s_2]\right)\phi_{j_1,j_2}[i_1,i_2] \\ & = \sum_{j_1,j_2=1}^{N}\left(\sum_{s_1,s_2=1}^{N}\sum_{n_1,n_2=1}^{\ell}\left(\sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\right)\right. \\ & \left.\times v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[s_1,s_2]\right)\phi_{j_1,j_2}[i_1,i_2] \\ & = \sum_{j_1,j_2=1}^{N}\sum_{n_1,n_2=1}^{\ell}\left(\sum_{s_1,s_2=1}^{N}\sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[s_1,s_2]\right) \\ & \times v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[i_1,i_2]. \end{aligned}$$

We define,

$$C_{(j_1,j_2),(n_1,n_2)} \quad := \quad \sum_{s_1,s_2=1}^{N}\sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[s_1,s_2].$$

Observing that

$$v_{n_1,n_2} * \phi_{j_1,j_2}[m_1,m_2] = \sum_{k_1,k_2=1}^{N} v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[m_1-k_1,m_2-k_2],$$

with a change of indices we have that

$$\begin{aligned}
\langle f, v_{n_1,n_2} * \phi_{j_1,j_2}\rangle &= \sum_{m_1,m_2=1}^{N} f[m_1,m_2]v_{n_1,n_2} * \phi_{j_1,j_2}[m_1,m_2] \\
&= \sum_{m_1,m_2=1}^{N} f[m_1,m_2] \sum_{k_1,k_2=1}^{N} v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[m_1-k_1,m_2-k_2] \\
&= \sum_{s_1,s_2=1}^{N} \sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[s_1,s_2] \\
&= C_{(j_1,j_2),(n_1,n_2)}.
\end{aligned}$$

Thus

$$\begin{aligned}
&f[i_1+\ell_1,i_2+\ell_2] \\
&= \sum_{j_1,j_2=1}^{N} \sum_{n_1,n_2=1}^{\ell} \left( \sum_{s_1,s_2=1}^{N} \sum_{k_1,k_2=1}^{N} f[s_1+k_1,s_2+k_2]v_{n_1,n_2}[k_1,k_2]\phi_{j_1,j_2}[s_1,s_2] \right) \\
&\times v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[i_1,i_2] \\
&= \sum_{j_1,j_2=1}^{N} \sum_{n_1,n_2=1}^{\ell} \left( C_{(j_1,j_2),(n_1,n_2)} \right) v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[i_1,i_2] \\
&= \sum_{j_1,j_2=1}^{N} \sum_{n_1,n_2=1}^{\ell} \langle f, v_{n_1,n_2} * \phi_{j_1,j_2}\rangle v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[i_1,i_2]
\end{aligned}$$

So, for any $(I_1, I_2)$ in the $N \times N$ lattice, we have

$$\begin{aligned}
f[I_1,I_2] &= \frac{1}{\ell^2} \sum_{i_1+\ell_1=I_1} \sum_{i_2+\ell_2=I_2} f[i_1+\ell_1,i_2+\ell_2] \\
&= \frac{1}{\ell^2} \sum_{j_1,j_2=1}^{N} \sum_{n_1,n_2=1}^{\ell} \langle f, v_{n_1,n_2} * \phi_{j_1,j_2}\rangle \sum_{i_1+\ell_1=I_1} \sum_{i_2+\ell_2=I_2} v_{n_1,n_2}[\ell_1,\ell_2]\phi_{j_1,j_2}[i_1,i_2]
\end{aligned}$$

$$= \frac{1}{\ell^2} \sum_{j_1,j_2=1}^{N} \sum_{n_1,n_2=1}^{\ell} \langle f, v_{n_1,n_2} * \phi_{j_1,j_2} \rangle v_{n_1,n_2} * \phi_{j_1,j_2}[I_1, I_2]. \quad \square$$

## 6.2 Numerical examples

We examine here two numerical examples in 2D. We use two different samples of size $200 \times 200$ from Lena which is widely use in signal processing, Figure 5. We also remark that numerical examples of energy compactification performed in [64] vectorize images.

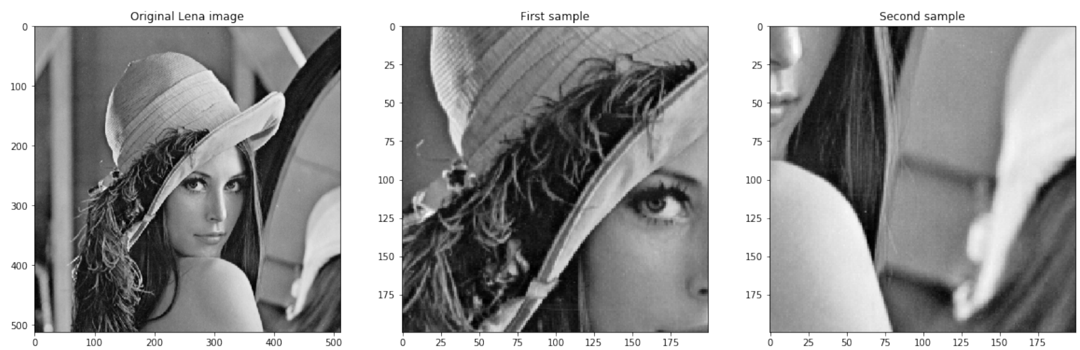


Figure 5: From left to right, Lena image, first sample, second sample

First to adapt the images to the 1D model, we vectorize both samples horizontally and vertically. Then we generate a Hankel Matrix taking patches from each vectorized sample image. We also take 2D patches and vectorize them horizontally and vertically. This is the Hankel Tensor (see Section 8). Then we again form another two Hankel matrices from those 2D patches.

Next, we compute the SVD decomposition and extract the 10 highest singular values. We then truncate those Hankel matrices to those singular values and obtain $\tilde{F} = \sum_{i=1}^{10} \sigma_i \phi_i v_i^T$. From here we now get back the reconstructed sample images, Figures 6 and 7.

Table 1: Error of reconstructed images

| L2-error | first sample | second sample |
|---|---|---|
| horizontal | 0.16422855726305777 | 0.11132959026263807 |
| vertical | 0.1429457243180916 | 0.06879351765953533 |
| 2D patch horizontal | 0.09820448166825274 | 0.050111964879811316 |
| 2D patch vertical | 0.09820448166825295 | 0.050111964879811316 |

We observe that the $L2-error$ of the remonstrated images compared to their original ones

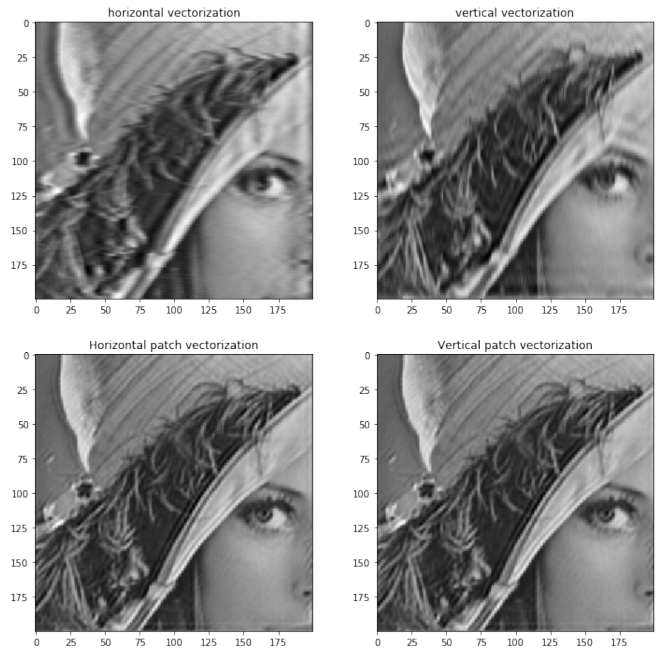


Figure 6: Reconstruction of first sample

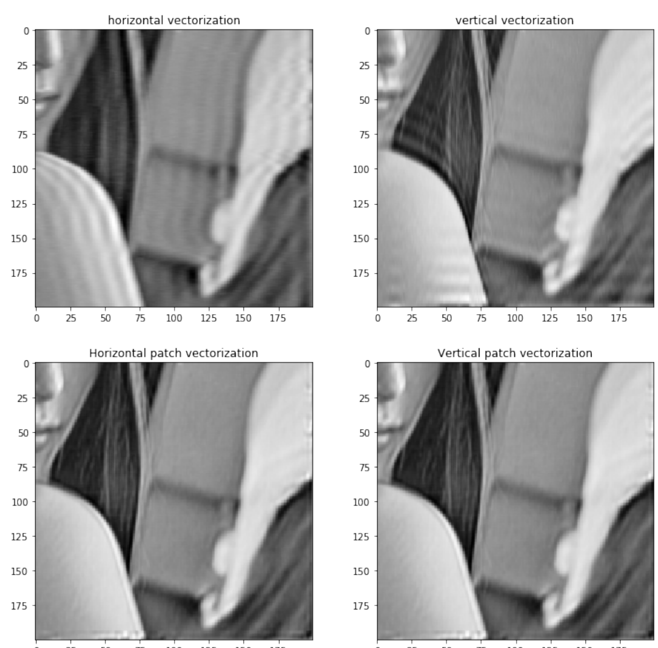


Figure 7: Reconstruction of second sample

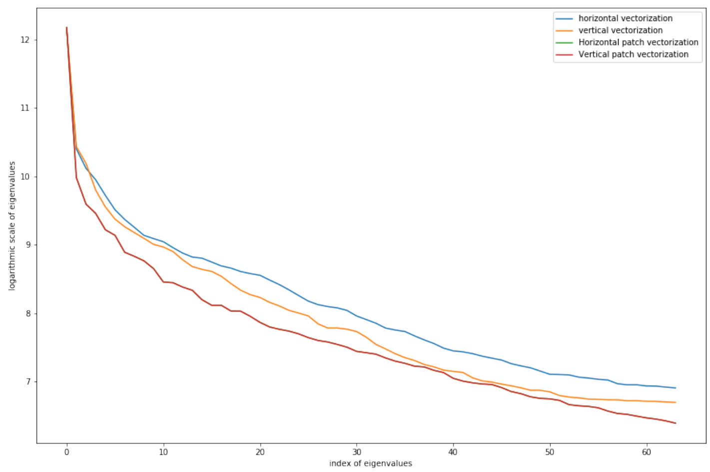


Figure 8: Singular values of the first sample

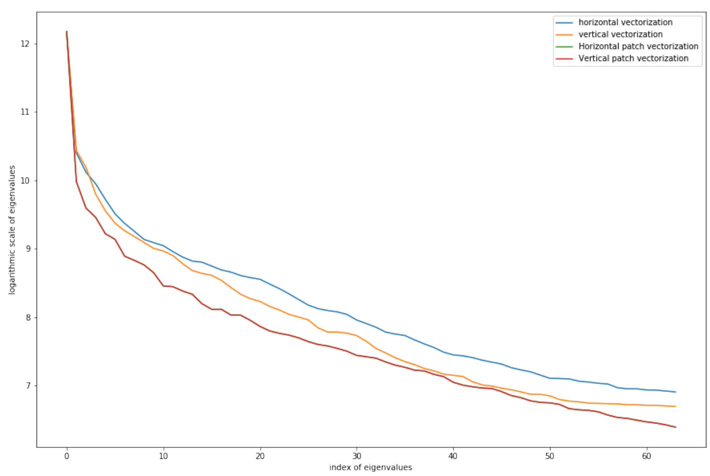


Figure 9: Singular values of the second sample

vary according on how the patches are taken and how those patches are vectorized, Table 1. We remark that we get less error in these experiments when we vectorized 2D patches rather thank taking patches from a vectorized image. We also observe in Figures 8 and 9 that energy is more compact when taking 2D patches. This leads us to consider a tensor approach when we represent 2D signal with framelet expansion. This is what motivates the following work.

# 7 Tensor approach

Our goal is to extend the above analysis to dimension two using tensor tools. We start by defining a global and a local basis. For $N \in \mathbb{N}$, we let $\{\phi_i\} \subset \mathbb{R}^{N\times N}$ be be a Parseval frame (our global orthonormal basis) and, for $\ell \leq N$, we let $\{v_j\} \subset \mathbb{R}^{\ell\times\ell}$ be another Parseval frame (our local orthonormal basis) supported inside the square sub-lattice $\{(\ell_1, \ell_2) : 1 \leq \ell_1, \ell_2 \leq \ell\}$.

Given $f \in \mathbb{R}^{N\times N}$, we want to write $f$ as

$$f = \frac{1}{\ell^2}\sum_i \sum_j \langle f, \phi_i * v_j\rangle \, \phi_i * v_j \tag{52}$$

where the inner product is given by $\langle A, B\rangle = tr(AB^T)$.

We recall that, in the proof of Proposition 1 from appendix B in [64], there is a matrix $V = [v_1, \cdots, v_p] \in \mathbb{R}^{\ell\times p}$ such that $VV^T = I_\ell$ which is the frame condition. This is equivalent to the equation

$$\delta(n-m) = I_\ell[n,m] = \sum_{i=1}^{p}(v_i v_i^T)[n,m] = \sum_{i=1}^{p} v_i[n]v_i[m] = \sum_{i=1}^{p} v_i \otimes v_i[n,m],$$

where $\delta(\cdot)$ is the Kronecker delta and $\otimes$ is the *tensor product* introduced in [40]. Notice that, for any two vectors $a, b \in \mathbb{R}^N$, we have $a \otimes b \in \mathbb{R}^{N\times N}$ where $(a\otimes b)[i,j] = a_i b_j$. So, we define $a \otimes b \in \mathbb{R}^{\ell\times\ell} \otimes \mathbb{R}^{\ell\times\ell}$ where $a, b \in \mathbb{R}^{\ell\times\ell}$ as

$$(a\otimes b)[n,m,p,q] = a[n,m]b[p,q].$$

**Remark 7.1.** *We may also regard the tensor $a \otimes b$ as a linear map $a\otimes b : \mathbb{R}^{\ell\times\ell} \to \mathbb{R}^{\ell\times\ell}$ given by $(a\otimes b)(v) = \langle v, a\rangle b$ where $\langle v, a\rangle = tr(va^T)$.*

Following the approach in [64], we have

**Proposition 7.2.** *Let* $\{v_j : 1 \le j \le p\} \subset \mathbb{R}^{\ell\times\ell}$ *be a Parseval frame, i.e., it satisfies*

$$I = \sum_{j=1}^{p} v_j \otimes v_j$$

*where* $I$ *is a* $4D$ *tensor with entries* $I[a,b,c,d] = \delta(a,b,c,d)$*. Then, for any* $f \in \mathbb{R}^{N\times N}$*, we have:*

$$f = \frac{1}{\ell^2}\sum_{j=1}^{p} f * v_j * v_j(-\cdot) \tag{53}$$

*here we take the convention for* $v(-\cdot)$ *in* (53) *as*

$$v(-\cdot)[n,m] = v[-n,-m] = v[N-n, N-m]$$

*assuming periodicity* $N$ *with respect to both matrix indices*

**Proof.** Direct computations show that

$$v_j * v_j(-\cdot)[n_1,n_2] = \sum_{m_1,m_2=0}^{\ell-1} v_j[n_1+m_1, n_2+m_2]v_j[m_1,m_2].$$

$$\begin{aligned}
\sum_{j=1}^{p} v_j * v_j(-\cdot)[n_1,n_2] &= \sum_{m_1,m_2=0}^{\ell-1}\sum_{j=1}^{p} v_j[n_1+m_1,n_2+m_2]v_j[m_1,m_2] \\
&= \sum_{m_1,m_2=0}^{\ell-1}\sum_{j=1}^{p} v_j\otimes v_j[n_1+m_1,n_2+m_2,m_1,m_2] \\
&= \sum_{m_1,m_2=0}^{\ell-1} I[n_1+m_1,n_2+m_2,m_1,m_2] \\
&= \sum_{m_1,m_2=0}^{\ell-1} \delta(n_1-n_2) \\
&= \ell^2\delta(n_1-n_2).
\end{aligned}$$

Therefore,

$$\sum_{j=1}^{p} f * v_j * v_j(-\cdot) \quad = \quad f * \left( \sum_{j=1}^{p} v_j * v_j(-\cdot) \right) = \ell^2 f * \delta = \ell^2 f. \quad \square$$

Using (53), and a Parseval frame $\{\phi_i : 1 \leq i \leq q\} \subset \mathbb{R}^{N \times N}$, we get

$$\begin{aligned} f &= \frac{1}{\ell^2} \sum_{j=1}^{p} f * v_j * v_j(-\cdot) \\ &= \frac{1}{\ell^2} \sum_{j=1}^{p} \Big( \sum_{i=1}^{q} \langle f * v_j(-\cdot), \phi_i \rangle \phi_i \Big) * v_j \\ &= \frac{1}{\ell^2} \sum_{j=1}^{p} \sum_{i=1}^{q} \langle f * v_j(-\cdot), \phi_i \rangle \phi_i * v_j. \end{aligned}$$

So, in order to have (52), we need

**Proposition 7.3.** $\langle f * v_j(-\cdot), \phi_i \rangle = \langle f, \phi_i * v_j \rangle$.

**Proof.** Direct computations show

$$\begin{aligned} \langle f, \phi_i * v_j \rangle &= tr\Big( f (\phi_i * v_j)^T \Big) \\ &= \sum_{n_1, n_2 = 0}^{N-1} f[n_1, n_2] (\phi_i * v_j)[n_1, n_2] \\ &= \sum_{n_1, n_2 = 0}^{N-1} f[n_1, n_2] \sum_{m_1, m_2 = 0}^{\ell - 1} \phi_i[n_1 - m_1, n_2 - m_2] v_j[m_1, m_2]. \end{aligned}$$

On the other side,

$$\begin{aligned} f * v_j(-\cdot)[n_1, n_2] &= \sum_{m_1, m_2 = 0}^{N-1} f[n_1 - m_1, n_2 - m_2] v_j[N - m_1, N - m_2] \\ &= \sum_{t_1, t_2 = 0}^{N-1} f[n_1 + t_1 - N, n_2 + t_2 - N] v_j[t_1, t_2] \\ &= \sum_{t_1, t_2 = 0}^{\ell - 1} f[n_1 + t_1, n_2 + t_2] v_j[t_1, t_2] \end{aligned}$$

We have used periodicity of the $2D$ signals in the sense that $f(\cdot, n) = f(\cdot, n+N) = f(\cdot, n-N)$. So, we get

$$\begin{aligned}
\langle f * v_j(-\cdot), \phi_i \rangle &= \sum_{n_1,n_2=0}^{N-1} (f * v_j(-\cdot))[n_1,n_2]\phi_i[n_1,n_2] \\
&= \sum_{n_1,n_2=0}^{N-1} \Big( \sum_{m_1,m_2=0}^{\ell-1} f[n_1+m_1, n_2+m_2]v_j[m_1,m_2] \Big)\phi_i[n_1,n_2] \\
&= \sum_{m_1,m_2=0}^{\ell-1} \Big( \sum_{n_1,n_2=0}^{N-1} f[n_1+m_1, n_2+m_2]\phi_i[n_1,n_2] \Big) v_j[m_1,m_2] \\
&= \sum_{m_1,m_2=0}^{\ell-1} \Big( \sum_{t_1,t_2=0}^{N-1} f[t_1,t_2]\phi_i[t_1-m_1, t_2-m_2] \Big) v_j[m_1,m_2] \\
&= \langle f, \phi_i * v_j \rangle. \qquad \square
\end{aligned}$$

# 8 Hankel tensor

We again review material from [64] for the analysis of 1D signals. Let $f = [f_0, f_1, \ldots, f_{N-1}]^T \in \mathbb{R}^N$ and let $F_m = [f_m, f_{m+1}, \ldots, f_{m+d-1}] \in \mathbb{R}^d$, with $d < N$, so we have the Hankel matrix $F = [F_0, F_1, \ldots, F_N]^T \in \mathbb{R}^{N \times d}$.

Also, from [64], we have

$$F = \sum_{i=1}^{N} \sum_{j=1}^{\ell} C_{i,j}\phi_i v_j^T \tag{54}$$

where $\{\phi_i\} \subset \mathbb{R}^N$ and $\{v_j\} \subset \mathbb{R}^\ell$ orthonormal bases and $C_{i,j} = tr(F(\phi_i v_j^T)^T) = \langle f, \phi_i * v_j \rangle$. Observe that $f(n) = \frac{1}{\ell}\sum_{a=0}^{\ell-1} F[n-a, a]$. Following [64],

$$\begin{aligned}
f(n) &= \frac{1}{\ell}\sum_{a=0}^{\ell-1} F[n-a,a] \\
&= \frac{1}{\ell}\sum_{a=0}^{\ell-1} \Big( \sum_{i=1}^{N}\sum_{j=1}^{\ell} C_{i,j}\phi_i v_j^T \Big)[n-a,a]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\ell}\sum_{i=1}^{N}\sum_{j=1}^{\ell} C_{i,j}\sum_{a=0}^{\ell-1}(\phi_i v_j^T)[n-a,a] \\
&= \frac{1}{\ell}\sum_{i=1}^{N}\sum_{j=1}^{\ell} C_{i,j}\sum_{a=0}^{\ell-1}(\phi_i[n-a]v_j[a]) \\
&= \frac{1}{\ell}\sum_{i=1}^{N}\sum_{j=1}^{\ell} C_{i,j}(\phi_i * v_j)[n].
\end{aligned}
$$

Notice here that we have a tensor product:

$$\phi_i v_j^T = \phi_i \otimes v_j \in \mathbb{R}^N \otimes \mathbb{R}^\ell$$

Thus, we see:

$$(\phi_i * v_j)[n] = \sum_{a=0}^{\ell-1}(\phi_i v_j^T)[n-a,a] = \sum_{a=0}^{\ell-1}(\phi_i \otimes v_j)[n-a,a].$$

*Our approach.* We now examine how to extend the above machinery to allow *patches in 2D.* We consider a $2D$ signal $f \in \mathbb{R}^{N\times N}$, and let $F[n,m] = f[n:n+\ell-1, m:m+\ell-1] \in \mathbb{R}^{\ell\times\ell}$ be a patch where $0 \le n,m \le N-1$ and $F \in \mathbb{R}^{(N\times N)} \otimes \mathbb{R}^{(\ell\times\ell)}$ be defined as a $4D$ tensor

$$F[n,m,i,j] := F[n,m][i,j] = f[n+i, m+j]$$

where $0 \le i,j \le \ell-1$. This would be our "Hankel Tensor". So, we have:

$$f[n,m] = \frac{1}{\ell^2}\sum_{i,j=0}^{\ell-1} F[n-i, m-j, i, j].$$

From [40] we have the following *elementary tensor* definition $(A \otimes B) \in \mathbb{R}^{N\times N} \otimes \mathbb{R}^{\ell\times\ell}$

$$(A \otimes B)[n,m,i,j] := A[n,m]B[i,j]$$

where $A \in \mathbb{R}^{N\times N}$ and $B \in \mathbb{R}^{\ell\times\ell}$. Thus, consider the orthonormal bases $\{v_{\ell_1,\ell_2} : 1 \leq \ell_1, \ell_2 \leq \ell\} \subset \mathbb{R}^{\ell\times\ell}$, $\{\phi_{j_1,j_2}\} \subset \mathbb{R}^{N\times N}$ which are a particular case of Parseval frames from Section 7. And also from [40], $\{\phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}\}$ is a basis for $\mathbb{R}^{N\times N} \otimes \mathbb{R}^{\ell\times\ell}$ which is orthonormal under the *induced inner product* defined on $\mathbb{R}^{N\times N} \otimes \mathbb{R}^{\ell\times\ell}$ as

$$\langle A, B\rangle := \sum_{n_1,n_2=0}^{N-1} \sum_{m_1,m_2=0}^{\ell-1} A[n_1, n_2, m_1, m_2]B[n_1, n_2, m_1, m_2]$$

where $A, B \in \mathbb{R}^{N\times N} \otimes \mathbb{R}^{\ell\times\ell}$.

So, we are looking for a similar decomposition of $F$ as in (54), so we have

**Proposition 8.1.** *Let $f \in \mathbb{R}^{N\times N}$ and $F$ its correspondent Hankel tensor, i.e., $F[n, m, i, j] = f[n + i, m + j]$. Also, let $\{v_{\ell_1,\ell_2} : 1 \leq \ell_1, \ell_2 \leq \ell\} \subset \mathbb{R}^{\ell\times\ell}$, $\{\phi_{j_1,j_2}\} \subset \mathbb{R}^{N\times N}$ be orthonormal bases. Then we can write*

$$F = \sum_{j_1,j_2=1}^{N} \sum_{\ell_1,\ell_2}^{\ell} C_{j_1,j_2,\ell_1,\ell_2}\ \phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}$$

*where $C_{j_1,j_2,\ell_1,\ell_2} = \langle f, \phi_{j_1,j_2} * v_{\ell_1,\ell_2}\rangle$.*

**Proof**. It follows from [40] the fact that $\{\phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}\}$ is an orthonormal basis where $C_{j_1,j_2,\ell_1,\ell_2} = \langle F, \phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}\rangle$. Notice that

$$\begin{aligned}
\langle F, \phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}\rangle &= \sum_{n_1,n_2=0}^{N-1} \sum_{m_1,m_2=0}^{\ell-1} F[n_1, n_2, m_1, m_2]\phi_{j_1,j_2} \otimes v_{\ell_1,\ell_2}[n_1, n_2, m_1, m_2] \\
&= \sum_{n_1,n_2=0}^{N-1} \sum_{m_1,m_2=0}^{\ell-1} f[n_1 + m_1, n_2 + m_2]\phi_{j_1,j_2}[n_1, n_2]v_{\ell_1,\ell_2}[m_1, m_2] \\
&= \langle f, \phi_{j_1,j_2} * v_{\ell_1,\ell_2}\rangle
\end{aligned}$$

where the last line is very similar to the computations performed in section 7. $\square$

**Remark 8.2.** *Making again another analogy from the Hankel matrix, we see:*

$$\begin{aligned}
f[n,m] &= \frac{1}{\ell^2}\sum_{i,j=0}^{\ell-1} F[n-i,m-j,i,j] \\
&= \frac{1}{\ell^2}\sum_{i,j=0}^{\ell-1}\left(\sum_{j_1,j_2=1}^{N}\sum_{\ell_1,\ell_2}^{\ell} C_{j_1,j_2,\ell_1,\ell_2}\phi_{j_1,j_2}\otimes v_{\ell_1,\ell_2}\right)[n-i,m-j,i,j] \\
&= \frac{1}{\ell^2}\sum_{j_1,j_2=1}^{N}\sum_{\ell_1,\ell_2}^{\ell} C_{j_1,j_2,\ell_1,\ell_2}\sum_{i,j=0}^{\ell-1}(\phi_{j_1,j_2}\otimes v_{\ell_1,\ell_2})[n-i,m-j,i,j] \\
&= \frac{1}{\ell^2}\sum_{j_1,j_2=1}^{N}\sum_{\ell_1,\ell_2}^{\ell} C_{j_1,j_2,\ell_1,\ell_2}\sum_{i,j=0}^{\ell-1}\phi_{j_1,j_2}[n-i,m-j]v_{\ell_1,\ell_2}[i,j] \\
&= \frac{1}{\ell^2}\sum_{j_1,j_2=1}^{N}\sum_{\ell_1,\ell_2}^{\ell} C_{j_1,j_2,\ell_1,\ell_2}\phi_{j_1,j_2} * v_{\ell_1,\ell_2}[n,m]
\end{aligned}$$

*So, we have again obtained an alternative derivation of* (52).

From here, we need to find an algorithm to perform some form of $SVD$ decomposition in the tensor space $\mathbb{R}^{N\times N}\otimes\mathbb{R}^{\ell\times\ell}$. As vector spaces, $\mathbb{R}^{N\times N}\otimes\mathbb{R}^{\ell\times\ell}$ is isomorphic to $\mathbb{R}^{N\times N\times\ell\times\ell}$ but they are not isomorphic as tensors. As an example, consider the vector spaces $U, V, W$ and the tensor spaces $U\otimes V\otimes W$ and $U\otimes(V\otimes W)$. They are isomorphic as vector spaces, but the tensor $u\otimes(v_1\otimes w_1+v_2\otimes w_2)$ has rank 2 in the first tensor space and rank 1 in the second space. A simpler example, consider the spaces $\mathbb{R}^{3\times 4}$ which is isomorphic as vector space to $\mathbb{R}^{12}$ via vectorization map $vec$. For $A\in\mathbb{R}^{3\times 4}$ with rank two, we see that $vec(A)\in\mathbb{R}^{12}$ has rank one since $vec(A)$ may be considered as a $1\times 12$ or $12\times 1$ matrix. This could be one of the reasons why the Hankel matrices of taking horizontal and vertical vectorization of a 2D signal have different concentration energy when taking SVD decomposition.

As far as we know, [64] does not consider tensor product in their work.

## 8.1 Related and future work

In [62], framelets are used to model a new type of neural networks called *deep convolutional framelets* where several layers of framelet expansion are used to encode and decode signals. They again vectorize 2D signals. In [62], numerical experiments are performed which show deep convolutional framelets improve over existing deep architectures. This success is attributed to the novel signal representation of using non-local basis combined with local basis.

In [14], we find an analysis of *convolutional arithmetic circuits* through tensors. The authors prove that besides a negligible set, all functions that can be implemented by a deep network of polynomial size, require exponential size in order to be realized (or even approximated) by a shallow network. So, the viewpoint of tensor decomposition implies that almost all tensors realized by Hierarchical Tucker (HT) decomposition, see [40], cannot be efficiently realized by the classic CP (rank-1) decomposition. Tensors can effectively model these arithmetical circuits because their pooling operation is just arithmetic multiplication which is bilinear. And again from [40], all tensors are related to bilinear (multilinear) forms. This is the *universal property of tensors*.

Our future work will try to model deep convolutional framelets from [62] with tensors in a similar way as [14] does and try to extract similar properties. This may give new theoretical understanding of why numerical experiments shown in [62] improve existing deep architectures.

# 9 Appendix

**Lemma A.1.** *Let $f \in C_c^\infty(\mathbb{R}^2)$ and $L$ be the differential operator $L = \Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_1^2}\Big)\Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_2^2}\Big)$. For any $N \in \mathbb{N}$, we have that*

$$L^{-N}\big(e^{2\pi i z\cdot x}\big) = (1+x_1^2)^{-N}(1+x_2^2)^{-N}\, e^{2\pi i z\cdot x}. \tag{55}$$

*and*

$$\int_{\mathbb{R}^2} f(z)\, e^{2\pi i z\cdot x} dz = \int_{\mathbb{R}^2} L^N\big(f(z)\big) L^{-N}\big(e^{2\pi i z\cdot x}\big)\, dz.$$

*Proof.* Writing $x = (x_1, x_2)$, we have

$$\begin{aligned} L\big(e^{2\pi i\langle z,x\rangle}\big) &= \Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_1^2}\Big)\Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_2^2}\Big)e^{2\pi i\langle z,x\rangle} \\ &= \Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_1^2}\Big)\Big[e^{2\pi i\langle z,x\rangle} - \frac{1}{(2\pi)^2}e^{2\pi i\langle z,x\rangle}(2\pi i x_1)^2\Big] \\ &= (1+x_1^2)\Big(I - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_1^2}\Big)e^{2\pi i\langle z,x\rangle} \\ &= (1+x_1^2)(1+x_2^2)e^{2\pi i\langle z,x\rangle}. \end{aligned}$$

This implies $L^{-1}\big(e^{2\pi i z\cdot x}\big) = (1+x_1^2)^{-1}(1+x_2^2)^{-1}\, e^{2\pi i z\cdot x}$ and, by induction, we obtain (55). Using these observations, by direct computation we have

$$\begin{aligned} &\int_{\mathbb{R}^2} L\big(f(z)\big)L^{-1}\big(e^{2\pi i z\cdot x}\big)\, dz \\ &= (1+x_1^2)^{-1}(1+x_2^2)^{-1}\int_{\mathbb{R}^2} L\big(f(z)\big)\, e^{2\pi i z\cdot x}\, dz \\ &= (1+x_1^2)^{-1}(1+x_2^2)^{-1}\int_{\mathbb{R}^2}\Bigg(f(z) - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_1^2}f(z) - \frac{1}{(2\pi)^2}\frac{\partial^2}{\partial z_2^2}f(z) \\ &+ \frac{1}{(2\pi)^4}\frac{\partial^2}{\partial z_1^2}\frac{\partial^2}{\partial z_2^2}f(z)\Bigg)\, e^{2\pi i z\cdot x}\, dz \end{aligned}$$

Integrating by parts and using the assumption that $f$ is compactly supported, from the last expression we get:

$$\begin{aligned}
&\int_{\mathbb{R}^2} L\big(f(z)\big)L^{-1}\big(e^{2\pi i z\cdot x}\big)\,dz\\
&= (1+x_1^2)^{-1}(1+x_2^2)^{-1}(1+x_1^2+x_2^2+x_1^2x_2^2)\int_{\mathbb{R}^2} f(z)\,e^{2\pi i z\cdot x}\,dz\\
&= \int_{\mathbb{R}^2} f(z)\,e^{2\pi i z\cdot x}\,dz.
\end{aligned}$$

The general case $N \in \mathbb{N}$ follows by induction. $\square$

**Lemma A.2.** *Let $\mathcal{M}_h = \{(x_1,x_2)\in\mathbb{R}^2 : |x_2|\le h\}$, where $h>0$ and $\phi\in C_c^\infty(\mathbb{R}^2)$. Then*

$$\big(\widehat{\mathbb{1}_{\mathcal{M}_h}} * \widehat{\phi}\big)(\xi) = \big(\widehat{\mathbb{1}_{\mathcal{M}_h}} * \widehat{\phi}\big)(\xi_1,\xi_2) = 2h\int_{\widehat{\mathbb{R}}} \operatorname{sinc}(2\pi h\eta_2)\,\widehat{\phi}\,\big((\xi_1,\xi_2)-(0,\eta_2)\big)\,d\eta_2.$$

*Proof.* Recall that the distributional Fourier transform of $\mathbb{1}_{\mathcal{M}_h}$ is

$$\widehat{\mathbb{1}_{\mathcal{M}_h}}(\xi_1,\xi_2) = 2h\,\operatorname{sinc}(2\pi h\xi_2)\delta_1(\xi_1,\xi_2),$$

where $\int\int_{\widehat{\mathbb{R}}^2}\delta_1(x_1,x_2)\phi(x_1,x_2)dx_1dx_2 = \int_{\widehat{\mathbb{R}}}\phi(0,x_2)\,dx_2$. Thus

$$\begin{aligned}
\big(\widehat{\mathbb{1}_{\mathcal{M}_h}} * \widehat{\phi}\big)(\xi) &= \iint_{\widehat{\mathbb{R}}^2} \widehat{\mathbb{1}_{\mathcal{M}_h}}(\eta)\,\widehat{\phi}(\xi-\eta)\,d\eta\\
&= \iint_{\widehat{\mathbb{R}}^2} 2h\,\operatorname{sinc}(2\pi h\eta_2)\,\delta_1(\eta_1,\eta_2)\,\widehat{\phi}((\xi_1,\xi_2)-(\eta_1,\eta_2))\,d\eta_1\,d\eta_2\\
&= 2h\int_{\widehat{\mathbb{R}}} \operatorname{sinc}(2\pi h\eta_2)\,\widehat{\phi}((\xi_1,\xi_2)-(0,\eta_2))\,d\eta_2. \quad\square
\end{aligned}$$

From [35] we have the following

**Lemma A.3.** *Let $F\in L^2(\mathbb{R}^n)$ such that $\check{F}\in C_c^\infty$, i.e., Fourier Transform of $F$ is smooth and compactly supported. Then for each $N\in\mathbb{N}$, there is a constant $C_N>0$ such that for any $x$*

$$|F(x)| \le (1+|x|^2)^{-N}.$$

# Bibliography


[1] Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G., and Verdera, J. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Transactions on Image Processing 10*, 8 (2001), 1200–1211.

[2] Beck, A. *Introduction to nonlinear optimization theory, algorithms, and applications with MATLAB*. Society for Industrial and Applied Mathematics, U.S., 2015.

[3] Bertalmio, M., Bertozzi, A. L., and 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* (2001), vol. 1, pp. I–I.

[4] Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. Image inpainting. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (USA, 2000), SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., p. 417–424.

[5] Bertalmio, M., Vese, L., Sapiro, G., and Osher, S. Simultaneous structure and texture image inpainting. *IEEE Trans. Image Process 12* (2003), 882–889.

[6] Cai, J.-F., Chan, R. H., and Shen, Z. Simultaneous cartoon and texture inpainting. *Inverse Problems and Imaging 4*, 3 (2010), 379–395.

[7] Cai, J.-F., Dong, B., Osher, S., and Shen, Z. Image restoration: Total variation, wavelet frames, and beyond. *Journal of the American Mathematical Society 25*, 4 (2012), 1033–1089.

[8] Candès, E. J., and Donoho, D. L. New tight frames of curvelets and optimal representations of objects with piecewise c2 singularities. *Communications on Pure and Applied Mathematics 57*, 2 (2004), 219–266.

[9] Chan, T. F., and Kang, S. H. Error analysis for image inpainting. *Journal of Mathematical Imaging and Vision 26*, 1-2 (2006), 85–103.

[10] Chan, T. F., Shen, J., and Zhou, H.-M. Total variation wavelet inpainting. *Journal of Mathematical Imaging and Vision 25*, 1 (2006), 107–125.

[11] Chen, Y., and Hu, H. An improved method for semantic image inpainting with gans: Progressive inpainting. *Neural Processing Letters 49*, 3 (2018), 1355–1367.

[12] Christensen, O. An introduction to frames and riesz bases. *Applied and Numerical Harmonic Analysis* (2003).

[13] Chui, C. K., and Shi., X. Inequalities of Littlewood-Paley type for frames and wavelets. *SIAM Journal on Mathematical Analysis 24*, 1 (1993), 263–277.

[14] Cohen, N., Sharir, O., and Shashua, A. On the expressive power of deep learning: A tensor analysis. *arXiv:1509.05009v3* (2016).

[15] Daubechies, I. *Ten lectures on wavelets*. Soc. for Industrial and Applied Mathematics, 2006.

[16] De Bonet, J. Multiresolution sampling procedure for analysis and synthesis of texture images,. *Proceedings of SIGGRAPH* (1997).

[17] Demir, U., and Gozde, U. Patch-based image inpainting with generative adversarial networks. *arXiv preprint arXiv:1803.07422* (2018).

[18] Dong, B., Ji, H., Li, J., Shen, Z., and Xu, Y. Wavelet frame based blind image inpainting. *Applied and Computational Harmonic Analysis 32*, 2 (2012), 268–279.

[19] Donoho, D., and Huo, X. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory 47*, 7 (2001), 2845–2862.

[20] Donoho, D., and Kutyniok, G. Microlocal analysis of the geometric separation problem. *Communications on Pure and Applied Mathematics 66*, 1 (2012), 1–47.

[21] Donoho, D. L., and Elad, M. Optimally sparse representation in general (nonorthogonal) dictionaries via $\ell_1$-minimization. *Proceedings of the National Academy of Sciences 100*, 5 (2003), 2197–2202.

[22] Donoho, D. L., and Leigh, D. Sparse components of images and optimal atomic decompositions. *Constructive Approximation 17*, 3 (2001), 353–382.

[23] Efros, A., and Leung, T. Texture synthesis by non-parametric sampling. *Proceedings of the Seventh IEEE International Conference on Computer Vision* (1999).

[24] Elad, M. *Sparse and redundant representations, from theory to application in signal and image processing*. Springer, 2010.

[25] Elad, M., and Bruckstein, A. A generalized uncertainty principle and sparse representation in pairs of bases. *IEEE Transactions on Information Theory 48*, 9 (2002), 2558–2567.

[26] Elad, M., Starck, J.-L., Querre, P., and Donoho, D. Simultaneous cartoon and texture image inpainting using morphological component analysis (mca). *Applied and Computational Harmonic Analysis 19*, 3 (2005), 340–358.

[27] Esedoglu, S., and Shen, J. Digital inpainting based on the Mumford–Shah–Euler image model. *European Journal of Applied Mathematics 13*, 04 (2002).

[28] Folland, G. B. *Real analysis: modern techniques and applications*. Wiley, 1999.

[29] Genzel, M. Kutyniok, G. Asymptotic analysis of inpainting via universal shearlet systems. *SIAM J. Imaging Sciences 7* (2014), 2301–2339.

[30] Guleryuz, O. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising - part i, theory,. *IEEE Trans. Image Process. 15* (2006), 539–554.

[31] Guo, K., and Labate, D. Optimally sparse multidimensional representation using shearlets. *SIAM Journal on Mathematical Analysis 39*, 1 (2007), 298–318.

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

[33] Guo, K., and Labate, D. Optimally sparse representations of 3d data with $c^2$ surface singularities using Parseval frames of shearlets. *SIAM Journal on Mathematical Analysis 44*, 2 (2012), 851–886.

[34] Guo, K., and Labate, D. The construction of smooth Parseval frames of shearlets. *Mathematical Modelling of Natural Phenomena 8*, 1 (2013), 82–105.

[35] Guo, K., and Labate, D. Efficient analysis and detection of edges through directional multiscale representations. *Harmonic and Applied Analysis* (2015), 39.

[36] Guo, K., and Labate, D. Characterization and analysis of edges in piecewise smooth functions. *Applied and Computational Harmonic Analysis 41*, 1 (2016), 139–163.

[37] Guo, K., and Labate, D. Detection of singularities by discrete multiscale transforms. *The Journal of Geometric Analysis 28*, 3 (2018), 2102–2128.

[38] Guo, K., and Labate, D. Geometric separation in r3. *to appear in J. Fourier A. Appl. https://doi.org/10.1007/s00041-017-9569-z.* (2019).

[39] Guo, K., Labate, D., and Rodriguez Ayllon, J. P. Image inpainting using sparse multiscale representations: Image recovery performance guarantees. *Applied and Computational Harmonic Analysis 49*, 2 (2020), 343–380.

[40] Hackbusch, W. *Tensor spaces and numerical tensor calculus.* Springer, 2012.

[41] Heil, C. *A basis theory primer.* Birkhäuser, 2011.

[42] Hennenfent, G., Fenelon, L., and Herrmann, F. J. Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach. *Geophysics 75*, 6 (2010).

[43] Herrmann, F. J., and Hennenfent, G. Non-parametric seismic data recovery with curvelet frames. *Geophysical Journal International 173*, 1 (2008), 233–248.

[44] Jean-François, Aubert, G., Blanc-Féraud, L., and Chambolle, A. Image decomposition: Application to textured images and SAR images. *Technical Report ISRN I3S/RR-2003-01-FR, INRIA—Project ARIANA, Sophia Antipolis* (2003).

[45] Jing, Z. On the stability of wavelet and Gabor frames (Riesz bases). *J. Fourier Anal. Appl. 51*, 1 (1999), 105–125.

[46] King, E. J., Kutyniok, G., and Zhuang, X. Analysis of data separation and recovery problems using clustered sparsity. *Wavelets and Sparsity XIV* (2011).

[47] King, E. J., Kutyniok, G., and Zhuang, X. Analysis of inpainting via clustered sparsity and microlocal analysis. *Journal of Mathematical Imaging and Vision 48*, 2 (2014), 205–239.

[48] Kutyniok, G. Labate, D. Shearlets: Multiscale analysis for multivariate data. *Birkhäuser* (2012).

[49] Kutyniok, G., and Lim, W.-Q. Compactly supported shearlets are optimally sparse. *Journal of Approximation Theory 163*, 11 (2011), 1564–1589.

[50] KUTYNIOK, G., AND PETERSEN, P. Classification of edges using compactly supported shearlets. *Appl. Comput. Harmon. Anal. 42*, 2 (2017), 245–293.

[51] MALLAT, S. *A wavelet tour of signal processing, third edition: the sparse way*, 3rd ed. Academic Press, Inc., USA, 2008.

[52] NAM, S., DAVIES, M. E., ELAD, M., AND GRIBONVAL, R. Cosparse analysis modeling - uniqueness and algorithms. *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2011).

[53] RUDIN, L. I., OSHER, S., AND FATEMI, E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena 60* (1992), 259–268.

[54] SHEN, J., AND CHAN, T. F. Mathematical models for local nontexture inpaintings. *SIAM Journal on Applied Mathematics 62*, 3 (2002), 1019–1043.

[55] SHEN, J., KANG, S. H., AND CHAN, T. F. Euler's elastica and curvature-based inpainting. *SIAM Journal on Applied Mathematics 63*, 2 (2003), 564–592.

[56] SHEN, L., XU, Y., AND ZENG, X. Wavelet inpainting with the $\ell_0$ sparse regularization. *Applied and Computational Harmonic Analysis 41* (2016), 26–53.

[57] STARCK, J.-L., ELAD, M., AND DONOHO, D. Redundant multiscale transforms and their application for morphological component separation. *Advances in Imaging and Electron Physics* (2004), 287–348.

[58] STARCK, J.-L., ELAD, M., AND DONOHO, D. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Transactions on Image Processing 14*, 10 (2005), 1570–1582.

[59] STEIN, E. M., AND MURPHY, T. S. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton University Press, 1995.

[60] VESE L.A., O. S. Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Comput. 19* (2003), 553–577.

[61] YANG, C., LU, X., LIN, Z., SHECHTMAN, E., WANG, O., AND LI, H. High-resolution image inpainting using multi-scale neural patch synthesis. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).

[62] YE, J. C., HAN, Y., AND CHA, E. Deep convolutional framelets: A general deep learning framework for inverse problems. *SIAM Journal on Imaging Sciences 11*, 2 (2018), 991–1048.

[63] YEH, R. A., CHEN, C., LIM, T. Y., SCHWING, A. G., HASEGAWA-JOHNSON, M., AND DO, M. N. Semantic image inpainting with deep generative models. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).

[64] YIN, R., GAO, T., LU, Y. M., AND DAUBECHIES, I. A tale of two bases: Local-nonlocal regularization on image patches with convolution framelets. *SIAM Journal on Imaging Sciences 10*, 2 (2017), 711–750.