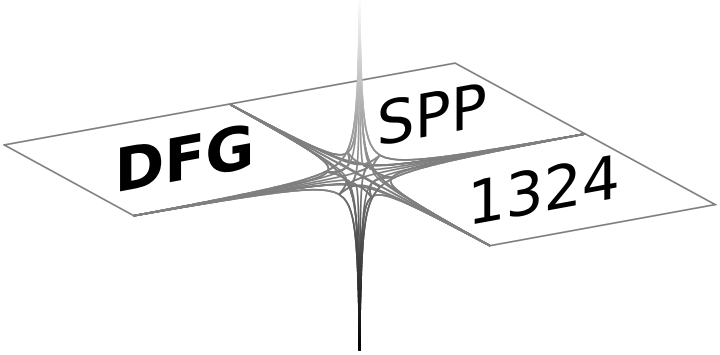
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen”

Shearlets on Bounded Domains

G. Kutyniok, W.-Q Lim

Preprint 54



Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

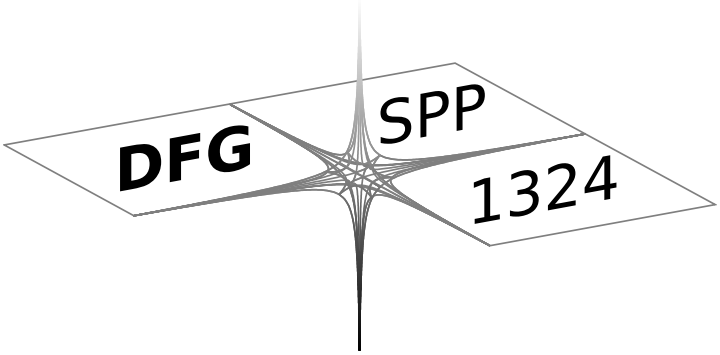
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen”

Shearlets on Bounded Domains

G. Kutyniok, W.-Q Lim

Preprint 54



The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

Shearlets on Bounded Domains

Gitta Kutyniok and Wang-Q Lim

Abstract Shearlet systems have so far been only considered as a means to analyze L^2-functions defined on $\mathbb{R}^2$, which exhibit curvilinear singularities. However, in applications such as image processing or numerical solvers of partial differential equations the function to be analyzed or efficiently encoded is typically defined on a non-rectangular shaped bounded domain. Motivated by these applications, in this paper, we first introduce a novel model for cartoon-like images defined on a bounded domain. We then prove that compactly supported shearlet frames satisfying some weak decay and smoothness conditions, when orthogonally projected onto the bounded domain, do provide (almost) optimally sparse approximations of elements belonging to this model class.

1 Introduction

It is by now well accepted that L^2-functions supported on the unit cube which are C^2 except for a C^2 discontinuity curve are a suitable model for images which are governed by edges. Of all directional representation systems which provide optimally sparse approximations of this model class, shearlet systems have distinguished themselves by the fact that they are the only system which provides a unified treatment of the continuum and digital setting, thereby making them particularly useful for both theoretical considerations as well as applications. However, most applications concern sparse approximations of functions on bounded domains, for instance, a numerical solver of a transport dominated equation could seek a solution

Gitta Kutyniok
Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany, e-mail: kutyniok@math.uos.de

Wang-Q Lim
Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany, e-mail: wlim@math.uos.de

on a polygonal shaped area. This calls for shearlet systems which are adapted to bounded domains while still providing optimally sparse expansions.

In this paper, we therefore consider the following questions:

(I) Which is a suitable model for a function on a bounded domain with curvilinear singularities?
(II) What is the 'correct' definition of a shearlet system for a bounded domain?
(III) Do these shearlet systems provide optimally sparse approximations of the model functions introduced in (I)?

In the sequel we will indeed provide a complete answer to those questions. These results push the door open for the usability of shearlet systems in all areas where 2D functions on bounded domains require efficient encoding.

1.1 Optimally Sparse Approximations of Cartoon-Like Images

The first complete model of cartoon-like images has been introduced in [1], the basic idea being that a closed C^2 curve separates smooth – in the sense of C^2 – functions. For the precise definition, we let $\rho : [0, 2\pi] \to \mathbb{R}^+$ be a C^2 radius function and define the set B by

$$B = \{x \in \mathbb{R}^2 : \|x\|_2 \leq \rho(\theta), x = (\|x\|_2, \theta) \text{ in polar coordinates}\}, \tag{1}$$

where

$$\sup |\rho''(\theta)| \leq \nu, \quad \rho \leq \rho_0 < 1. \tag{2}$$

This allows us to introduce $STAR^2(\nu)$, a class of sets B with C^2 boundaries ∂B and curvature bounded by ν, as well as $\mathscr{E}^2(\nu)$, a class of cartoon-like images.

Definition 1 ([1]). For $\nu > 0$, the set $STAR^2(\nu)$ is defined to be the set of all $B \subset [0,1]^2$ such that B is a translate of a set obeying (1) and (2). Further, $\mathscr{E}^2(\nu)$ denotes the set of functions f on $\mathbb{R}^2$ with compact support in $[0,1]^2$ of the form

$$f = f_0 + f_1 \chi_B,$$

where $B \in STAR^2(\nu)$ and $f_0, f_1 \in C^2(\mathbb{R}^2)$ with compact support in $[0,1]^2$ as well as $\sum_{|\alpha| \leq 2} \|D^\alpha f_i\|_\infty \leq 1$ for each $i = 0, 1$.

In [4], Donoho proved that the optimal rate which can be achieved under some restrictions on the representation system as well as on the selection procedure of the approximating coefficients is

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \quad \text{as } N \to \infty,$$

where f_N is the best N-term approximation.

1.2 Shortcomings of this Cartoon-Like Model Class

The first shortcoming of this model is the assumption that the discontinuity curve is C^2. Think, for instance, of an image, which pictures a building. Then the frames of the windows separate the dark interior of the windows from the presumably light color of the wall, however this frame is far from being C^2. Hence, a much more natural assumption would be to assume that the discontinuity curve is piecewise C^2.

The second shortcoming consists in the fact that the function is implicitly assumed to vanish on the boundary of $[0,1]^2$. More precisely, even if the function $f = f_0 + f_1\chi_B$ is non-zero on a section of positive measure of the boundary ∂B, this situation is not particularly treated at all. However, reminding ourselves of the very careful boundary treatment in the theory of partial differential equations, this situation should be paid close attention. Thus, a very natural approach to a careful handling of the boundary in a model for cartoon-like images seems to consist in regarding the boundary as a singularity curve itself.

The third and last shortcoming is the shape of the support $[0,1]^2$ of this model. Typically, in real-world situations the domain of 2D data can be very different from being a rectangle, and even a polygonal-shape model might not necessarily be sufficient. Examples to support this claim can be found, for instance, in fluid dynamics, where the flow can be supported on variously shaped domains. In this regard, a suitable model situation seems to be to allow the boundary to consist of any piecewise C^2 curve.

1.3 Our Model for Cartoon-Like Images on Bounded Domains

The model for cartoon-like images on bounded domains, which we now define, will indeed take all considerations from the previous subsection into account. For an illustration, we refer to Figure 1.

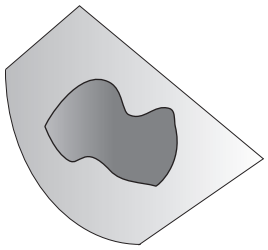


Fig. 1 Example of a function f belonging to our model class $\mathscr{E}^2_{\nu,L}(\Omega)$ of cartoon-like images on bounded domains.

We first introduce $STAR^2(\nu,L)$, a class of sets B with now piecewise C^2 boundaries ∂B and curvature on each piece bounded by ν. This will serve us for both modeling the bounded domain as well as modeling the discontinuity curve. For this, let $L \in \mathbb{Z}^+$ denote the number of C^2 pieces and let $\nu > 0$ be an upper estimate for the curvature on each piece. Then $B \in STAR^2(\nu,L)$, if B is a bounded subset of $[0,1]^2$ whose boundary ∂B is a simple closed curve and

$$\partial B = \bigcup_{i=1}^{L} \rho_i$$

where each curve ρ_i is parameterized by either $x_1 = E_i(x_2)$ or $x_2 = E_i(x_1)$ and $E_i \in C^2([a_i,b_i])$ such that

$$\max_{i=1,\dots,L} \max_{[a_i,b_i]} |E_i''| \leq \nu.$$

This allows us to introduce a model class of cartoon-like images in bounded domains. In accordance with modeling functions on bounded domains, we now consider functions defined on $[0,1]^2$; its 'true' domain is brought into play by requiring these functions to be supported on $\Omega \subseteq (0,1)^2$, which we model as piecewise C^2 bounded. This ensures that we treat $\partial\Omega$ as a singularity curve, which would not have been possible when defining the model on Ω itself.

Definition 2. For $\nu > 0$ and $L \in \mathbb{Z}^+$, let $\Omega, B \in STAR^2(\nu,L)$ be such that $B \subset \Omega^\circ$, where Ω° denotes the interior of the set Ω, and $\Omega \subset (0,1)^2$. Then, $\mathscr{E}^2_{\nu,L}(\Omega)$ denotes the set of functions f on $[0,1]^2$ with compact support in Ω of the form

$$f = f_0 + f_1 \chi_B,$$

where $f_0, f_1 \in C^2([0,1]^2)$ with compact support in Ω and $\sum_{|\alpha|\leq 2} \|D^\alpha f_i\|_\infty \leq 1$ for each $i = 0,1$.

Later it will become important to analyze the points on boundaries of sets in $STAR^2(\nu,L)$, in which the boundary is not C^2. For these points, we will employ the following customarily used notion.

Definition 3. For $\nu > 0$ and $L \in \mathbb{Z}^+$, let $B \in STAR^2(\nu,L)$. Then a point $x_0 \in \partial B$ will be called a *corner point*, if ∂B is not C^2 in x_0.

Since the model $\mathscr{E}^2_{\nu,L}(\Omega)$, while containing the previous model $\mathscr{E}^2_\nu$ as a special case, is considerably more complicated, we would like to make the reader aware of the fact that it is now not clear at all whether the optimal approximation rate is still

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \quad \text{as } N \to \infty.$$

1.4 Review of Shearlets

The directional representation system of *shearlets* has recently emerged – a first introduction dates back to 2005 in [17] – and rapidly gained attention due to the fact that, in contrast to other proposed directional representation systems, shearlets provide a unified treatment of the continuum and digital world similar to wavelets. We refer to, e.g., [7, 13] for the continuum theory, [16, 6, 18] for the digital theory, and [8, 5] for recent applications. Shearlets are scaled according to a parabolic scaling law and exhibit directionality by parameterizing slope by shearing, the later being the secret which allows the aforementioned unified treatment in contrast to rotation. Thus shearlets are associated with three parameters: scale, orientation, and position. A precise definition will be given in Section 2.

A few months ago, the theory of shearlets focussed entirely on band-limited generators although precise spatial localization is evidently highly desirable for, e.g., edge detection. Recently, motivated by this desideratum, compactly supported shearlets were studied by Kittipoom and the two authors. It was shown that a large class of compactly supported shearlets generate a frame for $L^2(\mathbb{R}^2)$ with controllable frame bounds alongside with several explicit constructions [10]. By the two authors it was then proven in [15] that a large class of these compactly supported shearlet frames does in fact provide (almost) optimally sparse approximations of functions in $\mathscr{E}_\nu^2$ in the sense of

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3 \quad \text{as } N \to \infty.$$

It should be mentioned that although the optimal rate is not completely achieved, the log-factor is typically considered negligible compared to the N^{-2}-factor, wherefore the term 'almost optimal' has been adopted into the language.

1.5 Surprising Result

We now aim to analyze the ability of shearlets to sparsely approximate elements of the previously introduced model for cartoon-like images on bounded domains, $\mathscr{E}_{\nu,L}^2(\Omega)$. For this, we first need to define shearlet systems for functions in $L^2(\Omega)$. Assume we are given a (compactly supported) shearlet frame for $L^2(\mathbb{R}^2)$. The most crude approach to transform this into a shearlet system defined on $L^2(\Omega)$, where $\Omega \in STAR^2(\nu, L)$, is to just truncate each element at the boundary of Ω. Since it is well known in classical frame theory that the orthogonal projection of a frame onto a subspace does not change the frame bounds (cf. [2]), this procedure will result in a (compactly supported) shearlet frame for $L^2(\Omega)$ with the same frame bounds as before.

We now apply this procedure to the family of compactly supported shearlet frames for $L^2(\mathbb{R}^2)$, which yielded (almost) optimally sparse approximations of functions in $\mathscr{E}_\nu^2$ (see [15, Thm. 1.3]). The main result of this paper then proves that the

resulting family of shearlet frames – now regarded as a system on $[0,1]^2$ with compact support in Ω – again provides (almost) optimally sparse approximations now of elements from our model of cartoon-like images on bounded domains $\mathscr{E}^2_{\nu,L}(\Omega)$ in the sense of

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3 \quad \text{as } N \to \infty.$$

The precise statement is phrased in Theorem 1 in Section 3.

This result is quite surprising in two ways:

- *Surprise 1*. Regarding a log-factor as negligible – a customarily taken viewpoint –, the previous result shows that even for our much more sophisticated model of cartoon-like images on bounded domains the *same* optimal sparse approximation rate as for the simple model detailed in Subsection 1.1 can be achieved. This is even more surprising taking into account that our model contains point singularities at the corner points of the singularity curves. Naively, one would expect that these should worsen the approximation rate. However, observing that 'not too many' shearlets intersect these 'sparsely occurring' points unravels this mystery.
- *Surprise 2*. Orthogonally projecting a shearlet system onto the considered bounded domain, thereby merely truncating it, seems an exceptionally crude approach to derive shearlets for a bounded domain. However, these 'modified' shearlet systems are indeed sufficient to achieve the optimal rate and no sophisticated adaptions are required, which is of significance for deriving fast algorithmic realizations.

1.6 Main Contributions

The main contributions of this paper are two-fold. Firstly, we introduce $\mathscr{E}^2_{\nu,L}(\Omega)$ as a suitable model for a function on a bounded domain with curvilinear singularities. Secondly, we show that the 'crude' approach towards a shearlet system on a bounded domain by simply orthogonally projecting still provides optimally sparse approximations of elements belonging to our model class $\mathscr{E}^2_{\nu,L}(\Omega)$.

We should mention that although not formally stated the idea of one piecewise C^2 discontinuity curve in a model for functions on $\mathbb{R}^2$ as an extension of Definition 1 is already lurking in [1]. Also a brief sketch of proof of (almost) optimally sparse approximations of curvelets is contained therein. These ideas are however very different from ours in two aspects. First of all, our goal is a suitable model for functions on bounded domains exhibiting discontinuity curves and also treating the boundary of the domain as a singularity curve. And secondly, in this paper we consider compactly supported shearlets – hence elements with superior spatial localization properties in contrast to the (band-limited) curvelets – which allows an elegant proof of the sparse approximation result in addition to a simplified treatment of the corner points.

1.7 Outline

In Section 2, after recalling the definition of shearlet systems, we introduce shearlet systems on bounded domains, thereby focussing in particular on compactly supported shearlet frames. The precise statement of our main result is presented in Section 3 together with a road map to its proof. The proof itself is then carried out in Section 4. Finally, in Section 5, we discuss our results and possible extensions of it.

2 Compactly Supported Shearlets

We first review the main notions and definitions related to shearlet theory, focussing in particular on compactly supported generators. For more details we would like to refer the interested reader to the survey paper [14]. Then we present our definition of shearlet systems on a bounded domain $\Omega \in STAR^2(\nu, L)$.

2.1 Compactly Supported Shearlet Frames for $L^2(\mathbb{R}^2)$

Shearlets are scaled according to a parabolic scaling law encoded in the *parabolic scaling matrices* A_{2^j} or $\tilde{A}_{2^j}$, $j \in \mathbb{Z}$, and exhibit directionality by parameterizing slope encoded in the *shear matrices* S_k, $k \in \mathbb{Z}$, defined by

$$A_{2^j} = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix} \qquad \text{or} \qquad \tilde{A}_{2^j} = \begin{pmatrix} 2^{j/2} & 0 \\ 0 & 2^j \end{pmatrix}$$

and

$$S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

respectively.

We next partition the frequency plane into four cones $\mathscr{C}_1 - \mathscr{C}_4$. This allow the introduction of shearlet systems which treat different slopes equally in contrast to the shearlet group-based approach. We though wish to mention that historically the shearlet group-based approach was developed first due to very favorable theoretical properties and it still often serves as a system for developing novel analysis strategies (see, for instance, [11]).

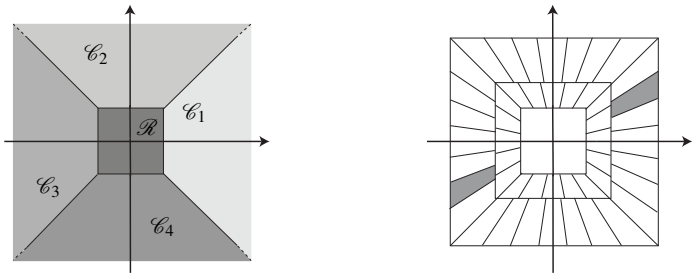
The four cones $\mathscr{C}_1 - \mathscr{C}_4$ are now defined by

$$\mathscr{C}_\iota = \begin{cases} \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \geq 1, |\xi_2/\xi_1| \leq 1\} : \iota = 1, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \geq 1, |\xi_1/\xi_2| \leq 1\} : \iota = 2, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \leq -1, |\xi_2/\xi_1| \leq 1\} : \iota = 3, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \leq -1, |\xi_1/\xi_2| \leq 1\} : \iota = 4, \end{cases}$$

and a centered rectangle

$$\mathscr{R} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \|(\xi_1, \xi_2)\|_\infty < 1\}.$$

For an illustration, we refer to Figure 2(a).



(a) (b)

Fig. 2 (a) The cones $\mathscr{C}_1 - \mathscr{C}_4$ and the centered rectangle $\mathscr{R}$ in frequency domain. (b) The tiling of the frequency domain induced by a cone-adapted shearlet system, where the (essential) support of the Fourier transform of one shearlet generator is exemplary high-lighted.

The rectangle $\mathscr{R}$ corresponds to the low frequency content of a signal and is customarily represented by translations of some scaling function. Anisotropy comes into play when encoding the high frequency content of a signal which corresponds to the cones $\mathscr{C}_1 - \mathscr{C}_4$, where the cones $\mathscr{C}_1$ and $\mathscr{C}_3$ as well as $\mathscr{C}_2$ and $\mathscr{C}_4$ are treated separately as can be seen in the following

Definition 4. For some sampling constant $c > 0$, the *cone-adapted shearlet system* $SH(\phi, \psi, \tilde{\psi}; c)$ generated by a *scaling function* $\phi \in L^2(\mathbb{R}^2)$ and *shearlets* $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is defined by

$$SH(\phi, \psi, \tilde{\psi}; c) = \Phi(\phi; c) \cup \Psi(\psi; c) \cup \tilde{\Psi}(\tilde{\psi}; c),$$

where

$$\Phi(\phi; c) = \{\phi_m = \phi(\cdot - cm) : m \in \mathbb{Z}^2\},$$

$$\Psi(\psi; c) = \{\psi_{j,k,m} = 2^{3j/4}\psi(S_k A_{2^j} \cdot - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\},$$

and

$$\tilde{\Psi}(\tilde{\psi}; c) = \{\tilde{\psi}_{j,k,m} = 2^{3j/4}\tilde{\psi}(S_k^T \tilde{A}_{2^j} \cdot - cm) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\}.$$

The tiling of frequency domain induced by $SH(\phi, \psi, \tilde{\psi}; c)$ is illustrated in Figure 2(b). From this illustration, the anisotropic footprints of shearlets contained in $\Psi(\psi; c)$ and $\tilde{\Psi}(\tilde{\psi}; c)$ can clearly be seen. The corresponding anisotropic footprints of shearlets *in spatial domain* are of size $2^{-j/2}$ times 2^{-j}.

The reader should keep in mind that although not indicated by the notation, the functions ϕ_m, $\psi_{j,k,m}$, and $\tilde{\psi}_{j,k,m}$ all depend on the sampling constant c. For the sake of brevity, we will often write ψ_λ and $\tilde{\psi}_\lambda$, where $\lambda = (j,k,m)$ index scale, shear, and position. For later use, we further let Λ_j and $\tilde{\Lambda}_j$ be the indexing sets of shearlets in $\Psi(\psi;c)$ and $\tilde{\Psi}(\tilde{\psi};c)$ at scale j, respectively, i.e.,

$$\Psi(\psi;c) = \{\psi_\lambda : \lambda \in \Lambda_j, j = 0,1,\ldots\} \text{ and } \tilde{\Psi}(\tilde{\psi};c) = \{\tilde{\psi}_\lambda : \lambda \in \tilde{\Lambda}_j, j = 0,1,\ldots\}.$$

Finally, we define

$$\Lambda = \bigcup_{j=0}^{\infty} \Lambda_j \quad \text{and} \quad \tilde{\Lambda} = \bigcup_{j=0}^{\infty} \tilde{\Lambda}_j.$$

The shearlet systems $SH(\phi,\psi,\tilde{\psi};c)$ have already been very well studied with respect to their frame properties for $L^2(\mathbb{R}^2)$, and we would like to refer to results in [7, 12, 3]. It should be mentioned that those results typically concern frame properties of $\Psi(\psi;c)$, which immediately imply frame properties of $\tilde{\Psi}(\tilde{\psi};c)$ likewise, whereas numerous frame properties for the low-frequency part $\Phi(\phi;c)$ can be found in the wavelet literature. Combining those leads to frame properties of $SH(\phi,\psi,\tilde{\psi};c)$.

Recent results in [10] establish frame properties specifically for the case of spatially compactly supported shearlet systems, i.e., shearlet systems with compactly supported generators ϕ, ψ, and $\tilde{\psi}$ which lead to a shearlet system consisting of compactly supported elements. These results give sufficient conditions for the so-called t_q conditions to be satisfied. As one class of examples with 'good' frame bounds, generating shearlets ψ and $\tilde{\psi}$ were chosen to be separable, i.e., of the form $\psi_1(x_1)\cdot\psi_2(x_2)$ and $\psi_1(x_2)\cdot\psi_2(x_1)$, respectively, where ψ_1 is a wavelet and ψ_2 a scaling function both associated with some carefully chosen (maximally flat) low pass filter. The separability has in addition the advantage to lead to fast accompanying algorithms.

We wish to mention that there is a trade-off between *compactly support* of the shearlet generators, *tightness* of the associated frame, and *separability* of the shearlet generators. The known constructions of tight shearlet frames do not use separable generators, and these constructions can be shown to *not* be applicable to compactly supported generators. Tightness is difficult to obtain while allowing for compactly supported generators, but we can gain separability, hence fast algorithmic realizations. On the other hand, when allowing non-compactly supported generators, tightness is possible, but separability seems to be out of reach, which makes fast algorithmic realizations very difficult.

2.2 Compactly Supported Shearlet Frames for $L^2(\Omega)$

Let now $\Omega \in STAR^2(v,L)$ be a bounded domain as defined in Subsection 1.3. The main idea to introduce a shearlet frame for $L^2(\Omega)$, preferably with compactly sup-

ported elements, is to start with a compactly supported shearlet frame for $L^2(\mathbb{R}^2)$ and apply the orthogonal projection onto $L^2(\Omega)$ to each element. To make this mathematically precise, we let $P_\Omega : L^2(\mathbb{R}^2) \to L^2(\Omega)$ denote the orthogonal projection onto $L^2(\Omega)$.

Definition 5. Let $\Omega \in STAR^2(\nu, L)$. For some sampling constant $c > 0$, the *cone-adapted shearlet system* $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ for $L^2(\Omega)$ generated by a *scaling function* $\phi \in L^2(\mathbb{R}^2)$ and *shearlets* $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is defined by

$$SH_\Omega(\phi, \psi, \tilde{\psi}; c) = P_\Omega(\Phi(\phi; c) \cup \Psi(\psi; c) \cup \tilde{\Psi}(\tilde{\psi}; c)),$$

where $\Phi(\phi; c)$, $\Psi(\psi; c)$, and $\tilde{\Psi}(\tilde{\psi}; c)$ are defined as in Definition 4.

As a direct corollary from well known results in frame theory (see [2]), we obtain the following result, which clarifies frame properties for systems $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ to the extent to which they are known for systems $SH(\phi, \psi, \tilde{\psi}; c)$. In the sequel, we will usually regard $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ as a system defined on $[0,1]^2$ – in accordance with our model $\mathscr{E}^2_{\nu,L}(\Omega)$ – by which we simply mean extension by zero. This system will be sometimes referred to as the *extension of* $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ *to* $[0,1]^2$. The following result also provides frame properties of these systems.

Proposition 1. *Let $c > 0$, let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$, and let $\Omega \in STAR^2(\nu, L)$ with positive measure. Then the following conditions are equivalent.*

(i) The shearlet system $SH(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\mathbb{R}^2)$ with frame bounds A and B.
(ii) The shearlet system $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\Omega)$ with frame bounds A and B.
(iii) The extension of the shearlet system $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ to $[0,1]^2$ is a frame with frame bounds A and B for functions $L^2([0,1]^2)$ with compact support in Ω.

3 Optimal Sparsity of Shearlets on Bounded Domains

We now have all ingredients to formally state the result already announced in Subsection 1.5, which shows that even with the 'crude' construction of shearlets on bounded domains *and* the significantly more sophisticated model for cartoon-like images on bounded domains we still obtain (almost) optimally sparse approximations.

Theorem 1. *Let $c > 0$, and let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported. Suppose that, in addition, for all $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, the shearlet ψ satisfies*

(i) $|\hat{\psi}(\xi)| \leq C_1 \cdot \min(1, |\xi_1|^\alpha) \cdot \min(1, |\xi_1|^{-\gamma}) \cdot \min(1, |\xi_2|^{-\gamma})$, *and*
(ii) $\left| \frac{\partial}{\partial \xi_2} \hat{\psi}(\xi) \right| \leq |h(\xi_1)| \cdot \left(1 + \frac{|\xi_2|}{|\xi_1|}\right)^{-\gamma}$,

where $\alpha > 5$, $\gamma \geq 4$, $h \in L^1(\mathbb{R})$, and C_1 is a constant, and suppose that the shearlet $\tilde{\psi}$ satisfies (i) and (ii) with the roles of ξ_1 and ξ_2 reversed. Further, let $\nu > 0, L \in \mathbb{Z}^+$ and $\Omega \in STAR^2(\nu, L)$, and suppose that $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ forms a frame for $L^2(\Omega)$.

Then, the extension of the shearlet frame $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ to $[0,1]^2$ (cf. Subsection 2.2) provides (almost) optimally sparse approximations of functions $f \in \mathscr{E}^2_{\nu,L}(\Omega)$ in the sense that there exists some $C > 0$ such that

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3 \qquad \text{as } N \to \infty,$$

where f_N is the nonlinear N-term approximation obtained by choosing the N largest shearlet coefficients of f.

3.1 Architecture of the Proof of Theorem 1

Before delving into the proof in the following section, we present some preparation before as well as describe the architecture of the proof for clarity purposes.

Let now $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ satisfy the hypotheses in Theorem 1, and let $f \in \mathscr{E}^2_{\nu,L}(\Omega)$. We first observe that, without loss of generality, we might assume the scaling index j to be sufficiently large, since f as well as all frame elements in the shearlet frame $SH_\Omega(\phi, \psi, \tilde{\psi}; c)$ are compactly supported in spatial domain, hence a finite number does not contribute to the asymptotic estimate we aim for. In particular, this means that we do not need to take frame elements from $\Phi(\phi; c)$ into account. Also, we are allowed to restrict our analysis to shearlets $\psi_{j,k,m}$, since the frame elements $\widetilde{\psi}_{j,k,m}$ can be handled in a similar way.

We further observe that we can drive the analysis for the frame $SH(\phi, \psi, \tilde{\psi}; c)$ and for the domain $[0,1]^2$ instead, since, by hypothesis, Ω is contained in the interior of $[0,1]^2$, we treat the boundary of Ω as a singularity curve in $[0,1]^2$, and the frame properties are equal as shown in Proposition 1. In this viewpoint, the function to be sparsely approximated vanishes on $[0,1]^2 \setminus \Omega$.

Our main concern will now be to derive appropriate estimates for the shearlet coefficients $\{\langle f, \psi_\lambda \rangle : \lambda \in \Lambda\}$ of f. Letting $|\theta(f)|_n$ denote the nth largest shearlet coefficient $\langle f, \psi_\lambda \rangle$ in absolute value and exploring the frame property of $SH(\phi, \psi, \tilde{\psi}; c)$, we conclude that

$$\|f - f_N\|_2^2 \leq \frac{1}{A} \sum_{n > N} |\theta(f)|_n^2,$$

for any positive integer N, where A denotes the lower frame bound of the shearlet frame $SH(\phi, \psi, \tilde{\psi}; c)$. Thus, for the proof of Theorem 1, it suffices to show that

$$\sum_{n > N} |\theta(f)|_n^2 \leq C \cdot N^{-2} \cdot (\log N)^3 \qquad \text{as } N \to \infty. \tag{3}$$

To derive the anticipated estimate in (3), for any shearlet ψ_λ, we will study two separate cases:

- *Case 1*. The compact support of the shearlet ψ_λ does not intersect the boundary of the set B (or $\partial\Omega$), i.e., $\mathrm{supp}(\psi_\lambda) \cap (\partial B \cup \partial\Omega) = \emptyset$.
- *Case 2*. The compact support of the shearlet ψ_λ does intersect the boundary of the set B (or $\partial\Omega$), i.e., $\mathrm{supp}(\psi_\lambda) \cap (\partial B \cup \partial\Omega) \neq \emptyset$.

Notice that this exact distinction is only possible due to the spatial compact support of all shearlets in the shearlet frame.

Case 2 will then throughout the proof be further subdivided into the situations – which we now do not state precisely, but just give the reader the intuition behind them:

- *Case 2a*. The support of the shearlet does intersect only one C^2 curve in $\partial B \cup \partial\Omega$.
- *Case 2b*. The support of the shearlet does intersect at least two C^2 curves in $\partial B \cup \partial\Omega$.
  - *Case 2b-1*. The support of the shearlet does intersect $\partial B \cup \partial\Omega$ in a corner point.
  - *Case 2b-2*. The support of the shearlet does intersect two C^2 curves in $\partial B \cup \partial\Omega$ simultaneously, but does not intersect a corner point.

4 Proof of Theorem 1

In this section, we present the proof of Theorem 1, following the road map outlined in Subsection 3.1. We wish to mention that Case 1 and Case 2a are similar to the proof of (almost) optimally sparse approximations of the class $\mathscr{E}^2(v)$ using compactly supported shearlet frames in [15]. However, Case 2b differs significantly from it, since it, in particular, requires a careful handling of the corner points of ∂B and $\partial\Omega$.

In the sequel – since we are concerned with an asymptotic estimate – for simplicity we will often simply use C as a constant although it might differ for each estimate. Also all the results in the sequel are independent on the sampling constant $c > 0$, wherefore we now fix it once and for all.

4.1 Case 1: The Smooth Part

We start with Case 1, hence the smooth part. Without loss of generality, we can consider some $g \in C^2([0,1]^2)$ and estimate its shearlet coefficients. The following proposition, which is taken from [15], implies the rate for optimal sparsity. Notice that the hypothesis on ψ of the following result is implied by condition (i) in Theorem 1.

Proposition 2 ([15]). *Let $g \in C^2([0,1]^2)$, and let $\psi \in L^2(\mathbb{R}^2)$ be compactly supported and satisfy*

$$|\hat{\psi}(\xi)| \leq C_1 \cdot \min(1,|\xi_1|^{\alpha}) \cdot \min(1,|\xi_1|^{-\gamma}) \cdot \min(1,|\xi_2|^{-\gamma}) \text{ for all } \xi = (\xi_1,\xi_2) \in \mathbb{R}^2,$$

where $\gamma > 3$, $\alpha > \gamma + 2$, *and* C_1 *is a constant. Then, there exists some* $C > 0$ *such that*

$$\sum_{n>N} |\theta(g)|_n^2 \leq C \cdot N^{-2} \qquad \text{as } N \to \infty.$$

This settles Theorem 1 for this case.

4.2 Case 2: The Non-Smooth Part

Next, we turn our attention to the non-smooth part, and aim to estimate the shearlet coefficients of those shearlets whose spatial support intersects the discontinuity curve ∂B or the boundary of the domain Ω. One of the main means of the proof will be the partitioning of the unit cube $[0,1]^2$ into dyadic cubes, picking those which contain such an intersection, and estimating the associated shearlet coefficients. For this, we first need to introduce the necessary notational concepts.

For any scale $j \geq 0$ and any grid point $p \in \mathbb{Z}^2$, we let $Q_{j,p}$ denote the dyadic cube defined by

$$Q_{j,p} = [-2^{-j/2}, 2^{-j/2}]^2 + 2^{-j/2} p.$$

Further, let Q_j be the collection of those dyadic cubes $Q_{j,p}$ which intersect $\partial B \cup \partial \Omega$, i.e.,

$$Q_j = \{Q_{j,p} : Q_{j,p} \cap (\partial B \cup \partial \Omega) \neq \emptyset, p \in \mathbb{Z}^2\}.$$

Of interest to us is also the set of shearlet indices, which are associated with shearlets intersecting the discontinuity curve inside some $Q_{j,p} \in Q_j$, hence, for $j \geq 0$ and $p \in \mathbb{Z}^2$ with $Q_{j,p} \in Q_j$, we will consider the index set

$$\Lambda_{j,p} = \{\lambda \in \Lambda_j : \operatorname{supp}(\psi_\lambda) \cap Q_{j,p} \cap (\partial B \cup \partial \Omega) \neq \emptyset\}.$$

Finally, for $j \geq 0$, $p \in \mathbb{Z}^2$, and $0 < \varepsilon < 1$, we define $\Lambda_{j,p}(\varepsilon)$ to be the index set of shearlets ψ_λ, $\lambda \in \Lambda_{j,p}$, such that the magnitude of the corresponding shearlet coefficient $\langle f, \psi_\lambda \rangle$ is larger than ε and the support of ψ_λ intersects $Q_{j,p}$ at the jth scale, i.e.,

$$\Lambda_{j,p}(\varepsilon) = \{\lambda \in \Lambda_{j,p} : |\langle f, \psi_\lambda \rangle| > \varepsilon\},$$

and we define $\Lambda(\varepsilon)$ to be the index set for shearlets so that $|\langle f, \psi_\lambda \rangle| > \varepsilon$ across all scales j, i.e.,

$$\Lambda(\varepsilon) = \bigcup_{j,p} \Lambda_{j,p}(\varepsilon).$$

The expert reader will have noticed that in contrast to the proofs in [1] and [9], which also split the domain into smaller scale boxes, we do not apply a weight function to obtain a smooth partition of unity. In our case, this is not necessary due to the spatial compact support of the frame elements. Finally we set

$$S_{j,p} = \bigcup_{\lambda \in \Lambda_{j,p}} \operatorname{supp}(\psi_\lambda),$$

which is contained in a cubic window of size $C \cdot 2^{-j/2}$ by $C \cdot 2^{-j/2}$, hence is of asymptotically the same size as $Q_{j,p}$. As mentioned earlier, we may assume that j is sufficiently large so that it is sufficient to consider the following two cases:

- *Case 2a*. There is only one edge curve $\Gamma_1 \subset \partial B$ (or $\partial\Omega$) which can be parameterized by $x_1 = E(x_2)$ (or $x_2 = E(x_1)$) with $E \in C^2$ inside $S_{j,p}$. For any $\lambda \in \Lambda_{j,p}$, there exists some $\hat{x} = (\hat{x}_1, \hat{x}_2) \in Q_{j,p} \cap \operatorname{supp}(\psi_\lambda) \cap \Gamma_1$.
- *Case 2b*. There are two edge curves $\Gamma_1, \Gamma_2 \subset \partial B$ (or $\partial\Omega$) which can be parameterized by $x_1 = E(x_2)$ (or $x_2 = E(x_1)$) with $E \in C^2$ inside $S_{j,p}$. For any $\lambda \in \Lambda_{j,p}$, there exist two distinct points $\hat{x} = (\hat{x}_1, \hat{x}_2)$ and $\hat{y} = (\hat{y}_1, \hat{y}_2)$ such that $\hat{x} \in Q_{j,p} \cap \operatorname{supp}(\psi_\lambda) \cap \Gamma_1$ and $\hat{y} \in Q_{j,p} \cap \operatorname{supp}(\psi_\lambda) \cap \Gamma_2$.

In the sequel, we only consider the edge curve ∂B to analyze shearlet coefficients associated with the non-smooth part, since the boundary of the domain Ω can be handled in a similar way; see also our elaboration on the fact that WLOG we can consider the approximation on $[0,1]^2$ rather than Ω in Subsection 3.1.

4.2.1 Case 2a: The Non-Smooth Part

This part was already studied in [15], where an (almost) optimally sparse approximation rate by the class of compactly supported shearlet frames $SH(\phi, \psi, \tilde{\psi}; c)$ under consideration was proven, and we refer to [15] for the precise argumentation. For intuition purposes as well as for later usage, we though state the key estimate, which implies (almost) optimally sparse approximation for Case 2a:

Proposition 3 ([15]). *Let $\psi \in L^2(\mathbb{R}^2)$ be compactly supported and satisfy the conditions (i), (ii) in Theorem 1 and assume that, for any $\lambda \in \Lambda_{j,p}$, there exists some $\hat{x} = (\hat{x}_1, \hat{x}_2) \in Q_{j,p} \cap \operatorname{supp}(\psi_\lambda) \cap \partial B$. Let s be the slope[1] of the tangent to the edge curve ∂B at $(\hat{x}_1, \hat{x}_2)$, i.e.,*

- *$s = E'(\hat{x}_2)$, if ∂B is parameterized by $x_1 = E(x_2)$ with $E \in C^2$ in $S_{j,p}$,*
- *$s = (E'(\hat{x}_1))^{-1}$, if ∂B is parameterized by $x_2 = E(x_1)$ with $E \in C^2$ in $S_{j,p}$, and*
- *$s = \infty$, if ∂B is parameterized by $x_2 = E(x_1)$ with $E'(\hat{x}_1) = 0$ and $E \in C^2$ in $S_{j,p}$.*

Then, there exists some $C > 0$ such that

$$|\langle f, \psi_\lambda \rangle| \leq C \cdot 2^{-\frac{9}{4} j}, \qquad \text{if } |s| > \frac{3}{2} \text{ or } |s| = \infty, \tag{4}$$

and

[1] Notice that here we regard the slope of the tangent to a curve $(E(x_2), x_2)$, i.e., we consider s of a curve $x_1 = s x_2 + b$, say. For analyzing shearlets $\tilde{\psi}_{j,k,m}$, the roles of x_1 and x_2 would need to be reversed.

$$|\langle f, \psi_\lambda \rangle| \leq C \cdot \frac{2^{-\frac{3}{4}j}}{|k+2^{j/2}s|^3}, \qquad \text{if } |s| \leq 3. \tag{5}$$

Similar estimates with ∂B substituted by $\partial\Omega$ hold if, for any $\lambda \in \Lambda_{j,p}$, there exists some $\hat{x} = (\hat{x}_1, \hat{x}_2) \in Q_{j,p} \cap \mathrm{supp}(\psi_\lambda) \cap \partial\Omega$.

4.2.2 Case 2b: The Non-Smooth Part

Letting $\varepsilon > 0$, our goal will now be to first estimate $|\Lambda_{j,p}(\varepsilon)|$ and, based on this, derive an estimate for $|\Lambda(\varepsilon)|$. WLOG we might assume $\|\psi\|_1 \leq 1$, which implies

$$|\langle f, \psi_\lambda \rangle| \leq 2^{-\frac{3}{4}j}.$$

Hence, for estimating $|\Lambda_{j,p}(\varepsilon)|$, it is sufficient to restrict our attention to scales $j \leq \frac{4}{3}\log_2(\varepsilon^{-1})$.

As already announced before, we now split Case 2b into the following two subcases:

- *Case 2b-1*. The shearlet ψ_λ intersects a corner point, in which two C^2 curves Γ_1 and Γ_2, say, meet (see Figure 3 (a)).
- *Case 2b-2*. The shearlet ψ_λ intersects two edge curves Γ_1 and Γ_2, say, simultaneously, but it does not intersect a corner point (see Figure 3 (b)).

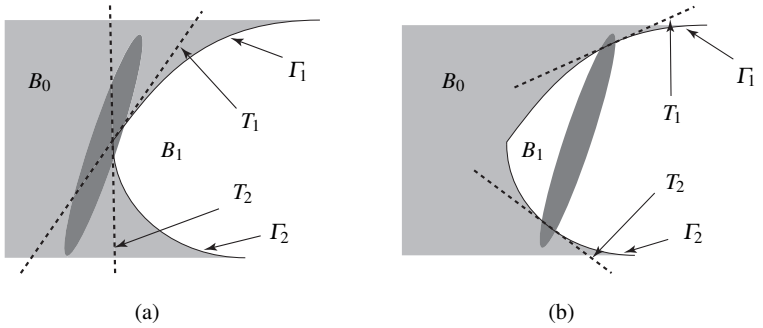


Fig. 3 (a) A shearlet ψ_λ intersecting a corner point where two edge curves Γ_1 and Γ_2 meet. T_1 and T_2 are tangents to the edge curves Γ_1 and Γ_2 in this corner point. (b) A shearlet ψ_λ intersecting two edge curves Γ_1 and Γ_2 which are a part of the boundary of sets B_0 and B_1. T_1 and T_2 are tangents to the edge curves Γ_1 and Γ_2 in points contained in the support of ψ_λ.

Case 2b-1. We first consider *Case 2b-1*. In this case, by a counting argument, it follows that

$$|\Lambda_{j,p}(\varepsilon)| \leq C \cdot 2^{j/2}.$$

Since there are only finitely many corner points with its number not depending on scale $j \geq 0$, we have

$$|\Lambda(\varepsilon)| \leq C \cdot \sum_{j=0}^{\frac{4}{3}\log_2(\varepsilon^{-1})} 2^{j/2} \leq C \cdot \varepsilon^{-\frac{2}{3}}.$$

The value $\varepsilon > 0$ can be written as a function of the total number N of coefficients, which yields $\varepsilon(N) \leq C \cdot N^{-\frac{3}{2}}$. This implies that

$$\sum_{n>N} |\theta(f)|_n^2 \leq C \cdot N^{-2},$$

and the optimal sparse approximation rate is proven for *Case 2b-1*.

Case 2b-2. Next, we consider *Case 2b-2*. In this case, WLOG, we might assume that, for any $\lambda \in \Lambda_{j,p}$, there exist two distinct points $\hat{x} = (\hat{x}_1, \hat{x}_2), \hat{y} = (\hat{y}_1, \hat{y}_2)$ such that $\hat{x} \in Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_1$ and $\hat{y} \in Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_2$, and the two edge curves Γ_1 and Γ_2 are parameterized by $x_1 = E(x_2)$ (or $x_2 = E(x_1)$) with $E \in C^2$ inside $S_{j,p}$. We can then write the function $f \in \mathscr{E}_{\nu,L}^2(\Omega)$ as

$$f_0 \chi_{B_0} + f_1 \chi_{B_1} = (f_0 - f_1)\chi_{B_0} + f_1 \qquad \text{on } S_{j,p},$$

where $f_0, f_1 \in C^2([0,1]^2)$ and B_0, B_1 are two disjoint subsets of $[0,1]^2$ (see Figure 3). By Proposition 2, the rate for optimal sparse approximation is achieved for the smooth part f_1. Thus, it is sufficient to consider $f = g_0 \chi_{B_0}$ with $g_0 = f_0 - f_1 \in C^2([0,1]^2)$.

Assume now that the tangents to the edge curves Γ_1 and Γ_2 at the points $\hat{x}$ and $\hat{y}$ are given by the equations

$$T_1 : x_1 + a_1 = s_1(x_2 + b_1) \quad \text{and} \quad T_2 : x_1 + a_2 = s_2(x_2 + b_2),$$

respectively, i.e., s_1 and s_2 are the slopes of the tangents to the edge curves Γ_1 and Γ_2 at $\hat{x}$ and $\hat{y}$, respectively. If the curve Γ_i, $i = 1, 2$, is parameterized by $x_2 = E(x_1)$ with $E'(\hat{x}_1) = 0$, we let $s_i = \infty$ so that the tangent is given by $x_2 = -b_i$ in this case.

Next, for fixed scale j and shear index k, let $N_{j,k}^1(Q_{j,p})$ denote the number of shearlets ψ_λ intersecting Γ_1 in $Q_{j,p}$, i.e.,

$$N_{j,k}^1(Q_{j,p}) = |\{\lambda = (j,k,m) : Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_1 \neq \emptyset\}|,$$

let $N_{j,k}^2(Q_{j,p})$ denote the number of shearlets ψ_λ intersecting Γ_2 in $Q_{j,p}$, i.e.,

$$N_{j,k}^2(Q_{j,p}) = |\{\lambda = (j,k,m) : Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_2 \neq \emptyset\}|,$$

and let $N_{j,k}(Q_{j,p})$ denote the number of shearlets ψ_λ intersecting Γ_1 and Γ_2 in $Q_{j,p}$, i.e.,

$$N_{j,k}(Q_{j,p}) = |\{\lambda = (j,k,m) : Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_1 \neq \emptyset \text{ and } Q_{j,p} \cap \text{supp}(\psi_\lambda) \cap \Gamma_2 \neq \emptyset\}|.$$

Then

$$N_{j,k}(Q_{j,p}) \leq \min\left(N_{j,k}^1(Q_{j,p}), N_{j,k}^2(Q_{j,p})\right). \tag{6}$$

By a counting argument, there exists some $C > 0$ such that

$$N^i_{j,k}(Q_{j,p}) \leq C \cdot 2^{j/2} \qquad \text{for } i = 1,2, \tag{7}$$

and the form of $\mathrm{supp}(\psi_\lambda)$ implies

$$N^i_{j,k}(Q_{j,p}) \leq C \cdot (|2^{j/2} s_i + k| + 1) \qquad \text{for } i = 1,2. \tag{8}$$

We now subdivide into three subcases, namely, $|s_1|, |s_2| \leq 2$, and $|s_1| \leq 2, |s_2| > 2$ (or vice versa), and $|s_1|, |s_2| > 2$, and show in each case the (almost) optimal sparse approximation rate claimed in Theorem 1. This then finishes the proof.

Subcase $|s_1|, |s_2| \leq 2$. In this case, (6) and (8) yield

$$N_{j,k}(Q_{j,p}) \leq C \cdot \min\left(|2^{j/2} s_1 + k| + 1, |2^{j/2} s_2 + k| + 1\right).$$

We first show independence on the values of s_1 and s_2 within the interval $[-2,2]$. For this, let s and s' be the slopes of the tangents to the edge curve Γ_1 (or Γ_2) at $t \in Q_{j,p} \cap \mathrm{supp}(\psi_\lambda)$ and $t' \in Q_{j,p} \cap \mathrm{supp}(\psi_{\lambda'})$, respectively, with $s \in [-2,2]$. Since Γ_1 (or Γ_2) is C^2, we have $|s - s'| \leq C \cdot 2^{-j/2}$, and hence

$$|2^{j/2} s' + k| \leq C \cdot (|2^{j/2} s + k| + 1).$$

This implies that the estimate for $N_{j,k}(Q_{j,p})$ asymptotically remains the same, independent of the values of s_1 and s_2. Further, we may assume $s' \in [-3,3]$ for $s \in [-2,2]$, since a scaling index j can be chosen such that $|s - s'|$ is sufficiently small. Therefore, one can apply inequality (5) from Proposition 3 for both points t and t'. In fact, it can be easily checked that one can use (5) with the slope s instead of s' for the point t' (or vice versa); this replacement will not change the asymptotic estimates which we will derive. Thus, we might from now on use universal values for the slopes s_1 and s_2 at each point in $Q_{j,p}$.

Now using (5) from Proposition 3, we have

$$|\langle f, \psi_{j,k,m} \rangle| \leq C \cdot \max\Big(\frac{2^{-\frac{3}{4}j}}{|2^{j/2} s_1 + k|^3}, \frac{2^{-\frac{3}{4}j}}{|2^{j/2} s_2 + k|^3} \Big). \tag{9}$$

Since $\frac{2^{-\frac{3}{4}j}}{|2^{j/2} s_i + k|^3} > \varepsilon$ implies

$$|2^{j/2} s_i + k| < \varepsilon^{-\frac{1}{3}} 2^{-\frac{1}{4}j} \qquad \text{for } i = 1,2,$$

the estimate (9) yields

$$\begin{aligned}|\Lambda_{j,p}(\varepsilon)| &\leq C \cdot \sum_{k \in K_j^1(\varepsilon) \cup K_j^2(\varepsilon)} \min(|2^{j/2}s_1+k|+1, |2^{j/2}s_2+k|+1) \\ &\leq C \cdot \sum_{i=1}^{2} \sum_{k \in K_j^i(\varepsilon)} (|2^{j/2}s_i+k|+1) \\ &\leq C \cdot (\varepsilon^{-\frac{1}{3}} 2^{-\frac{1}{4}j}+1)^2, \end{aligned} \tag{10}$$

where

$$K_j^i(\varepsilon) = \{k \in \mathbb{Z} : |2^{j/2}s_i+k| < C \cdot \varepsilon^{-\frac{1}{3}} 2^{-\frac{1}{4}j}\} \quad \text{for } i = 1,2.$$

By the hypothesis for Case 2b-2, we have $|Q_j| \leq C$, where the constant C is independent of scale $j \geq 0$. Therefore, continuing (10),

$$|\Lambda(\varepsilon)| \leq C \cdot \sum_{j=0}^{\frac{4}{3}\log_2(\varepsilon^{-1})} |\Lambda_{j,p}(\varepsilon)| \leq C \cdot \varepsilon^{-\frac{2}{3}}.$$

This allows us to write $\varepsilon > 0$ as a function of the total number of coefficients N, which gives

$$\varepsilon(N) \leq C \cdot N^{-\frac{3}{2}}.$$

Thus

$$\sum_{n>N} |\theta(f)|_n^2 \leq C \cdot N^{-2}, \tag{11}$$

which is the rate we seeked.

Subcase $|s_1| \leq 2$ *and* $|s_2| > 2$ *or vice versa.* In this case, (6)–(8) yield

$$N_{j,k}(Q_{j,p}) \leq C \cdot \min(|2^{j/2}s_1+k|+1, 2^{j/2}).$$

Again utilizing the fact that the edge curves are C^2, and using similar arguments as in the first subcase, WLOG we can conclude that the slopes s_1, s_2 at each point in $Q_{j,p}$ are greater than $\frac{3}{2}$.

Now, exploiting inequalities (4) and (5) from Proposition 3, we have

$$|\langle f, \psi_{j,k,m}\rangle| \leq C \cdot \max\Big(\frac{2^{-\frac{3}{4}j}}{|2^{j/2}s_1+k|^3}, 2^{-\frac{9}{4}j}\Big). \tag{12}$$

Since $\frac{2^{-\frac{3}{4}j}}{|2^{j/2}s_1+k|^3} > \varepsilon$ implies

$$|2^{j/2}s_1+k| < \varepsilon^{-\frac{1}{3}} 2^{-\frac{1}{4}j},$$

and $2^{-\frac{9}{4}j} > \varepsilon$ implies

$$j \leq \frac{4}{9}\log_2(\varepsilon^{-1}),$$

it follows from (12), that

$$
\begin{aligned}
|\Lambda(\varepsilon)| &\leq C \cdot \Big(\sum_{j=0}^{\frac{4}{3}\log_2(\varepsilon^{-1})} \sum_{k \in K_j^1(\varepsilon)} (|2^{j/2} s_1 + k| + 1) + \sum_{j=0}^{\frac{4}{9}\log_2(\varepsilon^{-1})} 2^{j/2} \Big) \\
&\leq C \cdot \Big(\sum_{j=0}^{\frac{4}{3}\log_2(\varepsilon^{-1})} (\varepsilon^{-\frac{1}{3}} 2^{-j/4} + 1)^2 + \sum_{j=0}^{\frac{4}{9}\log_2(\varepsilon^{-1})} 2^{j/2} \Big) \\
&\leq C \cdot \varepsilon^{-\frac{2}{3}}.
\end{aligned}
$$

The value $\varepsilon > 0$ can now be written as a function of the total number of coefficients N, which gives

$$\varepsilon(N) \leq C \cdot N^{-\frac{3}{2}}.$$

Thus, we derive again the seeked rate

$$\sum_{n>N} |\theta(f)|_n^2 \leq C \cdot N^{-2}.$$

Subcase $|s_1| > 2$ *and* $|s_2| > 2$. In this case, (6) and (7) yield

$$N_{j,k}(Q_{j,p}) \leq C \cdot 2^{j/2}.$$

Following similar arguments as before, we again derive the seeked rate (11).

5 Discussion

A variety of applications are concerned with efficient encoding of 2D functions defined on non-rectangular domains exhibiting curvilinear discontinuities, such as, e.g., a typical solution of a transport dominated partial differential equation. As an answer to this problem, our main result, Theorem 1, shows that compactly supported shearlets satisfying some weak decay and smoothness conditions, when orthogonally projected onto a given domain bounded by a piecewise C^2 curve, provide (almost) optimally sparse approximations of functions which are C^2 apart from a piecewise C^2 discontinuity curve. In this model the boundary curve is treated as a discontinuity curve.

Analyzing the proof of Theorem 1, it becomes evident that the presented optimal sparse approximation result for functions in $\mathscr{E}^2_{\nu,L}(\Omega)$ generalizes to an even more encompassing model, which does contain multiple piecewise C^2 possibly intersecting discontinuity curves separating C^2 regions in the bounded domain Ω.

In some applications it is though of importance to avoid discontinuities at the boundary of the domain. Tackling this question requires further studies to carefully design shearlets near the boundary, and this will be one of our objective for the future.

Acknowledgements Both authors would like to thank Wolfgang Dahmen, Philipp Grohs, Chunyan Huang, Demetrio Labate, Christoph Schwab, and Gerrit Welper for various discussions on related topics, and Jakob Lemvig for careful reading of an earlier version of this paper. They acknowledge support from DFG Grant SPP-1324, KU 1446/13. The first author also acknowledges support from DFG Grant KU 1446/14.

References

1. E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.
2. O. Christensen, *An introduction to frames and Riesz bases*, Birkhauser, Boston (2003).
3. S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke, *Shearlet Coorbit Spaces and associated Banach Frames*, Appl. Comput. Harmon. Anal. **27** (2009), 195–214.
4. D. L. Donoho, *Sparse components of images and optimal atomic decomposition*, Constr. Approx. **17** (2001), 353–382.
5. D. L. Donoho and G. Kutyniok, *Microlocal Analysis of the Geometric Separation Problems*, preprint (2010).
6. G. Easley, D. Labate, and W.-Q Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal. **25** (2008), 25–46.
7. K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators*, Wavelets and Splines (Athens, GA, 2005), Nashboro Press, Nashville, TN (2006), 189–201.
8. K. Guo, D. Labate, and W.-Q Lim, *Edge Analysis and identification using the Continuous Shearlet Transform*, Appl. Comput. Harmon. Anal. **27** (2009), 24–46.
9. K. Guo and D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, SIAM J. Math Anal. **39** (2007), 298–318.
10. P. Kittipoom, G. Kutyniok, and W.-Q Lim, *Construction of Compactly Supported Shearlets*, preprint (2010).
11. P. Kittipoom, G. Kutyniok, and W.-Q Lim, *Irregular Shearlet Frames: Geometry and Approximation Properties*, preprint (2010).
12. G. Kutyniok and D. Labate, *Construction of Regular and Irregular Shearlet Frames*, J. Wavelet Theory and Appl. **1** (2007), 1–10.
13. G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.
14. G. Kutyniok, J. Lemvig, and W.-Q Lim, *Compactly Supported Shearlets*, preprint (2010).
15. G. Kutyniok, and W.-Q Lim, *Compactly Supported Shearlets are Optimally Sparse*, preprint (2010).
16. G. Kutyniok, M. Shahram, and D. L. Donoho, *Development of a Digital Shearlet Transform Based on Pseudo-Polar FFT*, in Wavelets XIII (San Diego, CA, 2009), D. Van De Ville, V. K. Goyal und M. Papadakis, eds., 74460B-1 - 74460B-13, SPIE Proc. **7446**, SPIE, Bellingham, WA, 2009.
17. D. Labate, W-Q. Lim, G. Kutyniok, and G. Weiss. *Sparse multidimensional representation using shearlets*, Wavelets XI (San Diego, CA, 2005), 254-262, SPIE Proc. 5914, SPIE, Bellingham, WA, 2005.
18. W.-Q Lim, *Discrete Shearlet Transform: New Multiscale Directional Image Representation*, Proc. SAMPTA09, Marseille 2009.

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.