Image inpainting using sparse multiscale representations: image recovery performance guarantees

Kanghui Guo

Department of Mathematics, Missouri State University, Springfield, Missouri 65804, USA

Demetrio Labate[1]

Department of Mathematics, University of Houston, 651 Phillip G Hoffman, Houston, TX 77204-3008, USA

Jose Pedro Rodriguez Ayllon

Department of Mathematics, University of Houston, 651 Phillip G Hoffman, Houston, TX 77204-3008, USA

Abstract

Several strategies have been applied for the recovery of the missing parts in an image, with recovery performance depending significantly on the image type and the geometry of missing data. To provide a deeper understanding of such image restoration problem, King and al. recently introduced a rigorous multiscale analysis framework and proved that a shearlet based inpainting approach outperforms methods based on more conventional multiscale representations when missing data are line singularities. In this paper, we extend and improve the analysis of the inpainting problem to the more realistic and more challenging setting of images containing curvilinear singularities. We derive inpainting performance guarantees showing that exact image recovery is achieved if the size of the missing singularity is smaller than the size of the structure elements of appropriate functional representations of the image. Our proof relies critically on the microlocal and sparsity properties of the shearlet representation.

Keywords: Inpainting; ℓ^1 minimization; Shearlets; Sparse representations; Analysis of singularities; Wavelets.

Email addresses: `KanghuiGuo@MissouriState.edu` (Kanghui Guo), `dlabate@math.uh.edu` (Demetrio Labate), `jprayllon@gmail.com` (Jose Pedro Rodriguez Ayllon)

[1]Corresponding author.

1. Introduction

The term *inpainting*, originally referred to the art of repairing damaged paintings, describes in signal processing a technique for digitally recovering missing blocks of data in images or streamed videos. Due to its broad range of applications (e.g., restoration of damaged old photos, photo editing, text removal, dis-occlusion in vision analysis), a variety of inpainting methods have been proposed in the literature. State-of-the-art methods for inpainting include most prominently representation-based methods that set up the inpainting problem as an optimization task in a transform domain, e.g., using wavelets, curvelets or shearlets [1, 6, 8, 10, 22]. Other successful methods use PDEs or variational principles to recover the missing data from the close neighborhood of the region to be filled by imposing a criterion of regularity [2, 3, 4, 9]. More recently, following their success in many image processing applications, convolutional neural networks have also been applied to image inpainting with promising results [5, 24, 25, 26]. The performance of existing algorithms depends generally on the type of images considered and the geometry of the missing data.

To provide a deeper understanding of inpainting and assess the ultimate performance of different algorithmic strategies, a rigorous mathematical analysis of the inpainting problem was recently proposed by King et al. [19]. In their work, the inpainting problem is formulated in continuous-domain as the problem of 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. [19] proposed an approach relying on microlocal analysis and sparse approximations. Under the assumption that the unknown image x is sparse with respect to a certain representation system Φ, they search among all possible solutions x^* such that $P_K x^* = x_K$ for the one that minimizes the ℓ^1-norm of the representation coefficients of x^* with respect to Φ. 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. [19] proved that, if the missing information is a line segment, an ℓ^1-norm minimization approach in combination with an appropriate function representation Φ 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 Φ, 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* Φ. 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 [12, 20] – outperforms wavelets and similar conventional multiscale systems. A generalization using a more general shearlet system is given in [11].

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. King et al. remarked in [19] that, even though their proof is limited to linear singularities, it could possibly be extended to the curvilinear case using the Tubular Neighborhood Theorem [7], since the principles of geometric separation from [7] apply to more general singularities. However this task is far from trivial and no extension of the inpainting result to curvilinear singularities was developed to date.

In this paper, we remove the image model restriction of King et al. [19] 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 significant and non-trivial changes in the proofs. While our arguments involve the same concept of clustered sparsity employed in [19] and originally introduced in [7], the fundamental technical elements of our proof 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 by some of the authors in [16, 18]. 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 [19], we consider two strategies for inpainting: one based on ℓ^1 minimization and one based on thresholding. Using ℓ^1 minimization in combination with a shearlet representation, our result recovers the same rate found by King et al. [19] in the case of linear singularities. Interestingly, when a thresholding strategy is applied, we can improve the original result achieving a better convergence rate, i.e., we prove that our inpainting is successful even for a larger gap than the one allowed in [19].

The rest of the paper is organized as follows. In Section 2, we state our main results, namely, Theorems 1-4. In Section 3, we prove some lemmas that are needed for the proofs of our main theorems. We prove Theorems 1 and 2 about wavelets in Section 4, and prove Theorems 3 and 4 about shearlets in Section 5.

We start by establishing some notation and useful definitions.

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 ξ 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 $a = O(b)$ if the limit $\lim_{j\to\infty} \frac{a_j}{b_j}$ exists and $a = o(b)$ if the limit $\lim_{j\to\infty} \frac{a_j}{b_j} = 0$.

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

A set $E = \{e_\lambda : \lambda \in \Lambda\}$ in a Hilbert space $\mathcal{H}$ is a *frame* if there are 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

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

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 $P_Q f$ to be 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 the same symbol C can be used to denote 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 smooth Parseval frames of wavelets and shearlets consisting of smooth band-limited functions.

For the wavelet system, we consider a *Parseval frame of Meyer wavelets* $\Phi = \{\phi_\lambda : \lambda \in \Lambda\} \subset L^2(\mathbb{R}^2)$, where $\Lambda = \bigcup_{j\in\mathbb{Z}} \Lambda_j = \bigcup_{j\in\mathbb{Z}} \{\lambda = (j,k),\, k \in \mathbb{Z}^2\}$ and the functions $\phi_\lambda = \phi_{j,k}$ are defined in the Fourier domain by

$$\widehat{\phi}_{j,k}(\xi) = 2^{-2j}\, W(2^{-2j}\xi)\, e^{2\pi i 2^{-2j}\xi k}; \tag{1}$$

here $W \in C_0^\infty(\widehat{\mathbb{R}}^2)$ is an even function with support $\operatorname{supp}(W) \subset [-\frac{1}{2}, \frac{1}{2}]^2 \setminus [-\frac{1}{16}, \frac{1}{16}]^2$ and satisfying the condition $\sum_{j\in\mathbb{Z}} |W(2^{-2j}\xi)|^2 = 1$, for a.e. $\xi \in \widehat{\mathbb{R}}^2$. Hence the functions $W_j := W(2^{-2j}\cdot)$, with $j \in \mathbb{Z}$, have supports inside the Cartesian coronae

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

Our Parseval frame of shearlets is constructed as in [16] and is obtained by refining the Fourier-domain decomposition of the Parseval frame of wavelets (1) by adding an appropriate directional filtering. This operation has the effect of generating highly anisotropic waveforms ranging over multiple scales and orientations. More precisely, let us consider the following cone-shaped regions in the Fourier domain $\widehat{\mathbb{R}}^2$

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

and let $V \in C_0^\infty(\mathbb{R})$ be chosen so that $\operatorname{supp} V \subset [-1,1]$ and $|V(u-1)|^2+|V(u)|^2+|V(u+1)|^2 = 1$ 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 be the same window function used for the wavelet system above. Hence the *shearlet systems associated with the cone-shaped regions* $\mathcal{C}_\nu$, $\nu = 1,2$ are defined as the countable collection of functions

$$\{\psi^{(\nu)}_{j,\ell,k} : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\},$$

where

$$\hat{\psi}^{(\nu)}_{j,\ell,k}(\xi) = |\det A_{(\nu)}|^{-j/2}\, W(2^{-2j}\xi)\, G_{(\nu)}(\xi A^{-j}_{(\nu)} B^{-\ell}_{(\nu)})\, e^{2\pi i \xi A^{-j}_{(\nu)} B^{-\ell}_{(\nu)} k}, \qquad (3)$$

and

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

A *Parseval frame of shearlets* $\Psi = \{\psi_\eta : \eta \in M\}$, where M is a countable index set, is obtained by combining the shearlet systems associated with the cone-shaped regions $\mathcal{C}_\nu$ together with a coarse scale system and appropriate *boundary* shearlets. The boundary shearlets are slightly modified versions of the functions $\psi^{(\nu)}_{j,\ell,k}$, for $\ell = \pm 2^j$, where the modification is needed to ensure that all elements of system are C_0^∞ in the Fourier domain. Their regularity and localization properties are very similar to those of the shearlet functions $\psi^{(\nu)}_{j,\ell,k}$. 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. We refer to [15] for additional details about this construction. 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 [12, 14]. 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 [13, 18, 17, 21]. These properties of shearlets underpin several results derived in this paper.

2. Main results

Similar to [19], we adopt a continuous image model where the missing information to be recovered is associated with singularities in the plane $\mathbb{R}^2$ that are defined as distributions. We next present our two strategies to recover the missing information, namely ℓ^1 minimization and thresholding

2.1. Mathematical model of inpainting

Let S be a simple closed smooth curve contained in $[-1,1]^2 \subset \mathbb{R}^2$ that has nonvanishing curvature everywhere. We define a 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.

For $h > 0$, we denote as $\mathcal{M}_h$ the *horizontal strip domain* $\mathcal{M}_h = \{(x_1, x_2) \in \mathbb{R}^2 : |x_2| \le h\}$ 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. Clearly, we could also assume that the region to be inpainted is contained in a vertical strip domain of width h and our arguments below could be carried out in a very similar way.

To address the inpainting problem, it is convenient to decompose the image into frequency subbands. Hence, we project $\mathcal{T}$ into the subband regions associated with the Cartesian coronae Q_j, $j \in \mathbb{Z}$, given by (2). For $j \in \mathbb{Z}$, we let $\mathcal{T}_j \in L^2(\mathbb{R}^2) \subset \mathcal{S}'(\mathbb{R}^2)$ be defined in the Fourier domain by

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

where $W(2^{-2j}\cdot)$ is band-pass filter that appears in (1). Correspondingly, we have a sequence of masked images

$$f_j = P_{\mathbb{R}^2 \setminus \mathcal{M}_{h_j}} \mathcal{T}_j,$$

where now h_j depends on the scale parameter j

Following [19], we consider two strategies to recover $\mathcal{T}_j$, $j \in \mathbb{Z}$, from the masked image: one based on ℓ^1 minimization and another one based on thresholding. In both cases, we will establish a procedure to construct an approximate solution R_j and show that we can recover $\mathcal{T}_j$ asymptotically as

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

provided $h_j = o(2^{-\alpha j})$ for an appropriate $\alpha > 0$. We will prove that if the reconstruction approach is based on shearlets then α can be taken significantly smaller than in an analogous scheme based on wavelets. This indicates that, as compared with wavelets, shearlets can (asymptotically) recover an image where the width of the missing region is significantly larger.

The ℓ_1 minimization process to recover an approximate solution has the form

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

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

For the thresholding strategy, given a Parseval frame of wavelets or shearlets $E = \{e_\lambda\}_{\lambda \in \Lambda}$ and a sequence of thresholds σ_j, $j \in \mathbb{Z}$, we let $I_j = \{\lambda \in \Lambda :$

$|\langle f, e_\lambda \rangle| \geq \sigma_j\}$. Then the reconstructed image with respect to E is $R_j^\tau = F(\mathbb{1}_{I_j} F^* \mathcal{T}_j)$.

For the ℓ_1 minimization approach with wavelets, and for the thresholding approach with both wavelets and shearlets, we will follow [19] and assume the projection into the missing region to be $P_{\mathcal{M}_{h_j}} f(x) = \mathbb{1}_{|x_2| \leq h_j} f(x)$. However, for the ℓ_1 minimization with shearlets, for a technical reason, we slightly modify the setting in [19] by choosing $P_{\mathcal{M}_{h_j}} f(x) = h_j^{\Delta_0} \mathbb{1}_{|x_2| \leq h_j} f(x)$, for some fixed small $\Delta_0 > 0$.

Our main results are the following theorems.

Theorem 1. *Let Φ be a Parseval frame of wavelets on $L^2(\mathbb{R}^2)$ as defined in Section 1.2 and let R_j^ℓ be the reconstructed image of $\mathcal{T}_j$ obtained via ℓ_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. *Let Φ be a Parseval frame of wavelets on $L^2(\mathbb{R}^2)$ as defined in Section 1.2 and let R_j^τ 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 3. *Let Ψ be a Parseval frame of shearlets on $L^2(\mathbb{R}^2)$ as defined in Section 1.2 and let R_j^ℓ be the reconstructed image of $\mathcal{T}_j$ obtained via ℓ_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 4. *Let Ψ be a Parseval frame of shearlets on $L^2(\mathbb{R}^2)$ as defined in Section 1.2 and let R_j^τ 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*

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

Remark 1. *Our estimates for the ℓ_1 minimization case (Theorems 1 and 3) match those found by King et al. [19] under the simpler assumption that the missing information is a straight line segment. To extend such results to the more challenging setting where the missing information is curvilinear, our proof uses different techniques relying in part on ideas introduced by the authors in [16] and [18]. In the thresholding case (Theorems 2 and 4), our estimates improve those found by King et al. [19] indicating a better inpainting performance (i.e., the size of the missing gap can be larger) than ℓ_1 minimization for both wavelets* *and shearlets. We remark that our proofs of Theorems 2 and 4 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 [19]. Unlike the original argument, our proofs of Theorems 2 and 4 mainly rely on space domain techniques.*

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. [19] prove that, in the thresholding case (for linear gaps) the wavelet rate cannot be improved showing that shearlets perform better than wavelets. There is currently no proof of a similar negative wavelet result in the ℓ_1 *case.*

Finally we remark that our results can be directly extended to cartoon-like images with a smooth boundary of nonvanishing curvature since one can apply the divergence theorem to map the Fourier transform of a cartoon-like image to the Fourier transform of its boundary, as done in [13].

3. Useful technical results

Using a smooth partition of unity, we can decompose a curve as $S = \bigcup_1^M S_m$. We assume that each curve segment S_m has non vanishing curvature. For each $1 \leq m \leq M$, we can parametrize locally each curve S_m either as a vertical curve $(f(u), u)$ or a horizontal curve $(u, f(u))$, where $u \in (a_m, b_m)$ and $f \in C^\infty(a_m, b_m)$. In either case, there is a constant $\kappa > 0$ such that

$$|f''(u)| \geq \kappa > 0 \quad \text{for } u \in [a_m, b_m],$$

any $m \in [1, M]$, due to the nonvanishing curvature assumption. Here we assume that a *vertical curve* is defined so that the slope of the tangent lines to the curve is greater than or equal to 2 so that $|f'(u)| \leq \frac{1}{2}$; similarly a *horizontal curve* is such that the slope of the tangent line to the curve is smaller than or equal to 2 so that $|f'(u)| \leq 2$. According to this assumption, the function $y = \frac{1}{2}x^2$, for $x \in (-1, 1)$, is a horizontal curve while $y^2 = 8x$, for $y \in (-1, 1)$ is a vertical curve and will be written as $(\frac{1}{8}u^2, u), u \in (-1, 1)$.

Corresponding to each curve S_m, $1 \leq m \leq M$, we have a smooth density function $g_m \in C_0^\infty(S_m)$ so that, for any $\phi \in S(\mathbb{R}^2)$, we have

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

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

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

Consistently with the notation that we have introduced above, we let $\mathcal{T}_{m,j}$ be defined in the Fourier domain by $\widehat{\mathcal{T}}_{m,j}(\xi) = W(2^{-2j}\xi)\widehat{\mathcal{T}}_m(\xi)$ so that $\widehat{\mathcal{T}}_j(\xi) = \sum_{m=1}^{M} \widehat{\mathcal{T}}_{m,j}(\xi)$. Finally, it is convenient in many calculations to use polar coordinates where, for any $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$, we write $\xi = \rho\Theta(\theta)$ with $\rho = |\xi| = \sqrt{\xi_1^2 + \xi_1^2}$ and $\Theta(\theta) = (\cos\theta, \sin\theta)$ where $\Theta(0) = (1, 0)$ for $\xi = (0, 0)$ since, by convention, the angle at the origin is zero. Hence, for a vertical curve S_m we can write

$$\widehat{\mathcal{T}}_{m,j}(\rho, \theta) = W(2^{-2j}\rho\Theta(\theta)) \int_{a_m}^{b_m} e^{-2\pi i\rho\Theta(\theta)\cdot(f(u),u)} g_m(u)\, du. \tag{4}$$

Similarly for a horizontal curve we have

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

Below, we establish some estimates providing the analytical tools needed to prove our main results. The following lemma is stated for a vertical curve but a similar result holds for a horizontal curve.

Lemma 1. *Assume that the local curve S_m, for a fixed $m \in [1, M]$ 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 (3). Then, for any $N \in \mathbb{N}$, there exists a constant C_N, independent of j, ℓ, k such that $|\beta^{(2)}_{j,\ell,k}| \le C_N\, 2^{\frac{5}{2}j}\, 2^{-Nj}$.*

Proof. Using Plancherel theorem and (4), we have that

$$\begin{aligned}
\beta^{(2)}_{j,\ell,k} &= \langle \widehat{\psi}^{(2)}_{j,\ell,k}, \widehat{\mathcal{T}}_{m,j} \rangle \\
&= 2^{-\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi)\, V(2^j \frac{\xi_1}{\xi_2} - \ell)\, e^{-2\pi i\xi\left(A_{(2)}^{-j} B_{(2)}^{-\ell} k\right)} \overline{\widehat{\mathcal{T}}_{m,j}}(\xi) d\xi \\
&= 2^{-\frac{3}{2}j} \int_0^\infty \int_0^{2\pi} |W(2^{-2j}\rho\Theta(\theta))|^2\, V(2^j \cot(\theta) - \ell)\, e^{-2\pi i\rho\Theta(\theta)\left(A_{(2)}^{-j} B_{(2)}^{-\ell} k\right)} \\
&\times \int_{a_m}^{b_m} e^{2\pi i\rho\Theta(\theta)\cdot(f(u),u)} g_m(u)\, du\, \rho d\theta\, d\rho.
\end{aligned}$$

Since V is supported on $[-1, 1]$, then the integral above is non-zero only if $|2^j \cot(\theta) - \ell| \le 1$. This implies $|\cot(\theta)| \le 2^{-j}(|\ell| + 1) \le 1 + 2^{-j}$, which gives $|\theta - \frac{\pi}{2}| \le \frac{\pi}{4} + \epsilon_j$ or $|\theta - \frac{3\pi}{2}| \le \frac{\pi}{4} + \epsilon_j$ with $\epsilon_j \to 0$ as $j \to \infty$. Also, since W is supported on $[-\frac{1}{2}, \frac{1}{2}] \setminus [-\frac{1}{16}, \frac{1}{16}]$, then $\frac{1}{16}2^{2j} \le \rho \le 2^{2j}$ (the last inequality could be sharpened to $\rho \le \frac{1}{\sqrt{2}}2^{2j}$). Hence, using these observations we can write

$$\begin{aligned}
\beta^{(2)}_{j,\ell,k} &= 2^{-\frac{3}{2}j} \int_{\frac{1}{16}2^{2j}}^{2^{2j}} \left[\int_{\frac{\pi}{4}-\epsilon_j}^{\frac{3\pi}{4}+\epsilon_j} + \int_{\frac{5\pi}{4}-\epsilon_j}^{\frac{7\pi}{4}+\epsilon_j} \right] |W|^2(2^{-2j}\rho\Theta(\theta))\, V(2^j \cot(\theta) - \ell) \\
&\times\ e^{-2\pi i\rho\Theta(\theta)\left(A_{(2)}^{-j} B_{(2)}^{-\ell} k\right)} \int_{a_m}^{b_m} e^{2\pi i\rho\Theta(\theta)\cdot(f(u),u)} g_m(u)\, du\, \rho d\theta\, d\rho
\end{aligned}$$

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

$$\times \quad e^{-2\pi i 2^{2j} \rho\Theta(\theta)\left(A_{(2)}^{-j} B_{(2)}^{-\ell} k\right)} \int_{a_m}^{b_m} e^{2\pi i 2^{2j} \rho\Theta(\theta)\cdot(f(u),u)} g_m(u) du\, \rho d\theta\, d\rho. \quad (5)$$

Let $\varphi(u) = \Theta(\theta) \cdot (f(u), u)$. Since $|f'(u)| \le \frac{1}{2}$ for all $u \in [a, b]$, there exists $c > 0$ such that

$$|\varphi'(u)| = |\cos\theta f'(u) + \sin\theta| \ge |\sin\theta| - \frac{1}{2}|\cos\theta| \ge c,$$

for all $\theta \in [\frac{\pi}{4} - \epsilon_j, \frac{3\pi}{4} + \epsilon_j] \bigcup [\frac{5\pi}{4} - \epsilon_j, \frac{7\pi}{4} + \epsilon_j]$ and all $u \in [a, b]$. Finally, using repeated integration by parts N times with respect to the variable u in (5) yields that, for every $N \in \mathbb{N}$, there is a constant C_N, dependent on N such that

$$|\beta^{(2)}_{j,\ell,k}| \le C_N 2^{\frac{5}{2}j}\, 2^{-2Nj}. \qquad \square$$

The following lemma is a special case of the classical method of stationary phase (cf. Proposition 3 in [23, Chapter VIII]).

Lemma 2. *Let φ and ψ be smooth functions. Suppose $\varphi'(u_0) = 0$ and $\varphi''(u_0) \neq 0$. If ψ 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 $a(u_0) = \left(\frac{2\pi i}{\varphi''(u_0)}\right)^{\frac{1}{2}} \psi(u_0)$.

Remark 2. *In the following, we will apply Lemma 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 lemma (cf. Proposition 2 and its corollary in [23, Chapter VIII]).

Lemma 3. *(Van der Corput Lemma) Let $k \ge 2$, $\lambda > 0$, and $\phi(x)$ be a real-valued function defined on $[a, b]$ such that $|\phi^{(k)}(x)| \ge 1$ for all $x \in [a, b]$. Also, let ψ be smooth and compactly supported in $[a, b]$. Then*

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

where C_k depends only on k.

Lemma 4. *With the notation introduced above, for any $j \in \mathbb{Z}$, we have*

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

Proof. Using the decomposition $\mathcal{T}_j = \sum_{m=1}^{M} \mathcal{T}_{m,j}$, it is straightforward to see that $\|\mathcal{T}_j\|_2 \simeq \sum_{m=1}^{M} \|\mathcal{T}_{m,j}\|$. Hence to prove the lemma 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 $u \in (-\epsilon, \epsilon)$, 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 2 to get,

$$\widehat{\mathcal{T}}_{m,j}(\rho,\theta) = W(2^{-2j}(\rho\cos\theta, \rho\sin\theta))\, \rho^{-\frac{1}{2}} \left(a(u_\theta)\, e^{-2\pi i\rho\phi(u_\theta)} + O(\rho^{-\frac{1}{2}}) \right) \quad (6)$$

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 $\theta_0 \le |\theta| \le \frac{\pi}{4}$ and $\theta_0 \le |\theta - \pi| \le \frac{\pi}{4}$ and for $|u| \le \epsilon$, we have $|\phi''(u)| = 2\pi|\cos\theta||f''(u)| \ge c > 0$. In this case, we apply Lemma 3 with $k = 2$ to get

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

We have

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

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

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

Then we have

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

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 also need some preparation for the thresholding strategy of inpainting.

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 [19, 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* σ*.*

- **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 following lemma – originally stated in [19, Proposition 3] – gives an estimate of the approximation error of the one-step-thresholding algorithm. For completeness, we include a proof.

Lemma 5. *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

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

we have

$$\begin{aligned} \|x^* - x^0\|_2 &= \|F\mathbb{1}_I F^* P_K x^0 - \big(F\mathbb{1}_I F^* P_K x^0 + F\mathbb{1}_{I^c} F^* P_K x^0 + F\mathbb{1}_I F^* P_M x^0 \\ &+ F\mathbb{1}_{I^c} F^* P_M x^0\big)\,\|_2 \\ &= \|F\mathbb{1}_{I^c} F^* P_K x^0 + F\mathbb{1}_{I^c} F^* P_M x^0 + F\mathbb{1}_I F^* P_M x^0\|_2 \\ &= \|F\mathbb{1}_{I^c} F^* x^0 + F\mathbb{1}_I F^* P_M x^0]\|_2 \\ &\leq e\,(\|\mathbb{1}_{I^c} F^* x^0\|_1 + \|\mathbb{1}_I F^* P_M x^0\|_1). \end{aligned}$$

The last inequality follows from the observation that, due to the equal-norm condition on E, for any $x \in \mathcal{H}$ we have that

$$\|F\mathbb{1}_J F^* x\|_2 = \|\sum_{\lambda \in J} \langle x, e_\lambda \rangle e_\lambda\|_2 \leq \sum_{\lambda \in J} |\langle x, e_\lambda \rangle|\, \|e_\lambda\|_2 \leq e\|x\|_1. \quad \square$$

Given a Hilbert space $\mathcal{H}$ and a Parseval frame $E = \{e_\lambda : \lambda \in \Lambda\}$, a vector $x \in \mathcal{H}$ is δ *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 5 to be small, the signal x_0 must be δ 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 will be sufficient to analyze the situation for a section of the curve S_m, with a fixed $m \in [1, M]$. 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 1 (ℓ_1 minimization)

We will write the set of 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 denote as $S_{w,j} \subset \Lambda_j$ *the indices of the cluster of significant wavelet coefficients* and we assume it to be the set

$$S_{w,j} = \{k = (k_1, k_2)\,,\, |k_1| \leq 2 \cdot 2^{2j},\, |k_2| \leq 2 \cdot 2^{2j}\}.$$

As in [7], corresponding to the sets $S_{w,j} \subset \Lambda_j$, we define the *wavelet approximation error* at the level j as

$$\delta_j^w = \sum_{\lambda \in S_{w,j}^c} |\langle \mathcal{T}_j, \phi_\lambda \rangle| \tag{7}$$

and the *cluster coherences* as

$$\mu_c(S_{w,j}, P_{\mathcal{M}_{h_j}}\Phi; \Phi) = \max_{\lambda'} \sum_{\lambda \in S_{w,j}} |\langle P_{\mathcal{M}_{h_j}}\phi_\lambda, \phi_{\lambda'} \rangle|. \tag{8}$$

We recall the following useful observation from [19, Lemma 1].

Lemma 6. *For any $j \in \mathbb{Z}$ we have*

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

where R_j^ℓ, $\mathcal{T}_j$ are defined as in Theorem 2, δ_j^w is given by (7) *and μ_c by* (8).

Using the above lemma, Theorem 1 then follows directly from the two propositions below whose proofs are in the next subsection.

Proposition 1. *For any $j \in \mathbb{Z}$*

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

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

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

4.1.1. Proofs of Propositions 1 and 2.

Proof of Proposition 1.

Letting $\beta_{j,k} = \langle \phi_{j,k}, \mathcal{T}_j \rangle$ we can write (7) as $\delta_j^w = \sum_{k \in S^c_{w,j}} |\beta_{j,k}|$. Hence, the proof is completed if we show that $\sum_{k \in S^c_{w,j}} |\beta_{j,k}| = o(2^j)$.

Let L be the differential operator

$$L = \left(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial \eta_1^2} \right) \left(I - \frac{1}{(2\pi)^2} \frac{\partial^2}{\partial \eta_2^2} \right). \tag{9}$$

Using Lemma 8, for any natural number N we can write

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

Hence, using (A.1) in the Appendix, we have that, for any $N \in \mathbb{N}$, there is a constant C_N independent of j and k such that

$$|\beta_{j,k}| \le C_N 2^{2j} \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du. \tag{10}$$

If $k \in S^c_{w,j}$, then either $|k_1| > 2 \cdot 2^{2j}$ or $|k_2| > 2 \cdot 2^{2j}$. So, using the fact that $|u|, |f(u)| \le 1$, it follows that if $|k_1| > 2 \cdot 2^{2j}$ then

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

Similarly, if $|k_2| > 2 \cdot 2^{2j}$, then $1 + (k_2 - 2^{2j}|u|)^2 \ge 2^{4j}$. We can write $S^c_{w,j} = A \cup B$ where $A = \{(k_1,k_2) : |k_1| > 2 \cdot 2^{2j}\}$ and $B = \{(k_1,k_2) : |k_2| > 2 \cdot 2^{2j}\}$. Using these observations and inequality (10) for $N = 2$ we have

$$\begin{aligned} &\sum_{k \in S^c_{w,j}} |\beta_{j,k}| \\ \le\ & \sum_{k \in A} |\beta_{j,k}| + \sum_{k \in B} |\beta_{j,k}| \\ \le\ & C \sum_{k \in A} 2^{2j} \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-2} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)| \, du \\ +\ & C \sum_{k \in B} 2^{2j} \int_a^b \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-2} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)| \, du \\ \le\ & C 2^{2j} \int_a^b \sum_{|k_1| \ge 2 \cdot 2^{2j}} 2^{-4j} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-1} \sum_{k_2 \in \mathbb{Z}} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-2} |g(u)| du \end{aligned}$$

$$+\ C2^{2j}\int_a^b \sum_{|k_2|\geq 2\cdot 2^{2j}} 2^{-4j}\left(1+(k_2+2^{2j}u)^2\right)^{-1}\sum_{k_1\in\mathbb{Z}}\left(1+(k_1+2^{2j}f(u))^2\right)^{-2}|g(u)|du$$
$$\leq\ C\,2^{2j}\,2^{-2j}=o(2^j).\quad\square$$

Proof of Proposition 2.

Using Plancherel theorems and the Fourier transform of $\mathbb{1}_{\mathcal{M}_{h_j}}$ (see Lemma 9), we have

$$\begin{aligned}\langle P_{\mathcal{M}_{h_j}}\phi_{j,k},\phi_{j,k'}\rangle &= \langle\widehat{\mathbb{1}_{\mathcal{M}_{h_j}}}*\widehat{\phi_{j,k}},\widehat{\phi_{j,k'}}\rangle\\ &= 2h_j\int_{\widehat{\mathbb{R}}^2}\int_{\mathbb{R}}\operatorname{sinc}(2\pi h_j\eta_2)\,\widehat{\phi_{j,k}}\,(\xi-(0,\eta_2))\,d\eta_2\,\overline{\widehat{\phi_{j,k'}}}(\xi)\,d\xi\\ &= 2h_j2^{-4j}\int_{\widehat{\mathbb{R}}^2}\int_{\mathbb{R}}W(2^{-2j}(\xi_1,\xi_2-\eta_2))\operatorname{sinc}(2\pi h_j\eta_2)\\ &\times\ e^{-2\pi i 2^{-2j}\eta_2 k_2}d\eta_2\,\overline{W}(2^{-2j}\xi)\,e^{2\pi i\xi 2^{-2j}(k-k')}\,d\xi.\end{aligned}$$

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

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

where the function

$$g(\tau)=\int_{\mathbb{R}}W(\tau-(0,\gamma_2))\,\overline{W}(\tau)\operatorname{sinc}(2\pi h_j2^{2j}\gamma_2)\,e^{-2\pi i\gamma_2k_2}\,d\gamma_2$$

is smooth and compactly supported. Notice that by dominated convergence:

$$L(g(\tau))=\int_{\mathbb{R}}L\Big(W(\tau-(0,\gamma_2))\,\overline{W}(\tau)\Big)\operatorname{sinc}(2\pi h_j2^{2j}\gamma_2)\,e^{-2\pi i\gamma_2k_2}\,d\gamma_2$$

and

$$\begin{aligned}|L(g(\tau))| &\leq \int_{\mathbb{R}}\Big|L\Big(W(\tau-(0,\gamma_2))\,\overline{W}(\tau)\Big)\Big|\,|\operatorname{sinc}(2\pi h_j2^{2j}\gamma_2)|\,|e^{-2\pi i\gamma_2k_2}|\,d\gamma_2\\ &\leq \int_{\mathbb{R}}\Big|L\Big(W(\tau-(0,\gamma_2))\,\overline{W}(\tau)\Big)\Big|\,d\gamma_2\end{aligned}$$

which is bounded since W is smooth and compactly supported and $|\text{sinc}|\leq 1$. Hence we can apply Lemma 8 with (A.1) and the differential operator L given by (9) to obtain

$$\begin{aligned}&\langle P_{\mathcal{M}_{h_j}}\phi_{j,k},\phi_{j,k'}\rangle\\ &=2h_j2^{2j}\int_{\widehat{\mathbb{R}}^2}L\left(g(\tau)\right)L^{-1}\left(e^{2\pi i\tau(k-k')}\right)d\tau\\ &=2h_j2^{2j}(1+(k_1-k_1')^2)^{-1}(1+(k_2-k_2')^2)^{-1}\int_{\widehat{R}^2}L\left(g(\tau)\right)\,e^{2\pi i\tau(k-k')}d\tau.\end{aligned}$$

It follows that there is a constant C independent of k, k' and h_j such that

$$|\langle P_{\mathcal{M}_{h_j}}\phi_{j,k}, \phi_{j,k'}\rangle| \le C\, 2h_j 2^{2j}(1+(k_1-k_1')^2)^{-1}(1+(k_2-k_2')^2)^{-1}.$$

From the last estimate we have 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 $j\to\infty$. □

4.2. Proof of Theorem 2 (Thresholding)

We will 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 $\|\phi_{j,k}\|_2 = \|\phi\|_2$ 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 5, we obtain the following estimate.

Proposition 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*

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

Note that $\|\mathbb{1}_{I_j^c}F^*\mathcal{T}_j\|_1 = \sum_{k\in I_j^c}|\beta_{j,k}|$, that $\|\mathbb{1}_{I_j}F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 = \sum_{k\in I_j}|\gamma_{j,k}|$ and that $R_j^\tau = F[\mathbb{1}_{I_j}F^*P_{\mathbb{R}^2\setminus\mathcal{M}_{h_j}}\mathcal{T}_j]$. Since $\|\mathbb{1}_{I_j}F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 \le \|F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1$, it follows from Proposition 3 that Theorem 2 is proved if the following proposition holds.

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

$$(i)\quad \|F^*P_{\mathcal{M}_{h_j}}\mathcal{T}_j\|_1 = \sum_{k\in\mathbb{Z}^2}|\gamma_{j,k}| \le C\,2^{2j}\,h_j = o(\|\mathcal{T}_j\|_2); \tag{11}$$

$$(ii)\quad \sum_{k\in I_j^c}|\beta_{j,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2), \quad \text{as } j\to\infty. \tag{12}$$

Proof. Using Plancherel theorem and the change of variables $2^{-2j}\xi = \eta$, we have that

$$\begin{aligned}\gamma_{j,k} &= \langle\widehat{\phi}_{j,k}, \widehat{P_{\mathcal{M}_{h_j}}\mathcal{T}_j}\rangle\\ &= 2^{-2j}\int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi)\, e^{2\pi i 2^{-2j}\xi k}\,\overline{\widehat{P_{\mathcal{M}_{h_j}}T_j}(\xi)}\, d\xi\\ &= 2^{2j}\int_{\widehat{\mathbb{R}}^2} W(\eta)\, e^{2\pi i\eta k}\,\overline{\widehat{P_{\mathcal{M}_{h_j}}T_j}(2^{2j}\eta)}\, d\eta. \end{aligned}\tag{13}$$

A direct computation shows 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. It follows that

$$\begin{aligned}
& \widehat{P_{\mathcal{M}_{h_j}}\mathcal{T}_j}(2^{2j}\eta) \\
=\ & \int_{\mathbb{R}^2} \mathbb{1}_{\mathcal{M}_{h_j}}(x)\, \mathcal{T}_j(x)\, e^{-2\pi i 2^{2j}\eta x} dx \\
=\ & \int_{\mathbb{R}^2} e^{-2\pi i 2^{2j}\eta x}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x) \int_a^b 2^{4j}\, \check{W}(2^{2j}(x-(f(u),u)))\, g(u)\, du\, dx \\
=\ & \int_a^b \int_{\mathbb{R}^2} e^{-2\pi i 2^{2j}\eta x}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x)\, 2^{4j}\, \check{W}(2^{2j}(x-(f(u),u)))\, dx\, g(u)\, du \\
=\ & \int_a^b \int_{\mathbb{R}^2} e^{-2\pi i 2^{2j}\eta\cdot(x+(f(u),u))}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\, 2^{4j}\, \check{W}(2^{2j}x)\, dx\, g(u)\, du \\
=\ & \int_{\mathbb{R}^2} \int_a^b e^{-2\pi i 2^{2j}\eta\cdot(x+(f(u),u))}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\, g(u)\, du\, 2^{4j}\, \check{W}(2^{2j}x)\, dx \\
=\ & I_1(\eta) + I_2(\eta), \qquad (14)
\end{aligned}$$

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

$$I_1(\eta) = \int_{B_{\Delta_0}} 2^{4j}\, \check{W}(2^{2j}x) \int_a^b e^{-2\pi i 2^{2j}\eta\cdot(x+(f(u),u))}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\, g(u) du\, dx;$$

$$I_2(\eta) = \int_{B^c_{\Delta_0}} 2^{4j}\, \check{W}(2^{2j}x) \int_a^b e^{-2\pi i 2^{2j}\eta\cdot(x+(f(u),u))}\, \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\, g(u) du\, dx.$$

Substituting (14) into (13), we can then write $\gamma_{j,k} = \gamma_{j,k;I_1} + \gamma_{j,k;I_2}$, where

$$\gamma_{j,k;I_i} = 2^{2j} \int_{\widehat{\mathbb{R}}^2} W(\eta)\, e^{2\pi i \eta k}\, \overline{I}_i(\eta)\, d\eta, \quad i = 1, 2.$$

Using Lemma 8 and the differential operator L given by (9), we have

$$\gamma_{j,k;I_1} = 2^{6j} \int_{|x|\le 2^{-(2-\Delta_0)j}} \int_a^b \int_{\widehat{\mathbb{R}}^2} L\left(W(\eta)\right) L^{-1}\left(e^{2\pi i \eta\cdot(k+2^{2j}(x+(f(u),u)))}\right) d\eta$$

$$\times \quad \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))g(u)\,du\,\overline{\check{W}(2^{2j}x)}\,dx. \tag{15}$$

A similar computation gives

$$\begin{aligned}\gamma_{j,k;I_2} &= 2^{6j}\int_{|x|>2^{-(2-\Delta_0)j}}\int_a^b\int_{\widehat{\mathbb{R}}^2} L^2\left(W(\eta)\right)L^{-2}\left(e^{2\pi i\eta\cdot(k+2^{2j}(x+(f(u),u)))}\right)d\eta \\ &\times \quad \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))g(u)\,du\,\overline{\check{W}(2^{2j}x)}\,dx. \end{aligned} \tag{16}$$

To estimate the term $\gamma_{j,k;I_1}$, we observe that, since $W\in C_c^\infty(\mathbb{R}^2)$, then, for any $N\in\mathbb{N}$, there is a constant $C_N>0$ such that $|\check{W}(x)|\le C_N(1+|x|^2)^{-N}$ for all $x\in\mathbb{R}^2$. In addition, we have that $\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))|g(u)|du\le C\,h_j$ where C is a constant independent of $|x|\le 2^{-(2-\Delta_0)j}$. Hence from (15) we have that

$$\begin{aligned}|\gamma_{j,k;I_1}| &\le 2^{6j}\int_{|x|\le 2^{-(2-\Delta_0)j}}\int_a^b\int_{\widehat{\mathbb{R}}^2}|L\left(W(\eta)\right)|\,d\eta \\ &\times \left(1+\left(k_1+2^{2j}(x_1+f(u))\right)^2\right)^{-1}\left(1+\left(k_2+2^{2j}(x_2+u)\right)^2\right)^{-1} \\ &\times \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\,|g(u)|\,du\,\left|\check{W}(2^{2j}x)\right|\,dx\end{aligned}$$

Since W is a smooth function, for any $N\in\mathbb{N}$ we derive that there is a $C_N>0$ such that

$$\begin{aligned}&\sum_{k\in\mathbb{Z}^2}|\gamma_{j,k;I_1}| \\ \le\ & 2^{6j}C\int_{|x|\le 2^{-(2-\Delta_0)j}}\int_a^b\sum_{k\in\mathbb{Z}^2}\left(1+\left(k_1+2^{2j}(x_1+f(u))\right)^2\right)^{-1} \\ \times\ & \left(1+\left(k_2+2^{2j}(x_2+u)\right)^2\right)^{-1}\mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\,|g(u)|\,du\,\left|\check{W}(2^{2j}x)\right|\,dx \\ \le\ & 2^{6j}C\int_{|x|\le 2^{-(2-\Delta_0)j}}\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))\,|g(u)|\,du\,\left|\check{W}(2^{2j}x)\right|dx \\ \le\ & 2^{6j}h_jC_N\int_{|x|\le 2^{-(2-\Delta_0)j}}(1+|2^{2j}x|^2)^{-N}\,dx \\ =\ & 2^{2j}h_jC_N\int_{|u|\le 2^{\Delta_0 j}}(1+|u|^2)^{-N}\,du\end{aligned}$$

where C_N is a constant independent of j. By choosing N large enough, we conclude that:

$$\sum_{k\in\mathbb{Z}^2}|\gamma_{j,k;I_1}|\le C\,2^{2j}h_j \tag{17}$$

To estimate the term $\gamma_{j,k;I_2}$, we proceed similarly, with the difference that now we use the inequality $\int_a^b \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))|g(u)|du\le\int_a^b|g(u)|du\le C$ for some constant C independent of $x \in \mathbb{R}^2$. Hence from (16) we have that

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}^2} |\gamma_{j,k;I_2}| \\
\leq\ & C2^{6j}\int_{|x|>2^{-(2-\Delta_0)j}} \int_a^b \sum_{k\in\mathbb{Z}^2} \Big(1+\big(k_1+2^{2j}(x_1+f(u))\big)^2\Big)^{-1} \\
\times\ & \Big(1+\big(k_2+2^{2j}(x_2+u)\big)^2\Big)^{-1} \mathbb{1}_{\mathcal{M}_{h_j}}(x+(f(u),u))|g(u)|\,du\,\big|\check{W}(2^{2j}x)\big|\,dx \\
\leq\ & C2^{6j}\int_{|x|>2^{-(2-\Delta_0)}} \big|\check{W}(2^{2j}x)\big|\,dx \\
\leq\ & C_N\,2^{2j}\int_{|u|>2^{\Delta_0 j}} (1+|u|^2)^{-N}\,du \\
\leq\ & C_N\,2^{2j}\,2^{-(2N-2)\Delta_0 j}.
\end{aligned}$$

By choosing N large enough in (18) we have $2^{-(2N-2)\Delta_0 j} \leq h_j$ for j sufficiently large. Thus, combining (17) and (18), we have that there is a constant C independent of j such that

$$\|\mathbb{1}_{I_j} F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1 \leq \sum_{k\in\mathbb{Z}^2} |\gamma_{j,k}| \leq C\,2^{2j}\,h_j,$$

and this proves (11).

To prove (12), we estimate $\beta_{j,k}$. Using Plancherel formula, the change of variable $\eta = 2^{-2j}\xi$ and Lemma 8, we have that for any $N \in \mathbb{N}$

$$\begin{aligned}
\beta_{j,k} &= \langle \widehat{\phi}_{j,k}, \widehat{\mathcal{T}}_j \rangle \\
&= 2^{-2j}\int_{\widehat{\mathbb{R}}^2} |W2^{-2j}\xi)|^2\, e^{2\pi i 2^{-2j}\xi k} \int_a^b e^{2\pi i \xi\cdot(f(u),u)} g(u)\,du\,d\xi \\
&= 2^{2j}\int_a^b \int_{\widehat{\mathbb{R}}^2} |W(\eta)|^2\, e^{2\pi i \eta\cdot(k+2^{2j}(f(u),u))} d\eta\, g(u)\,du \\
&= 2^{2j}\int_a^b \int_{\widehat{\mathbb{R}}^2} L^N\big(|W(\eta)|^2\big)\, L^{-N}\Big(e^{2\pi i \eta\cdot(k+2^{2j}(f(u),u))}\Big)\, d\eta\, g(u)\,du.
\end{aligned}$$

Using the observation that W is a smooth function and (A.1) in Appendix, we have that, for any $N \in \mathbb{N}$, there is a constant C_N such that

$$|\beta_{j,k}| \leq C_N\,2^{2j}\int_a^b \big(1+(k_1+2^{2j}f(u))^2\big)^{-N}\big(1+(k_2+2^{2j}u)^2\big)^{-N}|g(u)|du. \tag{18}$$

Let $K_j = \{k : |k_1| \leq 2^{2j+1}, |k_2| \leq 2^{2j+1}\}$. 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 $|f(u)|, |u| \leq 1$, it follows that

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

Similarly, if $|k_2| > 2^{2j+1}$, we have $1 + (k_1 + 2^{2j} f(u))^2 \geq 2^{4j}$. It then follows from (18) that

$$\begin{aligned}
& \sum_{k \in K_j^c} |\beta_{j,k}| \\
\leq \; & C_N \, 2^{2j} \int_a^b \sum_{k \in K_j^c} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du \\
\leq \; & C_N \, 2^{2j} \int_a^b \sum_{|k_1| > 2^{2j+1}, k_2 \in \mathbb{Z}} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} \\
\times \; & \sum_{k_1 \in \mathbb{Z}, |k_2| > 2^{2j+1}} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du \\
\leq \; & C_N \, 2^{2j} \int_a^b 2^{(1-N)4j} \sum_{k \in \mathbb{Z}^2} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-1} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} \\
\times \; & 2^{(1-N)4j} \sum_{k \in \mathbb{Z}^2} \left(1 + (k_1 + 2^{2j} f(u))^2\right)^{-N} \left(1 + (k_2 + 2^{2j} u)^2\right)^{-N} |g(u)| du \\
\leq \; & C_N 2^{-2(2N-3)j}
\end{aligned} \tag{19}$$

Next observe that

$$\begin{aligned}
\sum_{k \in I_j^c} |\beta_{j,k}| \; & \leq \; \sum_{k \in I_j^c \bigcap K_j} |\beta_{j,k}| + \sum_{k \in I_j^c \bigcap K_j^c} |\beta_{j,k}| \\
& \leq \; \sum_{k \in I_j^c \bigcap K_j} |\alpha_{j,k}| + \sum_{k \in I_j^c \bigcap K_j} |\gamma_{j,k}| + \sum_{k \in I_j^c \bigcap K_j^c} |\beta_{j,k}| \\
& \leq \; \sum_{k \in I_j^c \bigcap K_j} |\alpha_{j,k}| + \sum_{k \in \mathbb{Z}^2} |\gamma_{j,k}| + \sum_{k \in K_j^c} |\beta_{j,k}|
\end{aligned} \tag{20}$$

We will now examine each term of the sum (20). For the first term, we see that, if $k \in I_j^c$, then $|\alpha_{j,k}| \leq \sigma_j \leq 2^{-4j}$. It then follows that

$$\sum_{k \in I_j^c \bigcap K_j} |\alpha_{j,k}| \leq \sum_{k \in K_j} 2^{-4j} \leq C \, 2^{4j} \, 2^{-4j} \leq C = o(2^j) = o(\|\mathcal{T}_j\|).$$

For the second term, we apply (19) with $N = 2$ to get

$$\sum_{k \in K_j^c} |\beta_{j,k}| \leq C \leq o(\|\mathcal{T}_j\|) \text{ as } j \to \infty.$$

For the third term, we apply (17) and use the assumption that $h_j = o(2^{-j})$ to get

$$\sum_{k \in \mathbb{Z}^2} |\gamma_{j,k}| \leq C \, 2^{2j} \, h_j = o(2^j) = o(\|\mathcal{T}_j\|) \text{ as } j \to \infty.$$

Combining these observations, we obtain (12). $\square$

5. Inpainting using shearlets

In this section, we examine the inpainting problem using the shearlet system $\Psi = \{\psi_\eta : \eta \in M\}$ defined in Section 1.2.

For this study, we need to analyze 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. Also, as in Section 4, we can fix m for the locally vertical curve S_m and to simplify the notation – since no horizontal curve need to be examined – we will denote S_m by S and $\mathcal{T}_{m,j}$ by $\mathcal{T}_j$ in the following.

5.1. Proof of Theorem 3 (ℓ_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 6. Corresponding to this set, we define the *shearlet approximation error* at the level j as $\delta^s_j = \sum_{\eta \in S^c_{s,j}} |\langle \mathcal{T}_j, \psi_\eta \rangle|$ 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 $S_{s,j,1} \subset M^{(1)}_j$ will be determined below. Since $S_{s,j,2} = \emptyset$, we have

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

As for the wavelet case, from Lemma 1 in [19] we have

Proposition 5.

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

Let $\beta^{(\nu)}_{j,\ell,k} = \langle \widehat{\psi^{(\nu)}_{j,\ell,k}}, \widehat{\mathcal{T}}_j\rangle$, $\nu = 1, 2$. The proof of Theorem 3 follows from the two propositions below where the set $S_{s,j,1}$ is also constructed. The proofs of these propositions in presented in the next subsection.

Proposition 6. *For any $j \in \mathbb{Z}$,*

$$\delta_j^s = \sum_{(\ell,k)\in M_j^{(1)}\setminus S^c_{s,j,1}} |\beta^{(1)}_{j,\ell,k}| + \sum_{(\ell,k)\in M_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2). \tag{21}$$

Proposition 7. *Assume that $h_j = o(2^{-j})$. Then*

$$\max_{\eta'} \sum_{\eta\in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}}\psi^{(1)}_{\eta'}, \psi^{(1)}_{\eta}\rangle| \to 0 \text{ as } j\to\infty; \tag{22}$$

$$\max_{\eta'} \sum_{\eta\in S_{s,j,1}} |\langle P_{\mathcal{M}_{h_j}}\psi^{(2)}_{\eta'}, \psi^{(1)}_{\eta}\rangle| \to 0 \text{ as } j\to\infty. \tag{23}$$

5.1.1. Proof of Propositions 6 and 7.

Proof of Proposition 6.

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

$$\begin{aligned}
\beta^{(1)}_{j,\ell,k} &= \int_{\widehat{\mathbb{R}}^2} \widehat{\psi}^{(1)}_{j,\ell,k}(\xi)\,\overline{\widehat{\mathcal{T}}_j}(\xi)\,d\xi \\
&= 2^{-\frac{3}{2}j}\int_a^b\int_{\widehat{\mathbb{R}}^2} |W(2^{-2j}\xi)|^2 V\Big(2^j\frac{\xi_2}{\xi_1}-\ell\Big) e^{2\pi i\xi\cdot(A^{-j}_{(1)}B^{-\ell}_{(1)}k+(f(u),u))}d\xi\, g(u)du.
\end{aligned} \tag{24}$$

Let $\eta = \xi A^{-j}_{(1)}B^{-\ell}_{(1)}$. Then,

$$\begin{aligned}
\xi\cdot(A^{-j}_{(1)}B^{-\ell}_{(1)}k + (f(u),u)) &= \eta\cdot(k + B^{\ell}_{(1)}A^{j}_{(1)}(f(u),u)) \\
&= \eta\cdot(k_1 + 2^{2j}f(u) + 2^j\ell u, k_2 + 2^j u)
\end{aligned}$$

and, thus

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

Applying Lemma 8, where L is the differential operator (9), we have that, for any $N\in\mathbb{N}$,

$$\beta^{(1)}_{j,\ell,k} = 2^{\frac{3}{2}j}\int_a^b\int_{\widehat{\mathbb{R}}^2} L^N\left(|W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))|^2 V\left(\frac{\eta_2}{\eta_1}\right)\right)$$

$$\times \quad L^{-N}\left(e^{2\pi i\left(\eta_1(k_1+2^{2j}f(u)+2^j\ell u)+\eta_2(k_2+2^j u)\right)}\right) d\eta\, g(u)du.$$

Since W and V are smooth and compactly supported, for any $N \in \mathbb{N}$, there is a constant C_N such that $\left|L^N\left(|W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))|^2 V(\frac{\eta_2}{\eta_1})\right)\right| \le C_N$. Hence, by (A.1) in Appendix,

$$|\beta^{(1)}_{j,\ell,k}| \le C_N\, 2^{\frac{3}{2}j} \int_a^b \left(1+(k_1+2^{2j}f(u)+2^j\ell u)^2\right)^{-N} \left(1+(k_2+2^j u)^2\right)^{-N} du. \tag{25}$$

Similarly for $\beta^{(2)}_{j,\ell,k}$, for any $N \in \mathbb{N}$, there is a constant C_N such that

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

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

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

We observe that, if $|k_2| \ge 2\cdot 2^j$, then $|k_2+2^j u| \ge 2^j$ for all $u \in [a,b]$. Also if $|k_1| \ge 3\cdot 2^{2j}$, and remembering $|\ell| \le 2^j$, then $|k_1+2^{2j}f(u)+2^j\ell u| \ge 2^{2j}$ for all $u \in [a,b]$. It then follows from inequality (25) that

$$\sum_{(\ell,k)\in M^{(1)}_j \setminus K^{(1)}_j} |\beta^{(1)}_{j,\ell,k}| \le C_N\, 2^{\frac{5}{2}j}\, 2^{-(N-1)2j} = C_N\, 2^{\frac{9}{2}j}\, 2^{-2Nj}. \tag{26}$$

Similarly $K^{(2)}_j = \{(j,\ell,k) \in M^{(2)}_j : \; |k_1| \le 2\cdot 2^j,\, |k_2| \le 3\cdot 2^{2j}\}$ 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

$$\sum_{(\ell,k)\in M^{(2)}_j \setminus K^{(2)}_j} |\beta^{(2)}_{j,\ell,k}| \le C_N\, 2^{\frac{7}{2}j}\, 2^{-2Nj}. \tag{27}$$

Since the set $K^{(2)}_j$ contains $O(2^{4j})$ elements, using Lemma 1 we have that, for any $N \in \mathbb{N}$, there is a constant C_N such that

$$\begin{aligned}\sum_{(\ell,k)\in K^{(2)}_j} |\beta^{(2)}_{j,\ell,k}| &\le \sum_{(\ell,k)\in K^{(2)}_j} C_N\, 2^{\frac{5}{2}j}\ 2^{-2Nj} \\ &\le C_N\, 2^{4j} 2^{\frac{5}{2}j} 2^{-2Nj} \\ &= C_N\, 2^{\frac{13}{2}j} 2^{-2Nj}.\end{aligned} \tag{28}$$

Thus, combining (27) and (28), we have that, for any $N \in \mathbb{N}$, there is a constant C_N such that

$$\sum_{(k,\ell)\in M^{(2)}_j} |\beta^{(2)}_{j,\ell,k}| = \Big(\sum_{(\ell,k)\in K^{(2)}_j} + \sum_{(\ell,k)\in M^{(2)}_j \setminus K^{(2)}_j}\Big)|\beta^{(2)}_{j,\ell,k}|$$

$$\begin{aligned} &\leq C_N \left(2^{\frac{7}{2}j}\, 2^{-2Nj} + 2^{\frac{13}{2}j} 2^{-2Nj}\right) \\ &\leq C_N\, 2^{\frac{13}{2}j}\, 2^{-2Nj}. \end{aligned} \tag{29}$$

Choosing N large enough in (29), it follows that

$$\sum_{(\ell,k)\in M_j^{(2)}} |\beta^{(2)}_{j,\ell,k}| = o(2^j), \text{ as } j \to \infty. \tag{30}$$

To complete the estimate of δ_j^s in (21), we need to define $S_{s,j,1}$ and show that $\sum_{(\ell,k)\in S^c_{s,j,1}} |\beta^{(1)}_{j,\ell,k}| = o(2^j)$.

In the integral of $\beta^{(1)}_{j,\ell,k}$, given by (24), we first make the change of variables $\eta = 2^{-2j}\xi$, next convert to polar coordinates with $\eta = \rho\Theta(\theta) = \rho(\cos\theta, \sin\theta)$, hence obtaining

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

As in the proof of Lemma 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 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 $[-(\epsilon_0+\delta_0), \epsilon_0+\delta_0]$. The map from $[-(\epsilon_0+\delta_0), \epsilon_0+\delta_0]$ onto $[-\theta_0, \theta_1]$ is strictly decreasing. 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 $|u| \leq \epsilon_0$, we have $f'(u) + \tan\theta \neq 0$.

It follows that there exists a constant $c > 0$ such that $|\phi'_u(u,\theta)| \geq c$ 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 $|u| \leq \epsilon_0$. Thus integration by parts 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

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

Also as in the proof of lemma 4, for $\frac{\pi}{4} \leq |\theta| \leq \frac{\pi}{2}$ or $\frac{\pi}{4} \leq |\theta - \pi| \leq \frac{\pi}{2}$, we have

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

Thus from the selection of the set $S_{s,j,1}$ to be found later, we see that in order to control $|\beta^{(1)}_{j,\ell,k}|$, we may write $\beta^{(1)}_{j,\ell,k}$ as

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

Since the discussion for the case $\theta \in [\pi - \theta_0, \pi + \theta_1]$ is the identical for the case $\theta \in [-\theta_0, \theta_1]$, we further write $\beta^{(1)}_{j,\ell,k}$ as

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

From the choice of θ_0 and θ_1, we see that for any $\theta \in [-\theta_0, \theta_1]$, there exists a unique $u_\theta \in [-(\epsilon_0 + \delta_0), \epsilon_0 + \delta_0]$ such that $\phi'_u(u_\theta, \theta) = 0$. We remark that unlike in the proof of Lemma 4, we will have $g(u_\theta) = 0$ if $\epsilon_0 \leq |u_\theta| \leq \epsilon_0 + \delta_0$. Now as in the proof of Lemma 4, we apply Lemma 2 to get

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

where $a(u_\theta) = \left(\frac{2\pi i}{\phi''_{u^2}(u_\theta,\theta)} \right)^{\frac{1}{2}} g(u_\theta)$.

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

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

Recall supp$\,(V) \subset [-1, 1]$, which means that for a given ℓ and for $\theta \in [-\theta_0, \theta_1]$, we must have $|2^j \tan\theta - \ell| \leq 1$. This is possible only when $\tan\theta \sim 2^{-j}\ell$ which means that $|2^{-j}\ell|$ needs to be small since $\theta \in [-\theta_0, \theta_1]$. Remember that $|\ell| \leq 2^j$, so in the integral of $\beta_{j,\ell,k}$, we make the change of variables $t = 2^j \tan\theta - \ell$, with $|t| \leq 1$, so that $\tan\theta(t) = 2^{-j}(t+\ell) = 2^{-j}\ell + 2^{-j}t$. Observing that $\theta(t) = \tan^{-1}(2^{-j}(t+\ell))$ and $u_\theta(t) = (f')^{-1}(-2^{-j}(t+\ell))$. Notice $u_\theta(t)$ is well defined, since for large values of $2^{-j}(t+\ell) = \tan\theta$, we have $|\theta| \geq \theta_0$ which corresponds to neglected part of $\beta^{(1)}_{j,\ell,k}$.

It follows that we can write $\beta^{(1)}_{j,\ell,k}$ as

$$\beta^{(1)}_{j,\ell,k} = 2^{\frac{3}{2}j} \int_0^\infty \int_{-1}^1 |W(\rho\Theta(\theta(t))|^2\, V(t)\, a(u_{\theta(t)})\, \frac{e^{2\pi i\rho\, G(t)\cos\theta(t)}}{1+2^{-2j}(t+\ell)^2}\, \rho^{\frac{1}{2}}\, dt\, d\rho, \quad (31)$$

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

$$G(t) = k_1 + tk_2 + 2^{2j} f\left((f')^{-1}(-2^{-j}(t+\ell)\right) + 2^j(t+\ell)\,(f')^{-1}(-2^{-j}(t+\ell)). \quad (32)$$

Note that G is continuous and compactly supported. Hence, for $k = (k_1,k_2) \in \mathbb{Z}^2$, $\ell \in \mathbb{Z}$ with $|\ell| \leq 2^j$, we can pick $t_{k,\ell} \in [-1,1]$ to be defined by the condition

$$|G(t_{k,\ell})| = \inf_{|t|\leq 1} |G(t)|. \quad (33)$$

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

$$S_{s,j,1} = \{(j,\ell,k) \in M^{(1)}_j : |k_1| \leq 3\cdot 2^{2j},\, |k_2| \leq 2\cdot 2^j,\, |G(t_{k,\ell})| \leq 2^{\Delta_0 j}\}. \quad (34)$$

Next, remember for $j > 0$ fixed, we have defined the sets $K^{(1)}_j = \{(j,\ell,k) \in M^{(1)}_j : |k_1| \leq 3\cdot 2^{2j},\, |k_2| \leq 2\cdot 2^j\}$. Similarly we define

$$Q^{(1)}_j = \{(j,\ell,k) \in M^{(1)}_j : |G(t_{k,\ell})| \leq 2^{\Delta_0 j}\}.$$

Since $S_{s,j,1} = K^{(1)}_j \cap Q^{(1)}_j$, then

$$M^{(1)}_j \setminus S_{s,j,1} = (M^{(1)}_j \setminus K^{(1)}_j) \cup (M^{(1)}_j \setminus Q^{(1)}_j) = (M^{(1)}_j \setminus K^{(1)}_j) \cup \Big((M^{(1)}_j \setminus Q^{(1)}_j) \cap K^{(1)}_j\Big).$$

Hence, we can write the first sum in (21) as

$$\sum_{(\ell,k)\in M^{(1)}_j \setminus S_{s,j,1}} |\beta^{(1)}_{j,\ell,k}| = \sum_{(\ell,k)\in M^{(1)}_j \setminus K^{(1)}_j} |\beta^{(1)}_{j,\ell,k}| + \sum_{(\ell,k)\in (M^{(1)}_j \setminus Q^{(1)}_j)\cap K_j} |\beta^{(1)}_{j,\ell,k}|. \quad (35)$$

From equation (26) we have that, for every $N \in \mathbb{N}$, there is a constant $C_N > 0$ such that $\sum_{(\ell,k)\in(M^{(1)}_j \setminus K^{(1)}_j)} |\beta^{(1)}_{j,\ell,k}| \leq C_N\, 2^{\frac{9}{2}j}\, 2^{-2Nj}$. Therefore, choosing N large enough in the last expression, we have that

$$\sum_{(\ell,k)\in(M^{(1)}_j \setminus K^{(1)}_j)} |\beta^{(1)}_{j,\ell,k}| = o(2^j). \quad (36)$$

To estimate the second sum in (35), we observe that, for $(\ell,k) \in (M^{(1)}_j \setminus Q^{(1)}_j) \cap K^{(1)}_j$, we have $|G(t)| \geq 2^{\Delta_0 j}$ for all $t \in [-1,1]$. By repeated integration by parts with respect to the variable ρ in the integral of $\beta^{(1)}_{j,\ell,k}$, given by (31), 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 $(\ell,k) \in (M^{(1)}_j \setminus Q^{(1)}_j) \cap K^{(1)}_j$ 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 $K^{(1)}_j$ is of order 2^{4j}, we have that

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

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 estimate with (36) in (35), and then using the estimate (30), we have that $\delta^s_j = o(2^j)$. $\square$

In order to prove Proposition 7, we need the following lemma.

Lemma 7. *For $j > 0$ fixed and $k \in \mathbb{Z}^2$, $\ell \in \mathbb{Z}$, let $t_{k,\ell}$ be defined by equation (33) in the proof of Proposition 6. 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 7. We start by proving the estimate (22), where the set $S_{s,j,1}$ is given by (34).

Similar to the proof of Proposition 2, using Plancherel theorem and the Fourier transform of $\mathbb{1}_{\mathcal{M}_{h_j}}$ (see Lemma 9), we have

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

Letting $\eta = \xi A_{(1)}^{-j} B_{(1)}^{-\ell}$ so that $\xi = (\xi_1, \xi_2) = \eta B_{(1)}^{\ell} A_{(1)}^{j} = (2^{2j}\eta_1, 2^j(\ell\eta_1 + \eta_2))$, we have

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

Letting $\gamma = 2^{-j}\tau_2$ and then applying Lemma 8, where L is the differential operator (9), we have that:

$$
\begin{aligned}
&\langle P_{\mathcal{M}_{h_j}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle \\
= \;& 2h_j^{1+\Delta_0} 2^j \int_{\widehat{\mathbb{R}}^2} \int_{\widehat{\mathbb{R}}} g_{j,\ell,\ell'}(\eta,\gamma)\, e^{-2\pi i \gamma k_2} \operatorname{sinc}(2\pi\gamma)\, d\gamma\, e^{2\pi i \eta(k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k')}\, d\eta \\
= \;& 2h_j^{1+\Delta_0} 2^j \int_{\widehat{\mathbb{R}}^2} L\left(\int_{\widehat{\mathbb{R}}} g_{j,\ell,\ell'}(\eta,\gamma) e^{-2\pi i \gamma k_2} \operatorname{sinc}(2\pi\gamma)\, d\gamma \right) \\
& L^{-1}\left(e^{2\pi i \eta(k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k')} \right) d\eta, \qquad (37)
\end{aligned}
$$

where

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

Since W, V are compactly supported and smooth, it follows that there is a uniform constant C, independent of j, ℓ, ℓ', such that $|L(g_{j,\ell,\ell'}(\eta,\gamma))| \le C$. Using this estimate and using (A.1) in Appendix with the observation that $(k - B_{(1)}^{\ell} B_{(1)}^{-\ell'} k') = (k_1 - k_1' - (\ell - \ell')k_2', k_2 - k_2')$, it follows from (37) that there is a constant C, independent of j, k, k', ℓ, ℓ', such that

$$
\begin{aligned}
|\langle P_{\mathcal{M}_{2^{-j}}} \psi^{(1)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'} \rangle| \;\le\; & C\, h_j^{1+\Delta_0} 2^j \left(1 + |(k_1 - k_1' - (\ell - \ell')k_2'|^2\right)^{-1} \\
\times\; & \left(1 + |k_2 - k_2'|^2\right)^{-1}. \qquad (38)
\end{aligned}
$$

Recalling the definition of $S_{s,j,1}$, given by (34), and next applying Lemma 7, it follows from (38) that

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

$$\leq C\, h_j^{1+\Delta_0} 2^j\, 2^{\frac{1}{2}\Delta_0 j}.$$

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 (22).

To prove (23), similarly to the computation above, we apply Plancherel theorem and the Fourier transform of $\mathbb{1}_{\mathcal{M}_{h_j}}$ to write

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

We next apply the change of variables $\eta = \xi A_{(1)}^{-j} B_{(1)}^{-\ell} = (2^{-2j}\xi_1, -\ell 2^{-2j}\xi_1 + 2^{-j}\xi_2)$, so that $\xi = \eta B_{(1)}^{\ell} A_{(1)}^{j} = (2^{2j}\eta_1, 2^j(\ell\eta_1 + \eta_2))$, and let $\alpha = (\alpha_1,\alpha_2) = B_{(1)}^{\ell} A_{(1)}^{j} A_{(2)}^{-j} B_{(2)}^{-\ell'} k'$. Hence we have

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

Similar to the calculation above, letting $\gamma = 2^{-j}\tau_2$ and then applying Lemma 8, where L is the differential operator (9), we have that

$$\begin{aligned}
\langle P_{\mathcal{M}_{h_j}} \psi^{(2)}_{j,\ell,k}, \psi^{(1)}_{j,\ell',k'}\rangle \;&=\; 2h_j^{1+\Delta_0} 2^j \int_{\widehat{\mathbb{R}}^2} L\Big(\int_{\widehat{\mathbb{R}}} \tilde{g}_{j,\ell,\ell'}(\eta,\gamma) \operatorname{sinc}(2\pi h_j 2^{-j}\gamma) \\
&\times\; e^{2\pi i \gamma k_2} d\gamma\Big) L^{-1}\left(e^{-2\pi i \eta(k-\alpha)}\right) d\eta, \qquad (39)
\end{aligned}$$

where

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

Using the fact that W, V are compactly supported and smooth, a direct computation gives that there is a uniform constant C, independent of j, ℓ, ℓ', such that $|L(\tilde{g}_{j,\ell,\ell'}(\eta,\gamma))| \leq C$. Therefore, using this observation in (39), we conclude that there is a constant C, independent of j, k, k', ℓ, ℓ', such that

$$|\langle P_{M_{h_j}} \psi^{(2)}_{j,\ell',k'}, \psi^{(1)}_{j,\ell,k}\rangle| \leq C\, h_j^{1+\Delta_o}\, 2^j\, (1+(k_1-\alpha_1)^2)^{-1}(1+(k_2-\alpha_2)^2)^{-1},$$

where the indices α_1, α_2 depend on ℓ. Using the definition of $S_{s,j,1}$, given by (34), and next applying Lemma 7 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

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

We finally prove Lemma 7.
Proof of Lemma 7.
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

$$|G(t_{k,\ell})| = |(k_1+t_{k,\ell}k_2) + 2^{2j}\big(f(y)-f'(y)y\big)| \simeq |(k_1+t_{k,\ell}k_2) - 2^{2j}\tfrac{1}{2}f''(c)y^2|. \quad (40)$$

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

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

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

Hence there is a constant C independent of j, k_1, k_2 such that $\#(Q_k) \le C\, 2^{\frac{1}{2}\Delta_0 j}$.

Case 2: $0 \le k_1 + t_{k,y}k_2 \le 2^{\Delta_0 j+1}$. The inequality (40) implies that

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

Therefore, in this case, we have that

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

Similar to case 1, it follows that

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

and

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

As in Case 1, it follows that there is a constant C independent of j, k_1, k_2 such that $\#(Q_k) \le C\, 2^{\frac{1}{2}\Delta_0 j}$.

Case 3: $k_1 + t_{k,y}k_2 \ge 2^{\Delta_0 j+1}$. The inequality (40) 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

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

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

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

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

5.2. Proof of Theorem 4 (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 ℓ_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 5 then implies the following estimate.

Proposition 8. *For any $j \in \mathbb{Z}$, let R^τ_j, I_j and δ^s_j be defined as above. Then there is a constant C independent of j and $\mathcal{T}$ such that*

$$\|R^\tau_j - \mathcal{T}_j\|_2 \leq C(\delta^s_j + \|\mathbb{1}_{I_j} F^* P_{\mathcal{M}_{h_j}} \mathcal{T}_j\|_1).$$

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

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

It then follows from Proposition 8 that Theorem 4 is true if the following proposition holds.

Proposition 9. *Let $j \geq 0$. For any $0 \leq \sigma_j \leq 2^{-4j}$ and $h_j = o(2^{-\frac{3}{4}j})$, we have*

$$\sum_{(\ell,k) \in M_j} |\gamma_{j,\ell,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2) \tag{41}$$

$$\sum_{(\ell,k) \in I^c_j} |\beta_{j,\ell,k}| = o(2^j) = o(\|\mathcal{T}_j\|_2), \;\; \textit{as } j \to \infty \tag{42}$$

Proof. A direct calculation with the change of variables $\eta = \xi A^{-j}_{(1)} B^{-\ell}_{(1)}$ gives that

$$\begin{aligned}
\gamma_{j,\ell,k} &= \langle \widehat{\psi}^{(\nu)}_{j,\ell,k}, \widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j} \rangle \\
&= 2^{-\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(2^{-2j}\xi)\, V(2^j \frac{\xi_2}{\xi_1} - \ell)\, e^{2\pi i \xi A^{-j}_{(1)} B^{-\ell}_{(1)} k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(\xi)}\, d\xi \\
&= 2^{\frac{3}{2}j} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V(\frac{\eta_2}{\eta_1})\, e^{2\pi i \eta k}\, \overline{\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}(\eta\, B^{\ell}_{(1)} A^{j}_{(1)})}\, d\eta.
\end{aligned}$$

Using the expression of $\widehat{P_{\mathcal{M}_{h_j}} \mathcal{T}_j}$ computed in Proposition 4, as in Proposition 4 we can then write $\gamma_{j,\ell,k} = \gamma^{(1)}_{j,\ell,k} + \gamma^{(2)}_{j,\ell,k}$ where

$$\gamma^{(1)}_{j,\ell,k} = 2^{\frac{11}{2}j} \int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1 + \eta_2))\, V(\tfrac{\eta_2}{\eta_1}) \int_{B_{\Delta_0}} \int_a^b e^{2\pi i \eta B^{\ell}_{(1)} A^{j}_{(1)} (x + (f(u), u))}$$

$$\times \quad \mathbb{1}_{h_j}(x+(f(u),u))\, g(u)\, du\, \overline{\check{W}}(2^{2j}x)\, dx\, e^{2\pi i\eta k} d\eta \tag{43}$$

$$\gamma^{(2)}_{j,\ell,k} = 2^{\frac{11}{2}j}\int_{\widehat{\mathbb{R}}^2} W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1}) \int_{B^c_{\Delta_0}}\int_a^b e^{2\pi i\eta B^{\ell}_{(1)}A^j_{(1)}(x+(f(u),u))}$$
$$\times \quad \mathbb{1}_{h_j}(x+(f(u),u))\, g(u)\, du\, \overline{\check{W}}(2^{2j}x)\, dx\, e^{2\pi i\eta k} d\eta$$

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

Using the same argument of Proposition 4 for the wavelet case, it follows that $\sum_{(\ell,k)\in M_j} |\gamma^{(2)}_{j,\ell,k}| = o(2^j)$. Thus to prove (41) it remains to show that, for $h_j = o(2^{-\frac{3}{4}j})$, we have $\sum_{(\ell,k)\in M_j} |\gamma^{(1)}_{j,\ell,k}| = o(2^j)$.

As in the proof of Lemma 4, we may assume that $f(0) = f'(0) = 0$, and that $a=-\epsilon, b=\epsilon$. So $f(u)\simeq f''(c)u^2/3$, where $c\in[-\epsilon,\epsilon]$. Since $|f''(u)| \le \frac{1}{2}M$ for some $M>0$ 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 $|\ell|\le 2^{j/4}$.

From (43), using Lemma 8, we have that

$$\gamma^{(1)}_{j,\ell,k} = 2^{\frac{11}{2}j}\int_{B_{\Delta_0}}\int_a^b\int_{\widehat{\mathbb{R}}^2} L\Big(W(\eta_1, 2^{-j}(\ell\eta_1+\eta_2))\, V(\tfrac{\eta_2}{\eta_1})\Big)$$
$$\times\ L^{-1}\Big(e^{2\pi i\eta\left(k+B^{\ell}_1A^j_1(x+(f(u),u))\right)}\Big) d\eta\, \mathbb{1}_{h_j}(x+(f(u),u)) g(u)\, du\, \overline{\check{W}}(2^{2j}x)\, dx.$$

Observing that there is a constant C independent of j and x such that

$$\int_a^b \mathbb{1}_{h_j}(x+(f(u),u))\, |g(u)|\, du \le C\, h_j,$$

and that

$$k+B^{\ell}_1A^j_1(x+(f(u),u)) = \big(k_1+2^{2j}(x_1+f(u))+2^j\ell(x_2+u),\, k_2+2^j(x_2+u)\big),$$

an argument similar to Proposition 4 gives that

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

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 η 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 $\phi'_u = \frac{\partial}{\partial u}\phi$. Note that, in the integral (43), $\frac{1}{16} \le |\eta_1| \le \frac{1}{2}$ 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, ℓ such that

$$\begin{aligned}|\phi'_u(\eta, x, u)| &\ge |\eta_1| 2^j \left(|\ell| - 2^j |f'(u)| - |\frac{\eta_2}{\eta_1}|\right) \\ &\ge |\eta_1| 2^j \left(|\ell| - \frac{1}{3} 2^{j/4} - \frac{1}{6} 2^{j/4}\right) \\ &\ge C\, 2^j |\ell|. \end{aligned} \tag{44}$$

To estimate $\gamma^{(1)}_{j,\ell,k}$, for fixed $|x| \le 2^{-(2-\Delta_0)j}$ and η, we examine the integral

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

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

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

where

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

Correspondingly, we have $\gamma^{(1)}_{j,\ell,k} = \gamma^{(1,1)}_{j,\ell,k} + \gamma^{(1,2)}_{j,\ell,k} + \gamma^{(1,3)}_{j,\ell,k}$, 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 8, where L is given by (9), we have that

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

Using inequality (44) and the fact that $2^{\frac14 j}\le|\ell|\le 2^j$, $\frac{1}{16}\le|\eta_1|\le\frac12$, and $|f'(h_j-x_2)|<\frac13 2^{-\frac34 j}$, a direct computation shows that there is a uniform constant C, independent of j,ℓ, such that

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

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

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

$$\begin{aligned}&\sum_{2^{\frac14 j}\le|\ell|\le 2^j}\ \sum_{k\in\mathbb{Z}^2}|\gamma^{(1,1)}_{j,\ell,k}|\\ &\le\ C\,2^{\frac92 j}\int_{B_{\Delta_0}}\sum_{2^{\frac14 j}\le|\ell|\le 2^j}|\ell|^{-1}\sum_{k\in\mathbb{Z}^2}\left(1+(k_1+2^{2j}(x_1+f(h_j-x_2))+2^j\ell h_j)^2\right)^{-1}\\ &\times\left(1+(k_2+2^j h_j)^2\right)^{-1}|\check{W}(2^{2j}(x))|\,dx\\ &\le C\,2^{\frac12 j}\sum_{2^{\frac14 j}\le|\ell|\le 2^j}|\ell|^{-1}\int_{|x|\le 2^{\Delta_0)j}}|\check{W}(x)|\,dx\\ &\le C\,j\,2^{\frac12 j}=o(2^j).\end{aligned}$$

A very similar argument shows that $\sum_{2^{\frac14 j}\le|\ell|\le 2^j}\sum_{k\in\mathbb{Z}^2}|\gamma^{(1,2)}_{j,\ell,k}|=o(2^j)$.

Finally, for the analysis of $\gamma^{(1,3)}_{j,\ell,k}$, applying again Lemma 8 as above, we have that

$$\begin{aligned}\gamma^{(1,3)}_{j,\ell,k} &= \frac{2^{\frac{11}{2}j}}{2\pi i}\int_{B_{\Delta_0}}\int_{-h_j-x_2}^{h_j-x_2}\int_{\widehat{\mathbb{R}}^2} L\left(W(\eta_1,2^{-j}(\ell\eta_1+\eta_2))\,V(\tfrac{\eta_2}{\eta_1})\left(\frac{g(u)}{\phi'_u(\eta,x,u)}\right)'_u\right)\\ &\quad\times L^{-1}\left(e^{-2\pi i\eta\cdot\left(k-\left(2^{2j}(x_1+f(u))+\ell 2^j(x_2+u), 2^j(x_2+u)\right)\right)}\right)d\eta\,du\,\overline{\check{W}(2^{2j}x)}\,dx.\end{aligned}\tag{45}$$

We observe that

$$\left[\frac{g(u)}{\phi'_u(\eta,x,u)}\right]'_u = -\frac{\phi''_{u^2}(\eta,x,u)\,g(u)}{(\phi'_u(\eta,x,u))^2} + \frac{g'(u)}{\phi'_u(\eta,x,u)}. \tag{46}$$

As above, in the integral (45), 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 $M <$. 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, ℓ such that

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

Also, from (44) we have that

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

Thus, applying these observations in (46), 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, ℓ 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

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

Using this estimate in (45), we have that

$$\sum_{k\in\mathbb{Z}^2} |\gamma^{(1,3)}_{j,\ell,k}| \le C\,|\ell|^{-1}2^{\frac{11}{2}j}\int_{|x|\le 2^{-(2-\Delta_0)j}} |\check{W}(2^{2j}x)|\,dx \le C\,|\ell|^{-1}h_j\,2^{\frac{3}{2}j}.$$

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 (25) derived above.

We remark that, in the integral of (25), we have $|f(u)| \le 1$ for all $u \in [a,b] \subset [-\epsilon,\epsilon]$, with ϵ small. For each j,ℓ, we set $K_{j,\ell} = \{k \in \mathbb{Z}^2 : |k_1| \le 2^{2j+2}, |k_2| \le 2^{j+1}\}$ and $G_{j,\ell} = \{k \in \mathbb{Z}^2 : (\ell,k) \in I^c_j\}$. 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

$|f(u)| \leq 1$, $|u| \leq \epsilon$). Hence it follows from (25) that, for any $N \in \mathbb{N}$, there is a constant C_N such that

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

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

$$\sum_{k \in K^c_{j,\ell}} |\beta_{j,\ell,k}| \leq C\, 2^{\frac{3}{2}j}\, 2^{-3j} = C\, 2^{-\frac{3}{2}j}. \tag{47}$$

We can write

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

Since $k \in G_{j,\ell}$ means $(\ell, k) \in I^c_j$ 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,

$$\sum_{|\ell| \leq 2^j} \sum_{k \in G_{j,\ell} \bigcap K_{j,\ell}} |\alpha_{j,\ell,k}| \leq C 2^j\, 2^{3j}\, 2^{-4j} = C = o(2^j). \tag{48}$$

Since $G_{j,\ell} \bigcap K^c_{j,\ell} \subset K^c_{j,\ell}$, the estimate (47) gives that

$$\sum_{|\ell| \leq 2^j} \sum_{k \in G_{j,\ell} \bigcap K^c_{j,\ell}} |\beta_{j,\ell,k}| \leq \sum_{|\ell| \leq 2^j} C\, 2^{-\frac{3}{2}j} \leq C 2^{-\frac{1}{2}j} = o(2^j). \tag{49}$$

Finally, since $\sum_{(\ell,k) \in M_j} |\gamma_{j,\ell,k}| = o(2^j)$ by (41), combining this estimate with (48) and (49), we have proved (42). □

Appendix A. Additional proofs

Lemma 8. *Let $f \in C^\infty_c(\mathbb{R}^2)$ 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*

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

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. By direct computation, writing $x = (x_1, x_2)$, we have

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

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 (A.1). Using these observations, by direct computation we have

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

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. □

Lemma 9. *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*

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

Proof. Recall that the distributional Fourier transform of $\mathbb{1}_{\mathcal{M}_h}$ is given by $\widehat{\mathbb{1}_{\mathcal{M}_h}}(\xi_1, \xi_2) = 2h \operatorname{sinc}(2\pi h \xi_2)\delta_1(\xi_1, \xi_2)$, where $\int\int_{\widehat{\mathbb{R}}^2} \delta_1(x_1, x_2)\phi(x_1, x_2) dx_1 dx_2 = \int_{\widehat{\mathbb{R}}} \phi(0, x_2)\, dx_2$. Thus

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

Acknowledgments. DL acknowledges support from NSF grants DMS 1720487 and 1720452.

References

[1] J.F. Cai, R. H. Cha, Z. Shen, Simultaneous cartoon and texture inpainting, Inverse Probl. Imag., 4 (2010), pp. 379-395.

[2] J.F. Cai, B. Dong, S. Osher, Z. Shen, Image restoration: Total variation, wavelet frames, and beyond, J. Amer. Math. Soc., 25 (2012), pp. 1033–1089.

[3] T.F. Chan, J. Shen, Mathematical models for local nontexture inpaintings, SIAM J. Appl. Math., 62(3) (2002), pp. 1019–1043.

[4] T. Chan, J. Shen, and H.-M. Zhou, Total variation wavelet inpainting, J. Math. Imaging Vision, 25 (2006), pp. 107–125

[5] Y. Chen, H. Hu, An Improved Method for Semantic Image Inpainting with GANs: Progressive Inpainting, Neural Processing Letters 20 (2018).

[6] B. Dong, H. Ji, J. Li, Z. Shen, Y. Xu., Wavelet framework based blind image inpainting, Appl. Comput. Harmon. Anal., 32 (2012), pp. 268-279.

[7] D. L. Donoho, G. Kutyniok, Microlocal analysis of the geometric separation problem, Comm. Pure Appl. Math., 66 (2013), pp. 1–47.

[8] M. Elad, J.L. Starck, P. Querre, D. L. Donoho, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), Appl. Comput. Harmon. Anal., 19 (2005), pp. 340-358.

[9] S. Esedoglu and J. Shen, Digital image inpainting by the Mumford-Shah-Euler image model, European J. Appl. Math., 13 (2002), pp. 353–370.

[10] O. Guleryuz, Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising - Part i, Theory, IEEE Trans. Image Process., 15 (2006), pp. 539–554.

[11] M. Genzel and G. Kutyniok, Asymptotic Analysis of Inpainting via Universal Shearlet Systems, SIAM J. Imaging Sciences, 7 (2014), pp. 2301–2339.

[12] K. Guo and D. Labate, Optimally sparse multidimensional representation using shearlets, SIAM J. Math. Anal., 39 (2007), pp. 298–318.

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

[14] K. Guo, and D. Labate, Optimally sparse representations of 3D data with C^2 surface singularities using Parseval frames of shearlets, SIAM J. Math. Anal., 44 (2012), pp. 851–886.

[15] K. Guo, and D. Labate, The construction of smooth Parseval frames of shearlets, Math. Model. Nat. Phenom., 8(1) (2013), pp. 82–105.

[16] K. Guo, and D. Labate, Geometric Separation in $\mathbb{R}^3$, J. Fourier Anal. Appl., 25(1) (2019), pp 108–130.

[17] K. Guo, and D. Labate, Characterization and analysis of edges in piecewise smooth functions, Appl. Comput. Harmon. Anal. 41(1) (2016), pp. 139–163.

[18] K. Guo, and D. Labate, Detection of singularities by discrete multiscale transforms, J. Geom. Anal., 28(3) (2018), pp 2102–2128.

[19] E. J. King, G. Kutyniok, X. Zhuang, Analysis of Inpainting via Clustered Sparsity and Microlocal Analysis, J. Math. Imag. Visi., 48 (2014), pp. 205–234.

[20] G. Kutyniok and D. Labate, Shearlets: Multiscale analysis for multivariate data, Birkhäuser, 2012

[21] G. Kutyniok, P. Petersen, Classification of edges using compactly supported shearlets Appl. Comput. Harmon. Anal., 42(2) (2017), pp. 245–293.

[22] L. Shen, Y. Xu, X. Zeng, Wavelet inpainting with the ℓ_0 sparse regularization, Appl. Comput. Harmon. Anal., 41(1) (2016), pp. 26–53.

[23] E. M. Stein, Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals, Princeton University Press, Princeton, 1993.

[24] U. Demir, U. Gozde, Patch-Based Image Inpainting with Generative Adversarial Networks, arXiv preprint arXiv:1803.07422 (2018).

[25] C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, H. Li, High-resolution image inpainting using multi-scale neural patch synthesis, CVPR, 1 (2017) pp. 3.

[26] R. A. Yeh, C. Chen, T.-Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, M. N. Do, Semantic image inpainting with deep generative models, CVPR, 2 (2017), pp. 4.