SHEARLAB: A RATIONAL DESIGN OF A DIGITAL PARABOLIC SCALING ALGORITHM

GITTA KUTYNIOK*, MORTEZA SHAHRAM†, AND XIAOSHENG ZHUANG‡

Abstract. Multivariate problems are typically governed by anisotropic features such as edges in images. A common bracket of most of the various directional representation systems which have been proposed to deliver sparse approximations of such features is the utilization of parabolic scaling. One prominent example is the shearlet system. Our objective in this paper is three-fold: We firstly develop a digital shearlet theory which is rationally designed in the sense that it is the digitization of the existing shearlet theory for continuous data. This implicates that shearlet theory provides a unified treatment of both the continuum and digital realm. Secondly, we analyze the utilization of pseudo-polar grids and the pseudo-polar Fourier transform for digital implementations of parabolic scaling algorithms. We derive an isometric pseudo-polar Fourier transform by careful weighting of the pseudo-polar grid, allowing exploitation of its adjoint for the inverse transform. This leads to a digital implementation of the shearlet transform; an accompanying Matlab toolbox called `ShearLab` is provided. And, thirdly, we introduce various quantitative measures for digital parabolic scaling algorithms in general, allowing one to tune parameters and objectively improve the implementation as well as compare different directional transform implementations. The usefulness of such measures is exemplarily demonstrated for the digital shearlet transform.

Key words. Curvelets, digital shearlet system, directional representation system, fast digital shearlet transform, parabolic scaling, performance measures, software package, tight frames

AMS subject classifications. Primary 42C40; Secondary 42C15, 65K99, 65T60, 65T99, 94A08

1. Introduction. In recent years, applied harmonic analysts have introduced several approaches for directional representations of image data, each one with the intent of efficiently representing highly anisotropic image features. Examples include curvelets [6, 7, 3], contourlets [9], and shearlets [14, 26]. These proposals are inspired by elegant results in theoretical harmonic analysis, which study functions defined on the continuum plane (i.e., not digital images) and address problems of efficiently representing certain types of functions and operators. One set of inspiring results concerns the possibility of highly compressed representations of 'cartoon' images, i.e., functions which are piecewise smooth with singularities along smooth curves. Another set of results concerns the possibility of highly compressed representations of wave propagation operators. In 'continuum theory', anisotropic directional transforms can significantly outperform wavelets in important ways.

Accordingly, one hopes that a digital implementation of such ideas would also deliver performance benefits over wavelet algorithms in real-world settings. Anticipated applications include [20], where missing sensors cause incomplete measurements, and

the problem of texture/geometry separation in image processing – for example in astronomy when images of galaxies require separated analyses of stars, filaments, and sheets [29, 10].

In many cases, however, there are no publicly available implementations of such ideas, or the available implementations are only sketchily tested, or the available implementations are only vaguely related to the continuum transforms they are reputed to represent. Accordingly, we have not yet seen a serious exploration of the potential benefit of such transforms, carefully comparing the expected benefits with those delivered by specific implementations.

In this paper we aim at providing both:

(1) A rationally designed shearlet transform implementation.
(2) A comprehensive framework for quantifying performance of directional representations in general.

For (1), we developed an implementation of the fast digital shearlet transform (FDST) based on a digital shearlet theory which is a very natural digitization of the existing shearlet theory for continuous data. Other parabolic-scaling transforms, for example, curvelets are inherently based on operations (rotation) which translate awkwardly into the digital realm. In contrast, when we consider shearlets, rotations are replaced by shearing, which has a natural digital realization, thus enabling a unified treatment for the continuum and digital realm similar to wavelets.

The framework in (2) has three benefits. First, it provides quantitative performance measures which can be used to tune the parameters of our implementation, which is publicly available at `www.ShearLab.org`. This allows us to specify 'recommended choices' for the parameters of our implementation. Second, the same 'measure and tune' approach may be useful to other implementers of directional transforms. Third, we show a way to improve the level of intellectual seriousness in applied mathematics which pretends to work in image processing. We believe that widespread adoption of this measure and tune framework can be very valuable, since many supposedly scientific presentations are now little more than vague, numbing 'advertising' or 'marketing' pitches. They could instead offer quantitative comparisons between algorithms, and thereby be far more informative. In fact the combination of quantitative evaluation with reproducible research [11] would be particularly effective at producing both intellectual seriousness and rapid progress.

1.1. Desiderata. We start by proposing the following desiderata for the fast digital shearlet transform FDST and its implementation:

[D1] *Algebraic Exactness.* The transform should be based on a shearlet theory for digital data on a pseudo-polar grid, than merely being 'somewhat close' to the shearlet theory for continuous data.

[D2] *Isometry of Pseudo-Polar Fourier Transform.* We introduce oversampling and weights to obtain an isometric pseudo-polar Fourier transform, which allows us to use the adjoint as inverse transform.

[D3] *Tight Frame Property.* The shearlet coefficients computed by the transform should be based on a tight frame decomposition, which ensures an isometric relation between the input image and the sequence of coefficients as well as allows us to use the adjoint as inverse transform. This property follows by combining [D1] and [D2], and allows the comparison with other transforms in contrast to those previous two tests.

[D4] *Time-Frequency-Localization.* The spatial portrait of the analyzing elements should 'look like' shearlets in the sense that they are sufficiently smooth as

well as time-localized. Localization and smoothness in frequency domain is ensured by definition.

[D5] *True Shear Invariance.* Since the orientation-related operator of shearlets is in fact the shear operator, we expect to see a shearing of the input image mirrored in a simple shift of the transform coefficients.

[D6] *Speed.* The transform should admit an algorithm of order $O(N^2 \log N)$ flops, where N^2 is the number of digital points of the input image.

[D7] *Geometric Exactness.* The transform should preserve geometric properties parallel to those of the continuum theory, for example, edges should be mapped to edges in shearlet domain.

[D8] *Robustness.* The transform should be resilient against impacts such as (hard) thresholded and quantized coefficients.

1.2. Definition of the Shearlet Transform. The main idea for the construction of the shearlet transform with discrete parameters for functions in $L^2(\mathbb{R}^2)$ is the choice of a two-parameter dilation group, where one parameter ensures the multiscale property, whereas the second parameter provides a means to detect directions. The choice for a direction sensitive parameter is particularly important, since the most canonical choice, the rotation, would prohibit a unified treatment of the continuum and digital realm due to the fact that the integer grid is not invariant under rotation. Shearlets parameterize directions by slope rather than angles. And the shear matrix does preserve the structure of the integer grid, which is key to enabling an exact digitization of the continuum domain shearlets.

For each $a > 0$ and $s \in \mathbb{R}$, let A_a denote the *parabolic scaling matrix* and S_s denote the *shear matrix* of the form

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \qquad \text{and} \qquad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix},$$

respectively. To provide an equal treatment of the x- and y-axis, the frequency plane is split into the four cones $\mathcal{C}_{11} - \mathcal{C}_{22}$ (see Figure 1.1), defined by

$$\mathcal{C}_\iota = \begin{cases} \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \geq 1, \ |\xi_1/\xi_2| \geq 1\} & : \ \iota = 21, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \geq 1, \ |\xi_1/\xi_2| \leq 1\} & : \ \iota = 11, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \leq -1, |\xi_1/\xi_2| \geq 1\} & : \ \iota = 22, \\ \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 \leq -1, |\xi_1/\xi_2| \leq 1\} & : \ \iota = 12. \end{cases}$$

Let now $\psi_1 \in L^2(\mathbb{R})$ be a wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subseteq [-4, -\frac{1}{4}] \cup [\frac{1}{4}, 4]$, and let $\psi_2 \in L^2(\mathbb{R})$ be a 'bump' function satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_2 \subseteq [-1, 1]$. We define $\psi \in L^2(\mathbb{R}^2)$ by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\tfrac{\xi_2}{\xi_1}). \tag{1.1}$$

For cone $\mathcal{C}_{21}$, at scale $j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, orientation $s = -2^j, \ldots, 2^j$, and spatial position $m \in \mathbb{Z}^2$, the associated *shearlets* are then defined by their Fourier transforms

$$\begin{aligned} \hat{\sigma}_\eta(\xi) &= 2^{-j\frac{3}{2}}\hat{\psi}(S_s^T A_{4^{-j}}\xi)\chi_{\mathcal{C}_{21}}(\xi)e^{2\pi i \langle A_{4^{-j}} S_s m, \xi \rangle} \\ &= 2^{-j\frac{3}{2}}\hat{\psi}_1(4^{-j}\xi_1)\hat{\psi}_2(s + 2^j \tfrac{\xi_2}{\xi_1})\chi_{\mathcal{C}_{21}}(\xi)e^{2\pi i \langle A_{4^{-j}} S_s m, \xi \rangle}, \end{aligned} \tag{1.2}$$

where $\eta = (j, s, m, \iota)$ index scale, orientation, position, and cone. The shearlets for $\mathcal{C}_{11}$, $\mathcal{C}_{21}$, and $\mathcal{C}_{22}$ are defined likewise by symmetry, as illustrated in Figure 1.1, and

we denote the resulting *discrete shearlet system* by

$$\{\sigma_\eta : \eta \in \mathbb{N}_0 \times \{-2^j, \ldots, 2^j\} \times \mathbb{Z}^2 \times \{11, 12, 21, 22\}\}. \qquad (1.3)$$

The definition shows that shearlets live on anisotropic regions of width 2^{-2j} and length 2^{-j} at various orientations.

It should be mentioned that discrete shearlets – 'discrete' referring to the set of parameters and *not* to the domain – can also be defined with respect to the dilation matrix $A_{2^{-j}}$. However, in this case the odd scales have to be handled particularly carefully. The attentive reader will have also observed that recently introduced compactly supported shearlets [22, 26] do not require projecting the shearlets to the respective cones; however, despite other advantageous properties, they do not form a tight frame for $L^2(\mathbb{R}^2)$. Finally, the generating window allows in fact more freedom than (1.1), but in this paper we restrict ourselves to this (customary) choice.
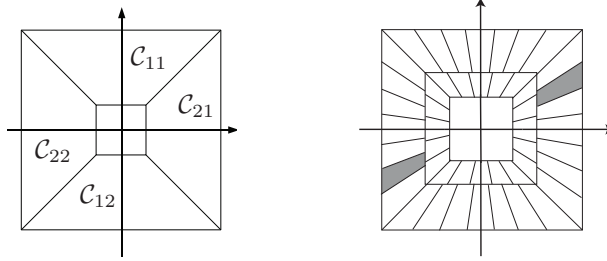

FIG. 1.1. *The cones $\mathcal{C}_{11} - \mathcal{C}_{22}$ and the tiling of the frequency domain induced by shearlets.*

Setting $\mathcal{C}^\vee = \bigcup_{\iota=11}^{22} \mathcal{C}_\iota$, we have the following theorem from [14, Thm. 3] concerning the frame properties of the discrete shearlet system. For the definition of a tight (sometimes called Parseval) frame, we refer to [8].

THEOREM 1.1 ([14]). *The system (1.3) is a tight frame for $\{f \in L^2(\mathbb{R}^2) : \text{supp } \hat{f} \subseteq \mathcal{C}^\vee\}$.* We remark that the low frequency part can be appropriately filled in to obtain a tight frame for $L^2(\mathbb{R}^2)$.

The transform associated with this system is the *discrete shearlet transform*, which for a given function $f \in L^2(\mathbb{R}^2)$ is defined to be the map

$$\mathbb{N}_0 \times \{-2^j, \ldots, 2^j\} \times \mathbb{Z}^2 \times \{11, 12, 21, 22\} \ni \eta \mapsto (\langle f, \sigma_\eta \rangle) \in \mathbb{C}.$$

It is this transform, which we aim to exactly digitize.

1.3. Ingredients of the Fast Digital Shearlet Transform (FDST). The shearlet transform for continuum domain data (see Figure 1.1) implicitly induces a trapezoidal tiling of frequency space which is evidently not cartesian. By introducing a special set of coordinates on the continuum 2D frequency space, the discrete shearlet transform can be represented as a cascade of five operations:
- Classical Fourier transformation.
- Change of variables to pseudo-polar coordinates.
- Weighting by a radial 'density compensation' factor.
- Decomposition into rectangular tiles.
- Inverse Fourier transform of each tiles.

Surprisingly, this process admits a natural translation into the digital domain. The key observation is that the pseudo-polar coordinates are naturally compatible

with digital image processing (compare Figure 1.2) and perfectly suited for a digitization of the discrete shearlet transform as a comparison with the frequency tiling generated by continuum domain shearlets in Figure 1.1 already visually evidences. Fortunately, in [1] a fast pseudo-polar Fourier transform (PPFT) is already developed. This transform evaluates the Fourier transform of an image of size N, say, on a pseudo-polar grid Ω of the form $\Omega = \Omega^1 \cup \Omega^2$, where

$$\Omega^1 = \{(-k \cdot \tfrac{2\ell}{N}, k) : -\tfrac{N}{2} \leq \ell \leq \tfrac{N}{2}, -N \leq k \leq N\},$$
$$\Omega^2 = \{(k, -k \cdot \tfrac{2\ell}{N}) : -\tfrac{N}{2} \leq \ell \leq \tfrac{N}{2}, -N \leq k \leq N\}.$$

Figure 1.2 shows an illustration of the case $N = 4$. For an $N \times N$ image $I := \{I(u,v) :$

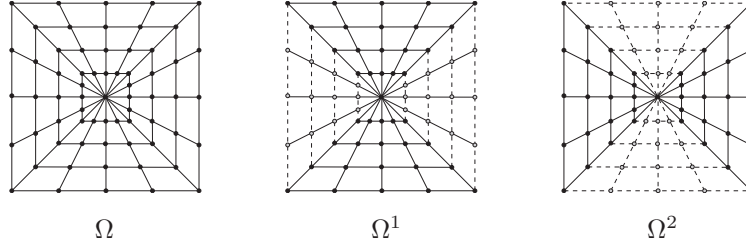Ω Ω^1 Ω^2

FIG. 1.2. *The pseudo-polar grid* $\Omega = \Omega^1 \cup \Omega^2$ *for* $N = 4$.

$-N/2 \leq u, v \leq N/2 - 1\}$, the pseudo-polar Fourier transform $\hat{I}$ of I evaluated on the pseudo-polar grid Ω is then defined to be

$$\hat{I}(\omega_x, \omega_y) = \sum_{u,v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i}{m_0}(u\omega_x + v\omega_y)}, \quad (\omega_x, \omega_y) \in \Omega,$$

where $m_0 \geq N$ is an integer which for the PPFT is chosen to be $m_0 = 2N + 1$ for computational reasons.

The existence of PPFT suggests that we can easily and naturally get a faithful FDST using this algorithm. However, besides the delicateness of digitizing the continuum domain shearlets so that they form a tight frame on the pseudo-polar grid, also the use of the PPFT is not at all straightforward. The PPFT as presented [1] is not an isometry. The main obstacle is the highly nonuniform arrangement of the points on the pseudo-polar grid. This intuitively suggests to downweight points in regions of very high density by using weights which correspond roughly to the density compensation weights underlying the continuous change of variables. In fact, we will show that isometry is possible with sufficient radial oversampling of the pseudo-polar grid; however, the weights will not be derivable from simple density compensation arguments.

Summarizing, the FDST of an $N \times N$ image cascades the following steps:
  1) Application of the PPFT with an oversampling factor R in radial direction.
  2) Weighting of the function values on the pseudo-polar grid by 'density-compensation-style' weights.
  3) Decomposing the pseudo-polar-indexed values by a scaled and sheared generating window into rectangular subbands followed by application of the 2D iFFT to each array.

This is an exact analogy of the discrete shearlet transform, in which the steps of Fourier transformation and pseudo-polar coordinate change as well as the steps of

decomposition into rectangular tiles and the inverse Fourier transform are collapsed into one step, respectively. With a careful choice of the weights and the windows, this transform is an isometry as we will show. Hence the inverse transform can be computed by merely taking the adjoint in each step.

1.4. Performance Measurement. The above sketch does not uniquely specify an implementation; there is freedom in choice of weights and windows. How can we decide if one choice is better than another one? It seems that currently researchers often use overall system performance on isolated tasks, such as denoising and compression of specific standard images like 'Lena', 'Barbara', etc. However, overall system performance for a system made up of a cascade of steps seems very opaque and at the same time very particular. It seems far better from an intellectual viewpoint to carefully decompose performance according to a more insightful array of tests, each one motivated by a particular well-understood property we are trying to obtain.

We have developed quantitative performance measures inspired by the desiderata we presented in Subsection 1.1. Each performance measure produces a real value or a real-valued curve, thus providing a standardized framework for evaluation and, especially, comparison.

1.5. Relation with Previous Work. Since the introduction of directional representation systems by many pioneer researchers ([4, 5, 6, 7, 9, 14]), various numerical implementations of their directional representation systems have been proposed. The closest ones are the curvelet, contourlet, and previous shearlet algorithms, whose main features we now briefly survey.

Curvelets [3]. The discrete curvelet transform is implemented in the software package *CurveLab*, which comprises two different approaches. One is based on unequispaced FFTs, which are used to interpolate the function in the frequency domain on different tiles with respect to different orientations of curvelets. The other is based on frequency wrapping, which wraps each subband indexed by scale and angle into a fixed rectangle around the origin. Both approaches can be realized efficiently in $O(N^2 \log N)$ flops with N being the image size. The disadvantage of this approach is the lack of an associated continuum domain theory.

Contourlets [9]. The implementation of contourlets is based on a directional filter bank, which produces a directional frequency partitioning similar to the one generated by curvelets. The main advantage of this approach is that it allows a tree-structured filter bank implementation, in which aliasing due to subsampling is allowed to exist. Consequently, one can achieve great efficiency in terms of redundancy and good spatial localization. A drawback of this approach is that various artifacts are introduced and that an associated continuum domain theory is missing.

Shearlets [12, 28]. In [12], Easley et. al. implemented the shearlet transform by applying the Laplacian pyramid scheme and directional filtering successively. One drawback is the deviation from the continuum domain theory. Another drawback is that the associated code was not made publicly available. In contrast to this implementation which is based on bandlimited subband tiling – similar to the implementation of curvelets in [3] – in [28], Lim provided an implementation of the shearlet transform based on compactly supported shearlet systems (see also [22]). These compactly supported shearlets are separable and provide excellent spatial localization. The drawback is that they do not form a tight frame, hence, the synthesis process needs to be performed by iterative methods. We further wish to mention two novel approaches [27] and [19] for which however no implementation is yet available nor was their focus on deriving an exact digitization of the continuum domain transform.

Summarizing, all the above implementations of directional representation systems have their own advantages and disadvantages, one of the most common shortcomings being the lack of providing the unified treatment of the continuum and digital world. Our effort will now be put to provide a natural digitization of the shearlet theory (bandlimited shearlets) fulfilling the unified treatment requirement as well as a software package `ShearLab` quantifying performances of directional representation systems.

1.6. Contribution of this Paper. The contributions of this paper are twofold. Firstly, we introduce a fast digital shearlet transform (FDST) which is *rationally designed* based on a *natural* digitization of shearlet theory. Secondly, we provide a variety of *quantitative performance measures* for directional representations, which allow tuning and comparison of implementations. Our digital shearlet implementation was tuned utilizing this framework, so we can provide the user community with an optimized representation.

All presented algorithms and tests are provided at `www.ShearLab.org` in the spirit of reproducible research [11].

1.7. Contents. Section 2 introduces the fast digital shearlet transform FDST and proves isometry. In Section 3, we then discuss two variants of an inverse digital shearlet transform, namely, a direct and an iterative approach. In Section 4, we prove several mathematical properties of the FDST such as decay properties of digital shearlet coefficients. The following section, Section 5, is concerned with details of the associated `ShearLab` implementation at `www.ShearLab.org`. The FDST is then analyzed in Section 7 according to the quantitative measures introduced in Section 6.

2. FDST for Finite Data. We start by discussing the three steps in the FDST as described in Subsection 1.3, which we for the convenience of the reader briefly repeat:

1) Application of the PPFT with an oversampling factor R in radial direction.
2) Weighting of the function values on the pseudo-polar grid by 'density-compensation-style' weights.
3) Decomposing the pseudo-polar-indexed values by a scaled and sheared generating window into rectangular subbands followed by application of the 2D iFFT to each array.

We will also show that careful selection of the oversampling factor, of the weights, and of the windows yields an isometric transform, which enables us to compute the inverse shearlet transform by its adjoint (see Section 3).

2.1. Weighted Pseudo-Polar Fourier Transform. Given an $N \times N$ image I, it is well known that the Fourier transform $\hat{I}$ of I evaluated on a rectangular $N \times N$ grid is an isometry:

$$\sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2 = \frac{1}{N^2} \sum_{\omega_x,\omega_y=-N/2}^{N/2-1} |\hat{I}(\omega_x,\omega_y)|^2. \qquad (2.1)$$

This is the Plancherel formula for a function defined on a finite group [21].

We now intend to obtain a similar formula for the Fourier transform of I evaluated on the pseudo-polar grid. For this, we first extend the definition of the pseudo-polar grid slightly by introducing an oversampling parameter $R > 0$ in radial direction. This new grid, which we will denote in the sequel by Ω_R, is defined by

$$\Omega_R = \Omega_R^1 \cup \Omega_R^2,$$

where

$$\Omega_R^1 = \{(-\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}, \tfrac{2k}{R}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, -\tfrac{RN}{2} \le k \le \tfrac{RN}{2}\}, \qquad (2.2)$$

$$\Omega_R^2 = \{(\tfrac{2k}{R}, -\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, -\tfrac{RN}{2} \le k \le \tfrac{RN}{2}\}. \qquad (2.3)$$

This grid is illustrated in Figure 2.1. Notice that the 'original' pseudo-polar grid (see

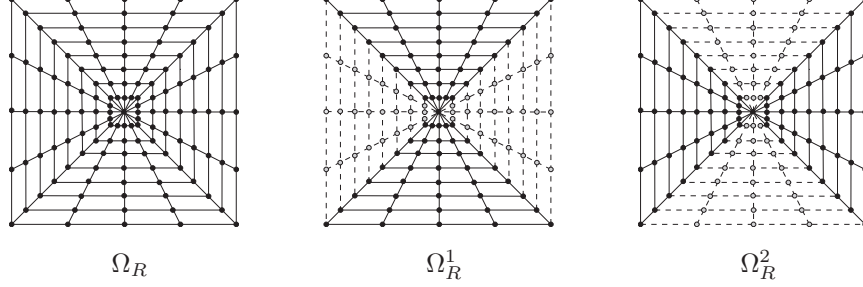$$\Omega_R \qquad\qquad \Omega_R^1 \qquad\qquad \Omega_R^2$$

FIG. 2.1. *The pseudo-polar grid* $\Omega_R = \Omega_R^1 \cup \Omega_R^2$ *for* $N = 4$ *and* $R = 4$.

Figure 1.2) as introduced in Subsection 1.3 is a special case of this definition when choosing $R = 2$. Also observe that the center

$$\mathcal{C} = \{(0,0)\}$$

appears $N + 1$ times in Ω_R^1 as well as Ω_R^2, and the points on the seam lines

$$\mathcal{S}_R^1 = \{(-\tfrac{2k}{R}, \tfrac{2k}{R}) : -\tfrac{RN}{2} \le k \le \tfrac{RN}{2}, k \ne 0\},$$

$$\mathcal{S}_R^2 = \{(\tfrac{2k}{R}, -\tfrac{2k}{R}) : -\tfrac{RN}{2} \le k \le \tfrac{RN}{2}, k \ne 0\},$$

appear in both Ω_R^1 and Ω_R^2. Later, we will also utilize a further partitioning of the sets Ω_R^1 and Ω_R^2 as

$$\Omega_R^1 = \Omega_R^{11} \cup \mathcal{C} \cup \Omega_R^{12} \qquad \text{and} \qquad \Omega_R^2 = \Omega_R^{21} \cup \mathcal{C} \cup \Omega_R^{22},$$

where

$$\Omega_R^{11} = \{(-\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}, \tfrac{2k}{R}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, 1 \le k \le \tfrac{RN}{2}\},$$

$$\Omega_R^{12} = \{(-\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}, \tfrac{2k}{R}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, -\tfrac{RN}{2} \le k \le -1\},$$

$$\Omega_R^{21} = \{(\tfrac{2k}{R}, -\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, 1 \le k \le \tfrac{RN}{2}\}$$

$$\Omega_R^{22} = \{(\tfrac{2k}{R}, -\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}) : -\tfrac{N}{2} \le \ell \le \tfrac{N}{2}, -\tfrac{RN}{2} \le k \le -1\}.$$

Now our goal is to choose weights $w : \Omega_R \to \mathbb{R}^+$ so that, for any $N \times N$ image I,

$$\sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2 = \sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot |\hat{I}(\omega_x, \omega_y)|^2, \qquad (2.4)$$

where here we modify the definition of the Fourier transform according to [1] and define it by

$$\hat{I}(\omega_x, \omega_y) = \sum_{u,v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i}{m_0}(u\omega_x + v\omega_y)}, \quad (\omega_x, \omega_y) \in \Omega_R, \qquad (2.5)$$

where $m_0 \geq N$. Also notice that the factor $1/N^2$ appearing in (2.1) will now be hidden in the weights $w(\omega_x, \omega_y)$.

We start by computing the right hand side of (2.4):

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot |\hat{I}(\omega_x, \omega_y)|^2$$

$$= \sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot \left| \sum_{u,v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i}{m_0}(u\omega_x + v\omega_y)} \right|^2$$

$$= \sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot \left[\sum_{u,v=-N/2}^{N/2-1} \sum_{u',v'=-N/2}^{N/2-1} I(u,v)\overline{I(u',v')} e^{-\frac{2\pi i}{m_0}((u-u')\omega_x + (v-v')\omega_y)} \right]$$

$$= \sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot \sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2$$

$$+ \sum_{\substack{u,v,u',v'=-N/2 \\ (u,v) \neq (u',v')}}^{N/2-1} I(u,v)\overline{I(u',v')} \cdot \left[\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot e^{-\frac{2\pi i}{m_0}((u-u')\omega_x + (v-v')\omega_y)} \right].$$

Choosing $I = c_{u_1,v_1}\delta(u-u_1, v-v_1) + c_{u_2,v_2}\delta(u-u_2, v-v_2)$ for all $-N/2 \leq u_1, v_1, u_2, v_2 \leq N/2-1$ and for all $c_{u_1,v_1}, c_{u_2,v_2} \in \mathbb{C}$, we can conclude that (2.4) holds if and only if

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot e^{-\frac{2\pi i}{m_0}(u\omega_x + v\omega_y)} = \delta(u,v), \quad -N+1 \leq u,v \leq N-1. \quad (2.6)$$

This is equivalent to the two conditions

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot \cos(\tfrac{2\pi}{m_0}(u\omega_x + v\omega_y)) = \delta(u,v), \quad (2.7)$$

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot \sin(\tfrac{2\pi}{m_0}(u\omega_x + v\omega_y)) = 0, \quad (2.8)$$

for all $-N+1 \leq u,v \leq N-1$. In view of the symmetry of the pseudo-polar grid, it is natural to impose the following symmetry conditions on the weights:

[S1] $w(\omega_x, \omega_y) = w(\omega_y, \omega_x), \quad (\omega_x, \omega_y) \in \Omega_R$,
[S2] $w(\omega_x, \omega_y) = w(-\omega_y, \omega_x), (\omega_x, \omega_y) \in \Omega_R$,
[S3] $w(\omega_x, \omega_y) = w(-\omega_x, \omega_y), (\omega_x, \omega_y) \in \Omega_R$,
[S4] $w(\omega_x, \omega_y) = w(\omega_x, -\omega_y), (\omega_x, \omega_y) \in \Omega_R$.

In this case, (2.8) automatically holds. By the sum formula for trigonometric functions, (2.7) is then equivalent to

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot [\cos(\tfrac{2\pi}{m_0}u\omega_x)\cos(\tfrac{2\pi}{m_0}v\omega_y) - \sin(\tfrac{2\pi}{m_0}u\omega_x)\sin(\tfrac{2\pi}{m_0}v\omega_y)] = \delta(u,v)$$

for all $-N+1 \leq u,v \leq N-1$. Again, by the symmetry of the weights, this is equivalent to

$$\sum_{(\omega_x, \omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot [\cos(\tfrac{2\pi}{m_0}u\omega_x)\cos(\tfrac{2\pi}{m_0}v\omega_y)] = \delta(u,v) \quad (2.9)$$

for all $-N+1 \le u, v \le N-1$. This is a linear system of equations with $RN^2/4 + RN/2 + 1$ unknows and $(2N-1)^2$ equations, wherefore, in general, we need the oversampling factor R to be at least 16 to enforce solvability.

For symmetry reasons, we can now restrict our attention to one quarter of a cone, say Ω_R^{21}. Using (2.2), (2.3), and (2.9), we then obtain the following equivalent condition to (2.6):

$$\delta(u,v) = w(0,0) + 4 \cdot \sum_{\ell=0,N/2}^{RN/2} \sum_{k=1}^{RN/2} w(\tfrac{2k}{R}, -\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}) \cdot \cos(2\pi u \cdot \tfrac{2k}{m_0 R}) \cdot \cos(2\pi v \cdot \tfrac{2k}{m_0 R} \cdot \tfrac{2\ell}{N})$$

$$+ 8 \cdot \sum_{\ell=1}^{N/2-1} \sum_{k=1}^{RN/2} w(\tfrac{2k}{R}, -\tfrac{2k}{R} \cdot \tfrac{2\ell}{N}) \cdot \cos(2\pi u \cdot \tfrac{2k}{m_0 R}) \cdot \cos(2\pi v \cdot \tfrac{2k}{m_0 R} \cdot \tfrac{2\ell}{N}) \quad (2.10)$$

for all $-N+1 \le u, v \le N-1$. Concluding, we have the following result.

THEOREM 2.1. *Let N be even, let $\Omega_R = \Omega_R^1 \cup \Omega_R^2$ be the pseudo-polar grid defined in (2.2) and (2.3), and let $w : \Omega_R \to \mathbb{R}^+$ be a weight function satisfying the symmetry conditions [S1] – [S4]. Then*

$$\sum_{u,v=-N/2}^{N/2-1} |I(u,v)|^2 = \sum_{(\omega_x,\omega_y) \in \Omega_R} w(\omega_x, \omega_y) \cdot |\hat{I}(\omega_x, \omega_y)|^2$$

holds if and only if the weights $w(\omega_x, \omega_y), (\omega_x, \omega_y) \in \Omega_R$ satisfy condition (2.10). Moreover, in general, R needs to be at least 16 for such weights to exist.

2.2. Weight Functions. To avoid high complexity in the computation of the weights satisfying Theorem 2.1, we relax the requirement for exact isometric weighting. Instead of representing the weights as the solution of a large system of equations, they will be represented in terms of an undercomplete basis for functions on the pseudo-polar grid. More precisely, we first design basis functions $w_1, \dots, w_n : \Omega_R \to \mathbb{R}^+$ such that $\sum_{j=1}^{n} w_j(\omega_x, \omega_y) \ne 0$ for all $(\omega_x, \omega_y) \in \Omega_R$. We then define the weight function $w : \Omega_R \to \mathbb{R}^+$ to be $w := \sum_{j=1}^{n} c_j w_j$, with $c_1, \dots, c_n$ being nonnegative constants. These coefficients are determined by solving (2.10) with respect to this weight function w using the least square method. We compute the coefficients in this expansion once for a given problem size; then hardwire them in the algorithm.

2.2.1. Recommended Choices of Weights. In what follows, we present several designs of basis functions, each one providing nearly isometric weighting. Notice that, slightly abusing notation, we will use (ω_x, ω_y) and (k, ℓ) interchangeably. The weighting for the three choice we recommend is displayed in Figure 2.2.

CHOICE 1: Our first choice are the following seven functions $w_1, \dots, w_7$:
Center: $w_1 = 1_{(0,0)}$ and $w_2 = 1_{\{(\omega_x,\omega_y):|k|=1\}}$,
Boundary: $w_3 = 1_{\{(\omega_x,\omega_y):|k|=NR/2, \omega_x=\omega_y\}}$ and $w_4 = 1_{\{(\omega_x,\omega_y):|k|=NR/2, \omega_x \ne \omega_y\}}$,
Seam lines: $w_5 = |k| \cdot 1_{\{(\omega_x,\omega_y):1<|k|<NR/2, \omega_x=\omega_y\}}, w_6 = 1_{\{(\omega_x,\omega_y):|k|=NR/2-3, \omega_x=\omega_y\}}$,
Interior: $w_7 = |k| \cdot 1_{\{(\omega_x,\omega_y):1<|k|<NR/2, \omega_x \ne \omega_y\}}$.

CHOICE 2: This is a simplified version of 'Choice 1' using the 5 functions:
Center: $w_1 = 1_{(0,0)}$,
Boundary: $w_2 = 1_{\{(\omega_x,\omega_y):|k|=NR/2, \omega_x=\omega_y\}}$ and $w_3 = 1_{\{(\omega_x,\omega_y):|k|=NR/2, \omega_x \ne \omega_y\}}$,
Seam lines: $w_4 = |k| \cdot 1_{\{(\omega_x,\omega_y):1 \le |k|<NR/2, \omega_x=\omega_y\}}$,

Interior: $w_5 = |k| \cdot 1_{\{(\omega_x, \omega_y): 1 \leq |k| < NR/2, \, \omega_x \neq \omega_y\}}$.

<u>CHOICE 3</u>: Finally, we suggest the following $N/2 + 2$ functions on the pseudo-polar grid:

Center: $w_1 = 1_{(0,0)}$,

Radial Lines: $w_{\ell+2} = 1_{\{(\omega_x, \omega_y): 1 < |k| < NR/2, \, \omega_y = \frac{\ell}{N/2} \omega_x\}}$, $\ell = 0, 1, \ldots, N/2$.
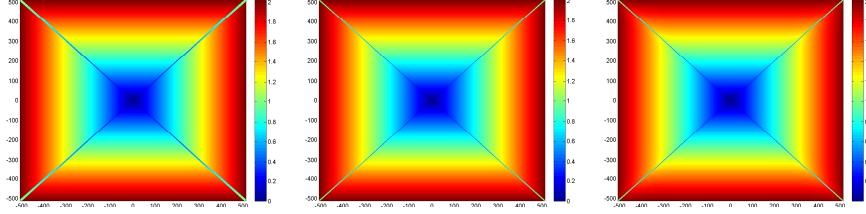

FIG. 2.2. *Recommended choices of weighting of the pseudo-polar grid for $N = 128$ and $R = 8$. The Choice 1 – 3 are displayed from left to right.*

2.2.2. Comparison of Weight Functions. The patterns of the weights are seemingly similar in view of Figure 2.2. However, their performances can be quite different depending on the chosen performance measure. We mention that the measures chosen below are also part of our framework of performance measures for parabolic scaling algorithms discussed in Section 6.

Letting $R = 8$, we generate a sequence of 5 random images $I_1, \ldots, I_5$ of size $N \times N$ with standard normally distributed entries. We use the following measure to compare the performance of different weights:

$$M_{isom} := \frac{1}{5} \sum_{i=1}^{5} \frac{\|P^\star w P I_i - I_i\|_2}{\|I_i\|_2},$$

where $P : I \rightarrow \hat{I}$ denotes the PPFT from (2.5) and $w : J \rightarrow J_w$ – by abusing notation – denotes the 'weighting operator' $J_w = w \cdot J$, for a image $J : \Omega_R \rightarrow \mathbb{C}$ and a weight function $w : \Omega_R \rightarrow \mathbb{R}^+$. Table 2.2.2 displays the performance of the weights from Choice 1 – 3 with respect to this measure.

TABLE 2.1
Comparison using random images

N	32	64	128	256	512
Choice 1	4.3E-3	2.6E-3	2.2E-3	1.4E-3	9.3E-4
Choice 2	4.2E-3	4.0E-3	1.8E-3	1.5E-3	8.8E-4
Choice 3	9.8E-3	6.2E-3	3.4E-3	2.1E-3	N/A

Next, we choose the real image 'Barbara', which we denote by I, and the measure $\frac{\|P^\star w P I - I\|_2}{\|I\|_2}$ to compare the performance of the different choices of weights.

From Tables 2.2.2 and 2.2.2, it can be seen that the operator $P^\star w P$ seems to converge to an identity operator as $N \rightarrow \infty$. The data also indicates that it is justifiable to define weights which are linearly increasing along the radial direction.

TABLE 2.2
Comparison using 'Barbara'

N	32	64	128	256	512
Choice 1	2.4E-3	1.6E-3	1.1E-3	5.2E-4	2.2E-4
Choice 2	2.8E-3	1.2E-3	8.3E-4	3.9E-4	1.5E-4
Choice 3	5.6E-3	2.8E-3	2.2E-3	9.1E-4	N/A

When using iterative methods, for instance, the conjugate gradient method, for computing the inverse, a weight w plays the role of a preconditioner. Therefore its performance as such can be effectively measured by the condition number of the operator $P^\star wP$, i.e., $cond(P^\star wP) = \lambda_{max}(P^\star wP)/\lambda_{min}(P^\star wP)$. This measure is displayed in Table 2.2.2 for our selected three choices of weights. Notice that, for each choice of a weight function with the exception of Choice 3, the condition numbers of $P^\star wP$ are always smaller than 2.

TABLE 2.3
Comparison of $cond(P^\star wP)$

N	32	64	128	256	512
Choice 1	1.328	1.483	1.621	1.726	1.834
Choice 2	1.379	1.503	1.621	1.731	1.833
Choice 3	1.760	1.887	2.001	2.104	N/A

2.3. Windowing. According to our discussion of the main steps of the FDST in Section 2, after performing a weighted pseudo-polar Fourier transform, the data has to be windowed using scaled and sheared versions of a generating window function.

In this section, we now define such a set of window functions, which we will coin *digital shearlets*, and prove that these – similar to the continuum domain – form a tight frame for functions $J : \Omega_R \to \mathbb{C}$. We remark that this construction is a digitization of the continuum domain discrete shearlets introduced in [14] (compare also (1.2) and (1.3)). However, it is far from obvious that again a tight frame is derived, since we here consider a finite domain.

For the convenience of the reader, we first briefly recall the notion of a tight frame in this particular situation. Let $f, g : \Omega_R \to \mathbb{C}$ be two functions defined on Ω_R. Then, the inner product $\langle f, g \rangle_{\Omega_R}$ is defined to be $\langle f, g \rangle_{\Omega_R} := \sum_{x \in \Omega_R}(x)\overline{g(x)}$. A sequence $\{\varphi_\lambda : \Omega_R \to \mathbb{C} : \lambda \in \Lambda\}$ with Λ being an indexing set is a *tight frame* for functions $J : \Omega_R \to \mathbb{C}$, if

$$\sum_{\lambda \in \Lambda} |\langle J, \varphi_\lambda \rangle_{\Omega_R}|^2 = \langle J, J \rangle_{\Omega_R}.$$

It then follows from basic frame theory (see [8]) that this allows recovery of a function $J : \Omega_R \to \mathbb{C}$ from its coefficients $(\langle J, \varphi_\lambda \rangle_{\Omega_R})_{\lambda \in \Lambda}$ by computing

$$J = \sum_{\lambda \in \Lambda} \langle J, \varphi_\lambda \rangle_{\Omega_R} \varphi_\lambda.$$

Despite the danger of repeating ourselves, let us mention that our fundamental goal is to introduce digital shearlets as the exact digitization of continuum domain

shearlets. We now describe the construction step by step, which will give evidence to the fact that we achieved this goal. There will be one step though – when defining the modulation – where we have to slightly deviate from an exact digitization, and we will explain the reasons for this.

We start by defining the scaling function and the generating digital shearlet. For this, let $j_L := -\lceil \log_4(R/2) \rceil$, which will soon be shown to be the lowest possible scale. Let W_0 be the Fourier transform of the Meyer scaling function such that

$$\text{supp } W_0 \subseteq [-1, 1] \quad \text{and} \quad W_0(\pm 1) = 0, \tag{2.11}$$

and let V_0 be a 'bump' function satisfying

$$\text{supp } V_0 \subseteq [-\tfrac{3}{2}, \tfrac{3}{2}] \qquad \text{with} \qquad V_0(\xi) \equiv 1 \text{ for } |\xi| \leq 1, \xi \in \mathbb{R}.$$

Then we define the *scaling function* ϕ for the digital shearlet system to be

$$\hat{\phi}(\xi_1, \xi_2) = W_0(4^{-j_L}\xi_1)V_0(4^{-j_L}\xi_2), \quad (\xi_1, \xi_2) \in \mathbb{R}^2.$$

We will later restrict this function to the pseudo-polar grid.

Let next W be the Fourier transform of the Meyer wavelet function with

$$\text{supp } W \subseteq [-4, \tfrac{1}{4}] \cup [\tfrac{1}{4}, 4] \quad \text{and} \quad W(\pm\tfrac{1}{4}) = W(\pm 4) = 0, \tag{2.12}$$

as well as

$$|W_0(4^{-j_L}\xi)|^2 + \sum_{j=j_L}^{\lceil \log_4 N \rceil} |W(4^{-j}\xi)|^2 = 1 \qquad \text{for all } |\xi| \leq N, \ \xi \in \mathbb{R}. \tag{2.13}$$

We further choose V to be a 'bump' function satisfying

$$\text{supp } V \subseteq [-1, 1] \quad \text{and} \quad V(\pm 1) = 0, \tag{2.14}$$

and also

$$|V(\xi - 1)|^2 + |V(\xi)|^2 + |V(\xi + 1)|^2 = 1 \qquad \text{for all } |\xi| \leq 1, \ \xi \in \mathbb{R}.$$

Notice that this implies

$$\sum_{s=-2^j}^{2^j} |V(2^j\xi - s)|^2 = 1 \qquad \text{for all } |\xi| \leq 1, \ \xi \in \mathbb{R} \text{ and } j \geq 0, \tag{2.15}$$

which will become important for the analysis of frame properties. For the choice of V_0, W_0, V, and W in our implementation, we refer to Section 5. Then the *generating shearlet* ψ is defined as

$$\hat{\psi}(\xi_1, \xi_2) = W(\xi_1)V(\tfrac{\xi_2}{\xi_1}), \quad (\xi_1, \xi_2) \in \mathbb{R}^2. \tag{2.16}$$

We will now define digital shearlets on Ω_R^{21} and extend the definition to the other cones by symmetry. At this time, we assume R and N are both positive, even integers and $N = 2^{n_0}$ for some integer $n_0 \in \mathbb{N}$.

For this, we first analyze the exact digitization of the coefficients of the discrete shearlet system from Subsection 1.2 for a function $J : \Omega_R \to \mathbb{C}$ by using the shearlets

ψ defined in (2.16). This will lead to the appropriate range of scales and to the support of a scaled and sheared version of the shearlet ψ.

When restricting to the cone Ω_R^{21}, the exact digitization of the coefficients of the discrete shearlet system is

$$\sum_{\omega:=(\omega_x,\omega_y)\in\Omega_R^{21}} J(\omega_x,\omega_y)2^{-j\frac{3}{2}}\overline{\hat{\psi}(S_s^T A_{4^{-j}}\omega)}e^{-2\pi i\langle A_{4^{-j}}S_s m,\omega\rangle}, \qquad (2.17)$$

where j, s, and m are to be determined. The choice of ψ leads to the coefficients

$$\sum_{\omega:=(\omega_x,\omega_y)\in\Omega_R^{21}} J(\omega_x,\omega_y)2^{-j\frac{3}{2}}\overline{W(4^{-j}\omega_x)V(s+2^j\frac{\omega_y}{\omega_x})}e^{-2\pi i\langle A_{4^{-j}}S_s m,\omega\rangle}$$

$$=\sum_{k=1}^{RN/2}\sum_{\ell=-N/2}^{N/2} J(\omega_x,\omega_y)2^{-j\frac{3}{2}}\overline{W(4^{-j}\frac{2k}{R})V(s-2^{j+1}\frac{\ell}{N})}e^{-2\pi i\langle m,S_s^T A_{4^{-j}}\omega\rangle}.$$

The support conditions (2.12) and (2.14) of W and V, respectively, imply

$$k=4^{j-1}\frac{R}{2}+n_1, \quad n_1=0,\ldots,4^{j-1}\cdot\frac{15R}{2},$$

as well as

$$\ell=2^{-j-1}N(s-1)+n_2, \quad n_2=0,\ldots,2^{-j}N,$$

if we assume k and ℓ to be positive integers.

We next analyze the support properties in radial direction. If $j<-\lceil\log(R/2)\rceil$, then $k<1$, which corresponds to only one point – the origin –, and this is dealt with by the scaling function. Hence the lowest possible scale is $j_L=-\lceil\log(R/2)\rceil$. If $j>\lceil\log_4 N\rceil$, we have $k\geq\frac{RN}{2}$. Hence the value $W(1/4)=0$ (cf. (2.12)) is placed on the boundary, and thus these scales can be omitted. This implies that the highest possible scale is $j_H:=\lceil\log_4 N\rceil$. Hence, the scaling parameter will be chosen to be

$$j\in\{j_L,\ldots,j_H\}.$$

The radial support of the windows associated with scales $j_L<j<j_H$ is

$$k=4^{j-1}\frac{R}{2}+n_1, \quad n_1=0,\ldots,4^{j-1}\cdot\frac{15R}{2}, \qquad (2.18)$$

and the radial support of the windows associated with the scales $j_L=-\lceil\log_4(R/2)\rceil$ and $j_H=\lceil\log_4 N\rceil$ is

$$\begin{aligned} k=n_1, & \quad n_1=1,\ldots,4^{j_L+1}\frac{R}{2}, & \text{for } j=j_L, \\ k=4^{j_H-1}\frac{R}{2}+n_1, & \quad n_1=0,\ldots,\frac{RN}{2}-4^{j_H-1}\frac{R}{2}, & \text{for } j=j_H. \end{aligned} \qquad (2.19)$$

We further analyze the precise support properties in angular direction. First, we examine the case $j\geq 0$. If $s>2^j$, we have $\ell\geq N/2$. Hence the value $V(-1)=0$ (cf. (2.14)) is placed on the seam line, and these parameters can be omitted. By symmetry, we also obtain $s\geq-2^j$. Thus the shearing parameter will be chosen to be

$$s\in\{-2^j,\ldots,2^j\}$$

The angular support of the windows at scale j associated with shears $-2^j<s<2^j$ is

$$\ell=2^{-j-1}N(s-1)+n_2, \quad n_2=0,\ldots,2^{-j}N, \qquad (2.20)$$

the angular support at scale j associated with the shear parameters $s_L = -2^j$ and $s_H = 2^j$ is

$$\ell = 2^{-j-1}N(s_L - 1) + n_2, \quad n_2 = 2^{-j}\tfrac{N}{2}, \ldots, 2^{-j}N, \quad \text{for } s = s_L,$$
$$\ell = 2^{-j-1}N(s_H - 1) + n_2, \quad n_2 = 0, \ldots, 2^{-j}\tfrac{N}{2}, \qquad \text{for } s = s_H. \tag{2.21}$$

For the case $j < 0$, we simply let $s = 0$ and $\ell = -N/2 + n_2$ with $n_2 = 0, \ldots, N$. Also, in this case, the window function $W(4^{-j}\omega_x)V(s + 2^j\tfrac{\omega_y}{\omega_x})$ is slightly modifed to be $W(4^{-j}\omega_x)V_0(s + 2^j\tfrac{\omega_y}{\omega_x})$ so that the tight frame property still holds.

These computations allow us to determine the support size of $W(4^{-j}\omega_x)V(s + 2^j\tfrac{\omega_y}{\omega_x})$ in terms of pairs (k, ℓ). In fact, the number of sampling points in radial and angular direction affected by a window at scale j and shear s are

$$\mathcal{L}_j^1 = \begin{cases} 4^{j+1}\tfrac{R}{2} & : \quad j = j_L, \\ 4^{j-1} \cdot \tfrac{15R}{2} + 1 & : \quad j_L < j < j_H, \\ \tfrac{RN}{2} - 4^{j-1}\tfrac{R}{2} + 1 & : \quad j = j_H, \end{cases} \tag{2.22}$$

and

$$\mathcal{L}_{j,s}^2 = \begin{cases} 2^{-j}N + 1 & : \quad -2^j < s < 2^j \ \text{with}\ j \geq 0, \\ 2^{-j}\tfrac{N}{2} + 1 & : \quad s \in \{-2^j, 2^j\} \ \text{with}\ j \geq 0, \\ N + 1 & : \quad j < 0, \end{cases} \tag{2.23}$$

respectively.

Notice that the support of the continuum function $\xi \mapsto W(4^{-j}\xi_1)V(s + 2^j\tfrac{\xi_2}{\xi_1})$ is of approximate size $4^j \times 2^j$ obeying parabolic scaling. The situation is however different in the digital realm. Since the sampling density in angular direction does change with growing radius – the sampling grid becomes in fact coarser –, parabolic scaling is not such directly mirrored in the relation between $\mathcal{L}_j^1$ and $\mathcal{L}_{j,s}^2$.

Let us next carefully examine the exponential term, which can be written as

$$e^{-2\pi i\langle m, S_s^T A_{4-j}\omega\rangle} = e^{-2\pi i\langle m, (4^{-j}\omega_x, 4^{-j}s\omega_x + 2^{-j}\omega_y)\rangle} = e^{-2\pi i\langle m, (4^{-j}\frac{2k}{R}, 4^{-j}s\frac{2k}{R} - 2^{-j}\frac{4\ell k}{RN})\rangle}.$$

We now adjust the exponential term as illustrated in Figure 2.3, which will be the only slight adaption we allow us to make when digitizing. The reason is to enable a direct application of the inverse fast Fourier transform. The necessity for this modification occurs because of two reasons:

1. We cannot make the change of variables $\tau := S_s^T A_{4-j}\omega$ in formula (2.17), which is the first step in the 'continuous' proof for tightness, due to the fact that the pseudo-polar grid is *not* invariant under the action of $S_s^T A_{4-j}$.
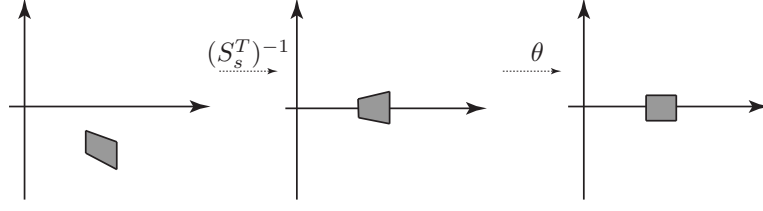2. The Fourier transform of a function defined on the pseudo-polar grid does *not* satisfy any Plancherel equation.

Defining $\theta : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ by $\theta(x, y) = (x, \tfrac{y}{x})$, we let the new exponential term be

$$e^{-2\pi i\langle m, (\theta \circ (S_s^T)^{-1})(4^{-j}\frac{2k}{R}, 4^{-j}s\frac{2k}{R} - 2^{-j}\frac{4\ell k}{RN})\rangle} = e^{-2\pi i\langle m, (4^{-j}\frac{2k}{R}, -2^{j+1}\frac{\ell}{N})\rangle}.$$

This exponential term can be rewritten as

$$e^{-2\pi i\langle m, (4^{-j}\frac{2k}{R}, -2^{j+1}\frac{\ell}{N})\rangle} = e^{-2\pi i(\frac{m_1}{4} + (1-s)m_2)}e^{-2\pi i\langle m, (4^{-j}\frac{2n_1}{R}, -2^{j+1}\frac{n_2}{N})\rangle},$$

FIG. 2.3. *Adjustment of the exponential term through the map $\theta \circ (S_s^T)^{-1}$.*

with n_1 and n_2 ranging over an appropriate set defined by (2.18), (2.19), and (2.20)–(2.21). The reformulation – recall the definitions of $\mathcal{L}_j^1$ and $\mathcal{L}_{j,s}^2$ in (2.22) and (2.23) –

$$\exp\left(-2\pi i \left\langle m, \left(\frac{\mathcal{L}_j^1 4^{-j}(2/R)}{\mathcal{L}_j^1} n_1, \frac{-\mathcal{L}_{j,s}^2 2^{j+1}(1/N)}{\mathcal{L}_{j,s}^2} n_2\right)\right\rangle\right), \quad n_1, \ n_2,$$

shows that we might regard the exponential terms as characters of a suitable locally compact abelian group (see [21]) with annihilator identified with the rectangle

$$\mathcal{R}_{j,s} = \{((\mathcal{L}_j^1)^{-1} 4^j \tfrac{R}{2} \cdot r_1, -(\mathcal{L}_{j,s}^2)^{-1} \tfrac{N}{2^{j+1}} \cdot r_2) : r_1 = 0, \ldots, \mathcal{L}_j^1 - 1, \ r_2 = 0, \ldots, \mathcal{L}_{j,s}^2 - 1\}.$$

For the low frequency square we further require the set

$$\mathcal{R} = \{(r_1, r_2) : r_1 = -1, \ldots, 1, \ r_2 = -\tfrac{N}{2}, \ldots, \tfrac{N}{2}\}.$$

We are now ready to define digital shearlets defined on the pseudo-polar grid Ω_R.

DEFINITION 2.2. *Retaining the definitions and notations derived in this subsection and Subsection 2.1, on the cone Ω_R^{21} we define* digital shearlets *at scale $j \in \{j_L, \ldots, j_H\}$, shear $s \in \{-2^j, \cdots, 2^j\}$, and spatial position $m \in \mathcal{R}_{j,s}$ by*

$$\sigma_{j,s,m}^{21}(\omega_x, \omega_y) = \frac{C(\omega_x, \omega_y)}{\sqrt{|\mathcal{R}_{j,s}|}} W(4^{-j}\omega_x) V^j\left(s + 2^j \tfrac{\omega_y}{\omega_x}\right) \chi_{\Omega_R^{21}}(\omega_x, \omega_y) e^{2\pi i \left\langle m, (4^{-j}\omega_x, 2^j \tfrac{\omega_y}{\omega_x})\right\rangle},$$

where $V^j = V$ for $j \geq 0$ and $V^j = V_0$ for $j < 0$, and

$$C(\omega_x, \omega_y) = \begin{cases} 1 & : \ (\omega_x, \omega_y) \notin \mathcal{S}_R^1 \cup \mathcal{S}_R^2, \\ \frac{1}{\sqrt{2}} & : \ (\omega_x, \omega_y) \in (\mathcal{S}_R^1 \cup \mathcal{S}_R^2) \setminus \mathcal{C}, \\ \frac{1}{\sqrt{2(N+1)}} & : \ (\omega_x, \omega_y) \in \mathcal{C}. \end{cases}$$

The shearlets $\sigma_{j,s,m}^{11}, \sigma_{j,s,m}^{12}, \sigma_{j,s,m}^{22}$ on the remaining cones are defined accordingly by symmetry with equal indexing sets for scale j, shear s, and spatial location m. For $\iota_0 = 1, 2$ and $n \in \mathcal{R}$, we further define the functions

$$\varphi_n^{\iota_0}(\omega_x, \omega_y) = \frac{C(\omega_x, \omega_y)}{\sqrt{|\mathcal{R}|}} \hat{\phi}(\omega_x, \omega_y) \chi_{\Omega_R^{\iota_0}}(\omega_x, \omega_y) e^{2\pi i \left\langle n, (\frac{k}{3}, \frac{\ell}{N+1})\right\rangle}.$$

Summarizing, we call the system

$$\mathcal{DSH} = \{\varphi_n^{\iota_0} : \iota_0 = 1, 2, n \in \mathcal{R}\} \cup \{\sigma_{j,s,m}^{\iota} : j \in \{j_L, \ldots, j_H\}, s \in \{-2^j, \cdots, 2^j\},$$
$$m \in \mathcal{R}_{j,s}, \iota = 11, 12, 21, 22\}$$

the digital shearlet system.

The just defined digital shearlet system forms indeed a tight frame for functions $J : \Omega_R \to \mathbb{C}$, as the following results shows.

THEOREM 2.3. *The digital shearlet system* $\mathcal{DSH}$ *defined in Definition 2.2 forms a tight frame for functions* $J : \Omega_R \to \mathbb{C}$.

Proof. Letting $J : \Omega_R \to \mathbb{C}$, we claim that

$$\langle J, J \rangle_{\Omega_R} = \sum_{\iota_0, n} |\langle J, \varphi_n^{\iota_0} \rangle_{\Omega_R}|^2 + \sum_{\iota, j, s, m} |\langle J, \sigma_{j,s,m}^{\iota} \rangle_{\Omega_R}|^2 \tag{2.24}$$

which proves the result.

We start by analyzing the first term on the RHS of (2.24). Let $\iota_0 \in \{1, 2\}$ and $J_C : \Omega_R \to \mathbb{C}$ be defined by $J_C(\omega_x, \omega_y) := C(\omega_x, \omega_y) \cdot J(\omega_x, \omega_y)$ for $(\omega_x, \omega_y) \in \Omega_R$. Using the support conditions of $\hat{\phi}$,

$$\sum_n |\langle J, \varphi_n^{\iota_0} \rangle_{\Omega_R}|^2 = \sum_n \Big| \sum_{(\omega_x, \omega_y) \in \Omega_R^{\iota_0}} J(\omega_x, \omega_y) \overline{\varphi_n^{\iota_0}(\omega_x, \omega_y)} \Big|^2$$

$$= \frac{1}{|\mathcal{R}|} \sum_n \Big| \sum_{(\omega_x, \omega_y) \in \Omega_R^{\iota_0}} J_C(\omega_x, \omega_y) \cdot \hat{\phi}(\omega_x, \omega_y) \cdot e^{-2\pi i \langle n, (\frac{k}{3}, \frac{\ell}{N+1}) \rangle} \Big|^2$$

$$= \frac{1}{|\mathcal{R}|} \sum_n \Big| \sum_{k=-1}^{1} \sum_{\ell=-N/2}^{N/2} J_C(\omega_x, \omega_y) \cdot \hat{\phi}(\omega_x, \omega_y) \cdot e^{-2\pi i \langle n, (\frac{k}{3}, \frac{\ell}{N+1}) \rangle} \Big|^2. \tag{2.25}$$

The choice of $\mathcal{R}$ now allows us to use the Plancherel formula. Exploiting again support properties, we conclude from (2.25) that

$$\sum_n |\langle J, \varphi_n^{\iota_0} \rangle_{\Omega_R}|^2 = \sum_{(\omega_x, \omega_y) \in \Omega_R^{\iota_0}} |C(\omega_x, \omega_y) \cdot J(\omega_x, \omega_y)|^2 \cdot |\hat{\phi}(\omega_x, \omega_y)|^2.$$

Combining $\iota_0 = 1, 2$ and using (2.11), we proved

$$\sum_{\iota_0} \sum_n |\langle J, \varphi_n^{\iota_0} \rangle_{\Omega_R}|^2 = \sum_{(\omega_x, \omega_y) \in \Omega_R} |J(\omega_x, \omega_y)|^2 \cdot |W_0(\omega_x)|^2. \tag{2.26}$$

Next we study the second term on the RHS in (2.24). By symmetry, it suffices to consider the case $\iota = 21$. By the support conditions on W and V (see (2.12) and (2.14)),

$$\sum_{j,s,m} |\langle J, \sigma_{j,s,m}^{21} \rangle_{\Omega_R}|^2 = \sum_{j,s} \sum_{m \in \mathcal{R}_{j,s}} \Big| \sum_{(\omega_x, \omega_y) \in \Omega_R^{21}} J(\omega_x, \omega_y) \overline{\sigma_{j,s,m}^{21}(\omega_x, \omega_y)} \Big|^2$$

$$= \sum_{j,s} \frac{1}{|\mathcal{R}_{j,s}|} \sum_{m \in \mathcal{R}_{j,s}} \Big| \sum_{(\omega_x, \omega_y) \in \Omega_R^{21}} J_C(\omega_x, \omega_y) \cdot \overline{W(4^{-j} \omega_x)}$$

$$\overline{\cdot V^j \big(s + 2^j \frac{\omega_y}{\omega_x}\big)} \cdot e^{-2\pi i \langle m, (4^{-j} \omega_x, 2^j \frac{\omega_y}{\omega_x}) \rangle} \Big|^2$$

$$= \sum_{j,s} \frac{1}{|\mathcal{R}_{j,s}|} \sum_{m \in \mathcal{R}_{j,s}} \Big| \sum_{k=4^{j-1}(R/2)}^{4^{j+1}(R/2)} \sum_{\ell=2^{-j-1}N(s-1)}^{2^{-j-1}N(s+1)} J_C(\omega_x, \omega_y)$$

$$\overline{\cdot W(4^{-j} \omega_x)} \cdot \overline{V^j \big(s + 2^j \frac{\omega_y}{\omega_x}\big)} \cdot e^{-2\pi i \langle m, (4^{-j} \frac{2k}{R}, -2^{j+1} \frac{\ell}{N}) \rangle} \Big|^2. \tag{2.27}$$

Similarly as before, the choice of $\mathcal{R}_{j,s}$ does allow us to use the Plancherel formula. Hence (2.27) equals

$$\sum_{j,s,m} |\langle J, \sigma_{j,s,m}^{21}\rangle_{\Omega_R}|^2 = \sum_{j,s} \sum_{(\omega_x,\omega_y)\in\Omega_R^{21}} \left| J_C(\omega_x,\omega_y) \cdot \overline{W(4^{-j}\omega_x)V^j(s+2^j\frac{\omega_y}{\omega_x})} \right|^2.$$

Next we use (2.15) to obtain

$$\sum_{j,s} \sum_{(\omega_x,\omega_y)\in\Omega_R^{21}} \left| J_C(\omega_x,\omega_y) \cdot \overline{W(4^{-j}\omega_x)} \cdot \overline{V^j(s+2^j\frac{\omega_y}{\omega_x})} \right|^2$$

$$= \sum_{(\omega_x,\omega_y)\in\Omega_R^{21}} |J_C(\omega_x,\omega_y)|^2 \sum_{j=j_L}^{j_H} |W(4^{-j}\omega_x)|^2 \cdot \sum_{s=-2^j}^{2^j} |V^j(s+2^j\frac{\omega_y}{\omega_x})|^2$$

$$= \sum_{(\omega_x,\omega_y)\in\Omega_R^{21}} |J_C(\omega_x,\omega_y)|^2 \sum_{j=j_L}^{j_H} |W(4^{-j}\omega_x)|^2.$$

Hence the second term on the RHS in (2.24) equals

$$\sum_{\iota} \sum_{j,s,m} |\langle J, \sigma_{j,s,m}^\iota\rangle_{\Omega_R}|^2 = \sum_{(\omega_x,\omega_y)\in\Omega_R} |J(\omega_x,\omega_y)|^2 \cdot \sum_{j=j_L}^{j_H} .|W(4^{-j}\omega_x)|^2. \qquad (2.28)$$

Finally, our claim (2.24) follows from combining (2.26), (2.28), and (2.13). □

3. Inverse FDST for Finite Data. In this section, we will analyze the inverse of the FDST. For this, let P and w denote the operators for the pseudo-polar Fourier transform and the weighting defined as before. By slight abuse of notation, we will further let W denote the windowing with respect to the digital shearlets, i.e., for $J : \Omega_R \to \mathbb{C}$,

$$WJ := \{\langle J, \varphi_n^{\iota_0}\rangle_{\Omega_R} : \iota_0, n\} \cup \{\langle J, \sigma_{j,s,m}^\iota\rangle_{\Omega_R} : \iota, j, s, m\}$$

with $\varphi_n^{\iota_0}$ and $\sigma_{j,s,m}^\iota$ being the digital shearlets (see Definition 2.2) for $j = j_L, \ldots, j_H$, $s \in \{-2^j, \cdots, 2^j\}$, and $n \in \mathcal{R}, m \in \mathcal{R}_{j,s}$. Then the FDST, which we abbreviate by S, takes the form

$$S = W\sqrt{w}P. \qquad (3.1)$$

To be more precise, letting I be an image of size $N \times N$ and $J_w := \sqrt{w}PI$ be the weighted pseudo-polar Fourier transform of I, the set of shearlet coefficients of I generated by the FDST can be written as

$$SI = \{c_n^{\iota_0} : \iota_0, n\} \cup \{c_{j,s,m}^\iota : \iota, j, s, m\}, \qquad (3.2)$$

where $c_n^{\iota_0} = \langle J_w, \varphi_n^{\iota_0}\rangle_{\Omega_R}$ for $\iota_0 = 0, 1$ and $c_{j,s,m}^\iota = \langle J_w, \sigma_{j,s,m}^\iota\rangle_{\Omega_R}$ for $\iota = 11, 12, 21, 22$.

Aiming to derive a closed form of S^{-1}, let $P^\star$ denote the adjoint operator of P, which, for a function $J : \Omega_R \to \mathbb{C}$, is given by

$$P^\star J(u,v) = \sum_{(\omega_x,\omega_y)\in\Omega_R} J(\omega_x,\omega_y)e^{\frac{2\pi i}{m_0}(u\omega_x+v\omega_y)}, \quad u, v = -\frac{N}{2}, \ldots, \frac{N}{2} - 1.$$

Further, let $W^\star$ denote the adjoint operator of W, which, given a sequence of shearlet coefficients $C_I = SI$ as in (3.2), is defined by

$$W^\star C_I = \sum_{\iota_0, n \in \mathcal{R}} c_n^{\iota_0} \varphi_n^{\iota_0} + \sum_{\iota, j, s, m} c_{j,s,m}^{\iota} \sigma_{j,s,m}^{\iota}.$$

Then we have the following result, which shows that the inverse FDST equals the adjoint FDST provided the weight function w satisfies (2.10).

PROPOSITION 3.1. *If w satisfies* (2.10), *then $S^{-1} = P^\star \sqrt{w} W^\star$.*

Proof. By hypothesis on w, we have $P^\star w P = \text{Id}$. Moreover, $W^\star W = \text{Id}$ by Theorem 2.3. Since $S = W \sqrt{w} P$ by (3.1), the result is proved. ∎

Let us now assume that the weight function w does not fulfill the requirements in (2.10), i.e., $P^\star w P \neq \text{Id}$, which prevents simply using the adjoint operator. In this case, we mention two methods to compute the inverse. The first method is a direct approach by resampling trigonometric polynomials from the pseudo-polar grid to the Cartesian grid, which is discussed in detail in [1]. The second method is an iterative approach by using the conjugate gradient method, which we now describe.

Suppose we are given a sequence of shearlet coefficients C_I of some image I. By Theorem 2.3, the inverse digital shearlet windowing of I is $\sqrt{w} P I = W^\star C_I =: J_w$. Consequently, the original image I can be computed by solving

$$\min_{I \in \mathbb{R}^{N \times N}} \| \sqrt{w} P I - J_w \|_2.$$

Solving this problem is equivalent to solving the linear system of equations

$$P^\star w P I = P^\star \sqrt{w} J_w. \tag{3.3}$$

This shows that w plays the role of a preconditioner for the following normal equations:

$$P^\star P I = P^\star J_w.$$

We refer to Subsection 2.2 for a selection of choices for weight functions and a discussion about their performance as preconditioners. Since the matrix corresponding to $P^\star w P$ is symmetric and positive definite, the conjugate gradient method can be used to solve (3.3). Algorithmic details will be discussed in Section 5. Let us just mention two main issues: The number of iterations required by the conjugate gradient method depends on the condition number of $P^\star w P$. Moreover, the conjugate gradient method only requires applications of $P^\star$ and P to a vector, which can be computed $O(N^2 \log N)$ flops, in contrast to forming the complete matrices.

4. Mathematical Properties of the FDST. Results on decay properties of the discrete shearlet coefficients and even more its sparse approximation properties are well-known, see [15, 23, 25]. These continuum domain results do however not directly imply similar statements for the introduced digital shearlets due to the fundamentally different nature of a digital grid. Therefore, in this section, we will analyze decay properties of the digital shearlet coefficients for linear singularities. Moreover, we will prove shear invariance of the FDST.

4.1. Decay Properties of FDST Coefficients for Linear Singularities. One main advantage of shearlets over wavelets is their ability to precisely resolve curvilinear – hence, in particular, linear – singularities due to their anisotropic shape and their additional direction-sensitive shear parameter [24]. Our computations will

show that this property carries over to the digital setting. Our model for a linear singularity will be a line. We remark that similar results can be shown for a Heaviside functions as a model.

Let I be an image of size $N \times N$ with an edge through the origin of slope t satisfying $|t| < 1$, i.e., $I(u,v) = \delta(tu - v)$, $-N/2 \leq u, v \leq N/2 - 1$. Figure 4.1 indicates that
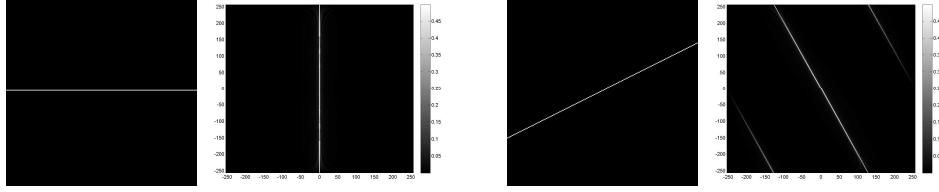

FIG. 4.1. *From left to right: A horizontal line in spatial domain and its pseudo-polar Fourier transform, and a line with slope < 1 in spatial domain and its pseudo-polar Fourier transform.*

for $t \neq 0$ an aliasing effect occurs. This effect is quite small in comparison with the intensity of the line, wherefore we will ignore it in the calculations. Also, by symmetry, we can assume $t \geq 0$ without loss of generality. Setting $m_0 = \frac{2}{R}(RN + 1)$, for $(\omega_x, \omega_y) = (\frac{2k}{R}, -\frac{2k}{R} \cdot \frac{2\ell}{N}) \in \Omega_R^2$, we have

$$\hat{I}(\omega_x, \omega_y) = \sum_{u=-N/2}^{N/2-1} I(u,tu)e^{-\frac{2\pi i}{m_0}(u\omega_x + tu\omega_y)} = \sum_{u=-N/2}^{N/2-1} e^{-\frac{2\pi i}{RN+1}u\cdot k\cdot(1-\frac{2\ell}{N}t)}.$$

If $k = 0$, we obtain $\hat{I}(\omega_x, \omega_y) = N$. If $k \neq 0$,

$$\hat{I}(\omega_x, \omega_y) = e^{\frac{\pi i k}{RN+1}(1-\frac{2\ell}{N}t)} \cdot \frac{\sin(\pi N k(1-\frac{2\ell}{N}t)/(RN+1))}{\sin(\pi k(1-\frac{2\ell}{N}t)/(RN+1))}.$$

Concluding,

$$\hat{I}(\omega_x, \omega_y) = N, \qquad\qquad \text{for } k = 0$$

$$|\hat{I}(\omega_x, \omega_y)| \leq \frac{RN+1}{2|k||1-\frac{2\ell}{N}t|}, \quad \text{for } |k| = 1, \ldots, \frac{RN}{2}.$$

Similarly, for $(\omega_x, \omega_y) = (-\frac{2k}{R} \cdot \frac{2\ell}{N}, \frac{2k}{R}) \in \Omega_R^1$,

$$\hat{I}(\omega_x, \omega_y) = N, \qquad\qquad \text{for } \ell = t \cdot \frac{N}{2}$$

$$|\hat{I}(\omega_x, \omega_y)| \leq \frac{(RN+1)N}{4|k||\ell - \frac{N}{2}t|}, \quad \text{for } |k| = 1, \ldots, \frac{RN}{2}, \ \ell \neq t \cdot \frac{N}{2}.$$

We now distinguish two cases. On Ω_R^{21} (similarly on Ω_R^{22}), for those digital shearlets *not* on the seam lines, i.e., $s \notin \{-2^j, 2^j\}$, we obtain

$$|\langle \hat{I}, \sigma_{j,s,m}^{21} \rangle_{\Omega_R}| \leq \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{(\omega_x, \omega_y) \in \text{supp } \sigma_{j,s,0}^{21}} |\hat{I}(\omega_x, \omega_y)|$$

$$\leq \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{\ell=2^{-j-1}N(s-1)}^{2^{-j-1}N(s+1)} \frac{N/2}{|N/2-t\ell|} \sum_{k=4^{j-1}R/2}^{4^{j+1}R/2} \frac{RN+1}{2|k|}$$

$$= \frac{(RN+1)N}{4\sqrt{|\mathcal{R}_{j,s}|}} \sum_{\ell=2^{-j-1}N(s-1)}^{2^{-j-1}N(s+1)} \frac{1}{|N/2-t\ell|} \sum_{k=4^{j-1}R/2}^{4^{j+1}R/2} \frac{1}{|k|}$$

$$\leq \frac{(RN+1)N}{4\sqrt{|\mathcal{R}_{j,s}|}} \sum_{\ell=2^{-j-1}N(s-1)}^{2^{-j-1}N(s+1)} \frac{1}{N/2(1-2^{-j}(s+1)t)} \sum_{k=4^{j-1}R/2}^{4^{j+1}R/2} \frac{1}{|k|}$$

$$\leq \frac{(RN+1)N}{4\sqrt{|\mathcal{R}_{j,s}|}} \cdot \frac{1}{2^j(1-t)} \cdot 4\log(2). \tag{4.1}$$

On Ω_R^{11} (similarly on Ω_R^{12}),

$$|\langle \hat{I}, \sigma_{j,s,m}^{11} \rangle_{\Omega_R}| \leq \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{(\omega_x,\omega_y)\in\text{supp } \sigma_{j,s,0}^{11}} |\hat{I}(\omega_x,\omega_y)|$$

$$\leq \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{\ell=2^{-j-1}N(s-1)}^{2^{-j-1}N(s+1)} \sum_{k=4^{j-1}R/2}^{4^{j+1}R/2} \Big(\delta(s-t\cdot\tfrac{N}{2})\delta(\ell-t\cdot\tfrac{N}{2})N$$

$$+(1-\delta(s-t\cdot\tfrac{N}{2})\delta(\ell-t\cdot N/2))\frac{(RN+1)N}{4|k||\ell-t\cdot\tfrac{N}{2}|}\Big)$$

$$\leq \frac{(RN+1)N}{4\sqrt{|\mathcal{R}_{j,s}|}} \sum_{\ell=2^{-j-1}N(s-1),\ell\neq t\cdot\frac{N}{2}}^{2^{-j-1}N(s+1)} \frac{1}{|\ell-t\cdot N/2|} \sum_{k=4^{j-1}R/2}^{4^{j+1}\cdot\frac{R}{2}+1} \frac{1}{|k|}$$

$$+\frac{\delta(s-t\cdot\tfrac{N}{2})\cdot(4^{j+1}\cdot\tfrac{R}{2}+1)N}{\sqrt{|\mathcal{R}_{j,s}|}}$$

$$\leq \frac{(RN+1)N\log(N/2)4\log(2)}{4\sqrt{|\mathcal{R}_{j,s}|}} + \frac{\delta(s-t\cdot\tfrac{N}{2})(4^{j+1}\cdot\tfrac{R}{2}+1)N}{\sqrt{|\mathcal{R}_{j,s}|}} \tag{4.2}$$

Digital shearlets on the seam lines can be dealt with similarly. Thus, by (4.1) and (4.2), we have upper bounds for shearlet coefficients which as $j \to \infty$ have asymptotic decay

- $O(2^{-3j/2})$, if the shearlet is in Ω_R^2 and is not aligned with the line singularity,
- $O(2^{-j/2})$, if the shearlet is in Ω_R^1 and is not aligned with the line singularity,
- $O(2^{3j/2})$, if the shearlet is in Ω_R^1 and is aligned with the line singularity.

We now aim to show that line singularities can indeed be detected in the transform domain. For this, we will show a lower bound on the decay of the shearlet coefficients if the shearlet is aligned with the line singularity. To avoid unnecessary technicalities, we now assume $\hat{I}(\omega_x,\omega_y) = N$ for $\omega_y/\omega_x = -1/t$ and $\hat{I}(\omega_x,\omega_y) = 0$ for $\omega_y/\omega_x \neq -1/t$ with $0 < t < 1$. The case for $t \leq 0$ can be handled similarly. Notice that this condition merely assumes that the aliasing effects are negligible for the decay analysis. We further assume that the window function W and the bump function V are positive on their supports. On Ω_R^{11},

$$\max_{j,s,m}\{|\langle \hat{I}, \sigma_{j,s,m}^{11} \rangle_{\Omega_R}|\} \geq |\langle \hat{I}, \sigma_{j,s,0}^{11} \rangle_{\Omega_R}|$$

$$= \left| \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{(\omega_x,\omega_y)\in\Omega_R^{11}} \hat{I}(\omega_x,\omega_y)\overline{W(4^{-j}\omega_x)V^j(s+2^j\tfrac{\omega_y}{\omega_x})} \right|$$

$$= \left| \frac{1}{\sqrt{|\mathcal{R}_{j,s}|}} \sum_{(\omega_x, \omega_y) \in \Omega_R^{21}} \delta(s - t \cdot \tfrac{N}{2}) \cdot N \cdot \overline{W(4^{-j}\omega_x) V^j(s + 2^j \tfrac{\omega_y}{\omega_x})} \right|$$

$$= \frac{\delta(s - t \cdot \tfrac{N}{2}) \cdot N}{\sqrt{|\mathcal{R}_{j,s}|}} S_{j,s},$$

where $S_{j,s} = |\sum_{(\omega_x, \omega_y) \in \Omega_R} W(4^{-j}\omega_x) V(s + 2^j \tfrac{\omega_y}{\omega_x})|$ denotes the area of the window function with respect to j and s. Since $S_{j,s}$ is approximately the same order as $|\mathcal{R}_{j,s}|$, we conclude that

$$\max_{j,m}\{|\langle \hat{I}, \sigma_{j,tN/2,m}^{11} \rangle_{\Omega_R}|\} \geq \frac{N \cdot S_{j,s}}{\sqrt{|\mathcal{R}_{j,s}|}} = O(2^{j/2}) \quad \text{as } j \to \infty.$$

This estimate combined with (4.2) shows that the asymptotic decay as $j \to \infty$ in $\Omega_R^{1\cdot}$ is

- $O(2^{-j/2})$, if the shearlet is not aligned with the line singularity,
- $\Omega(2^{j/2})$, if the shearlet is aligned with the line singularity.

In this sense there is a strong difference between the decay rates of shearlet coefficients between those aligned with the line singularity and those *not* aligned with the line singularity.

Experimental results strongly support this analysis. Table 4.1 presents the maximal absolute values of shearlet coefficients of an 512×512-image with a horizontal line ($t = 0$) and different scales, both aligned with the line singularity (c_{max}^0), and not aligned with the line singularity (c_{max}^1). The data in Table 4.1 does clearly indicate a significant difference of decay rates between these two classes. We shall confirm this behavior from a different viewpoint, more precisely, a particular quantitative measure, in Subsection 7.5.

TABLE 4.1
Maximal coefficients of a horizontal line

j	-1	0	1	2	3	4	5
c_{max}^0	0.160	0.098	0.068	0.038	0.018	0.013	2.6E-3
c_{max}^1	0.149	0.048	0.007	1.9E-3	4.8E-4	2.3E-4	2.4E-5

4.2. Shear Invariance of the FDST. Since the shearing operator is quite distinctive in the definition of shearlets, one might ask whether the FDST is in fact even shear invariant. In the continuum setting, when choosing ψ to be a shearlet generator as defined in (1.1), one can easily verify that, for any $f \in L^2(\mathbb{R}^2)$,

$$\langle 2^{3j/2}\psi(S_k^{-1}A_{4^j} \cdot -m), f(S_s \cdot) \rangle = \langle 2^{3j/2}\psi(S_{k+2^j s}^{-1}A_{4^j} \cdot -m), f \rangle.$$

This identity can be viewed as a manifestation of shear invariance, since it states that the shearlet coefficient of a sheared image $f(S_s \cdot)$ at scale j, shear k, and spatial position m equals the shearlet coefficient for the original image f at the same scale j, same spatial position m, but with a shear parameter shifted by $2^j s$.

The following results shows that the FDST associated with digital shearlets has a similar property.

PROPOSITION 4.1. *Let I be an $N \times N$ image . Let j be a scale, s be a shear, t with $|t| < 1$ be a slope, and $m = (m_1, m_2)$ be a spatial position such that $2^j t \in \mathbb{Z}$*

and $-2^j < s, s + 2^j t < 2^j$. *Let* $I_t := I(S_t \cdot)$ *be the sheared image of* I *such that* $\hat{I}_t(\omega_x, \omega_y) = \hat{I}(\omega_x, \omega_y - t\omega_x)$ *for all* $(\omega_x, \omega_y) \in \text{supp } \sigma^\iota_{j,s,m}$. *Then*

$$\langle \hat{I}_t, \sigma^\iota_{j,s,m} \rangle_{\Omega_R} = \langle \hat{I}, \sigma^\iota_{j,s+2^j t,m} \rangle_{\Omega_R} \cdot e^{-2\pi i m_2 \cdot 2^j t}.$$

Proof. It is sufficient to prove the claim for the case $\iota = 21$, since the other cones can be handled similarly. In this case, $(\omega_x, \omega_y) = (\frac{2k}{R}, \frac{2k}{R} \cdot \frac{-2\ell}{N})$ and hence

$$(\tilde{\omega}_x, \tilde{\omega}_y) := (\omega_x, \omega_y - t\omega_x) = (\frac{2k}{R}, \frac{2k}{R} \cdot \frac{-2(\ell + tN/2)}{N}).$$

Since $\hat{I}_t = \hat{I}((S_t^{-1})^T \cdot)$ on the support of $\sigma^\iota_{j,s,m}$, we have

$$\sqrt{|\mathcal{R}_{j,s}|} \cdot \langle \hat{I}_t, \sigma^{21}_{j,s,m} \rangle_{\Omega_R}$$
$$= \sum_{(\omega_x, \omega_y) \in \Omega_R^{21}} \hat{I}_t(\omega_x, \omega_y) \overline{W(4^{-j}\omega_x)V^j(s + 2^j \frac{\omega_y}{\omega_x})} e^{-2\pi i \langle m, (4^{-j}\omega_x, 2^j \frac{\omega_y}{\omega_x}) \rangle}$$
$$= \sum_{(\tilde{\omega}_x, \tilde{\omega}_y) \in (S_t^{-1})^T \Omega_R^{21}} \hat{I}(\tilde{\omega}_x, \tilde{\omega}_y) \overline{W(4^{-j}\tilde{\omega}_x)V^j(s + 2^j t + 2^j \frac{\tilde{\omega}_y}{\tilde{\omega}_x})} e^{-2\pi i \langle m, (4^{-j}\omega_x, 2^j \frac{\omega_y}{\omega_x}) \rangle}$$
$$= \sum_{(\tilde{\omega}_x, \tilde{\omega}_y) \in (S_t^{-1})^T \Omega_R^{21}} \hat{I}(\tilde{\omega}_x, \tilde{\omega}_y) \overline{W(4^{-j}\tilde{\omega}_x)V^j(s + 2^j t + 2^j \frac{\tilde{\omega}_y}{\tilde{\omega}_x})} e^{-2\pi i \langle m, (4^{-j}\tilde{\omega}_x, 2^j \frac{\tilde{\omega}_y}{\tilde{\omega}_x} + 2^j t) \rangle}$$
$$= \sqrt{|\mathcal{R}_{j,s}|} \cdot \langle \hat{I}, \sigma^{21}_{j,s+2^j t,m} \rangle e^{-2\pi i m_2 \cdot 2^j t}.$$

This proves the claim. □

5. Implementation of the FDST and its Inverse. After the formal introduction of the FDST and its theoretical analysis, we now turn to discuss the details of our implementation, including the forward transform, the adjoint transform, and the inverse transform. The associated code `ShearLab-PPFT-1.0` can be downloaded from `www.ShearLab.org`.

Figure 5.1 provides an overview of the main steps of of the FDST and its inverse, which were defined in Sections 2 and 3, respectively.
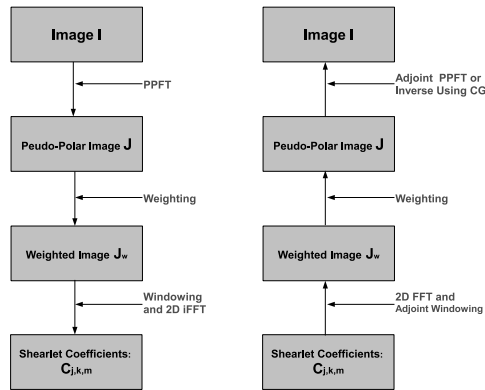

FIG. 5.1. *Flowcharts of the FDST (left) and its inverse (right).*

5.1. Choices of Parameters. Before discussing the implementation details of each transform, let us start by elaborating on the choices of the main parameters involved in our design of the FDST.

5.1.1. Choice of the Parameter m_0. The definition of the pseudo-polar Fourier transform (2.5) has m_0 as a free parameter. In the paper [1], in which only the oversampling rate $R = 2$ was considered, m_0 was chosen to be $2N + 1$. This particular choice enabled utilization of the 1D-FFT for the implementation of the fast pseudo-polar Fourier transform. In the situation of an arbitrary oversampling rate R, which we consider, the fast pseudo-polar Fourier transform is based on a (discrete) fractional Fourier transform. In order to utilize the 1D-FFT along one direction, m_0 has necessarily to be chosen as $\frac{2}{R}(RN + 1)$; for details see Subsection 5.2. When choosing a different parameter m_0, the fractional Fourier transform needs to be applied on both directions, which certainly would significantly lower the speed of the FDST. The (discrete) fractional Fourier transform can be implemented with the same complexity as the 1D-FFT, but a different constant; in fact it is about 5 times slower than the 1D-FFT, see [2]. To accelerate the speed, in our `ShearLab` package, the default m_0 is consequently set to be $\frac{2}{R}(RN + 1)$.

5.1.2. Choice of Weight Function w. A second free parameter is the weight function w, for which the criterion for isometry in Theorem 2.1 is required; and we discussed some choices in Subsection 2.2.2. As could be seen, the performance in terms of almost isometry M_{isom} differs depending on the types of images which are considered. Hence the weight needs to chosen depending on the application. It should be also emphasized that the particular type of weighting we considered in Subsection 2.2.2 provide good preconditioners for the conjugate gradient method, in case an even more accurate inverse than the adjoint is required. In our `ShearLab` package, various choices of w are available.

5.1.3. Choice of Window Functions W_0, V_0 and W, V. The final main parameter is the choice of the window functions W_0, V_0 and W, V. In our implementation, we use Meyer wavelets and define W_0 and W to be the Fourier transform of the Meyer scaling function and wavelet function, respectively, i.e.,

$$W_0(\xi) = \begin{cases} 1 & : \quad |\xi| \leq \frac{1}{4}, \\ \cos\left[\frac{\pi}{2}\nu(\frac{4}{3}|\xi| - \frac{1}{3})\right] & : \quad \frac{1}{4} \leq |\xi| \leq 1, \\ 0 & : \quad \text{otherwise}, \end{cases}$$

and

$$W(\xi) = \begin{cases} \sin\left[\frac{\pi}{2}\nu(\frac{4}{3}|\xi| - \frac{1}{3})\right] & : \quad \frac{1}{4} \leq |\xi| \leq 1, \\ \cos\left[\frac{\pi}{2}\nu(\frac{1}{3}|\xi| - \frac{1}{3})\right] & : \quad 1 \leq |\xi| \leq 4, \\ 0 & : \quad \text{otherwise}, \end{cases}$$

where ν is a C^k function or C^∞ function such that

$$\nu(x) = \begin{cases} 0 & : \quad x \leq 0, \\ 1 - \nu(1 - x) & : \quad 0 \leq x \leq 1, \\ 1 & : \quad x \geq 1. \end{cases} \tag{5.1}$$

In `ShearLab`, ν is set to be $\nu(x) = 2x^2$ for $0 \leq x \leq 1/2$ and $\nu(x) = 1 - 2(1 - x)^2$ for $1/2 \leq x \leq 1$, which is a C^1 function. Other choices are also available. The choice of ν then fixes W_0 and W. Since $|W_0(\xi)|^2 + |W(\xi)|^2 = 1$ for $|\xi| \leq 1$, the required condition (2.13) is satisfied. These choices for W_0, W, and ν are illustrated in Figure 5.2.
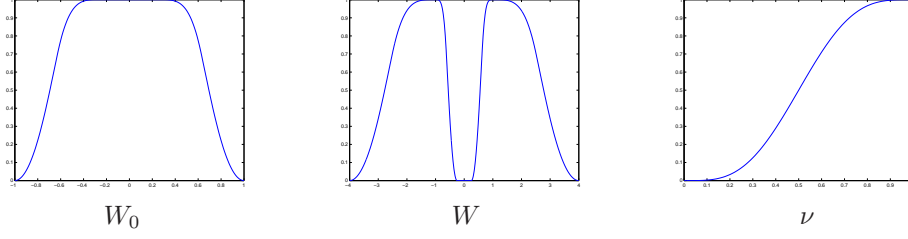
FIG. 5.2. *The graphs of W_0, W, and ν.*

Provided ν satisfies (5.1), it can be used to design the bump function V by setting $V(\xi) = \sqrt{\nu(1+\xi) + \nu(1-\xi)}$ for $-1 \le \xi \le 1$. This automatically satisfies (2.15), and is our choice in the implementation. The function V_0 is defined to be $V_0 \equiv 1$.

5.2. FDST. We now present our implementation of the FDST, which mainly consists of three parts: The fast pseudo-polar Fourier transform, the weighting on the pseudo-polar grid, and the digital shearlet windowing on the pseudo-polar grid followed by 2D-iFFT. These parts correspond to the operators P, w, and W introduced and discussed in Section 3. To facilitate our presentation of the implementation details, we require the following operators:

The *(unaliased) fractional Fourier transform* (frFT) of a vector $c \in \mathbb{C}^{N+1}$ with respect to a fraction $\alpha \in \mathbb{C}$ is defined to be

$$(F_{N+1}^{\alpha} c)(k) := \sum_{j=-N/2}^{N/2} c(j) e^{-2\pi i \cdot j \cdot k \cdot \alpha}, \quad k = -\tfrac{N}{2}, \ldots, \tfrac{N}{2}.$$

It was shown in [2], that the fractional Fourier transform $F_{N+1}^{\alpha} c$ can be computed using $O(N \log N)$ operations. In the special case of $\alpha = 1/(N+1)$, the fractional Fourier transform becomes the (unaliased) 1D Fast Fourier Transform (1D-FFT), which in the sequel we will denote by F_1. Similarly, we will denote the 2D Fast Fourier Transform (2D-FFT) by F_2, and the inverse of the F_2 by F_2^{-1} (2D-iFFT).

Let now N be even and $m > N$ be an odd integer. Then the *padding operator* $E_{m,n}$ acting on a vector $c \in \mathbb{C}^N$ gives a symmetrically zero padding version of c in the sense that

$$(E_{m,N} c)(k) = \begin{cases} c(k) & k = -\tfrac{N}{2}, \ldots, \tfrac{N}{2} - 1, \\ 0 & k \in \{-\tfrac{m}{2}, \ldots, \tfrac{m}{2}\} \setminus \{-\tfrac{N}{2}, \ldots, \tfrac{N}{2} - 1\}. \end{cases}$$

The fast pseudo-polar Fourier transform was introduced in [1], however only for the oversampling rate $R = 2$. We will next extend this algorithm to an arbitrary oversampling rate and show that its complexity is also $O(N^2 \log N)$. For this, let I be an image of size $N \times N$. We restrict to the cone Ω_R^1. Choosing $m_0 = \tfrac{2}{R}(RN+1)$ to utilize the 1D-FFT, for $(\omega_x, \omega_y) \in \Omega_R^1$,

$$\hat{I}(\omega_x, \omega_y) = \sum_{u,v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i}{m_0}(u\omega_x + v\omega_y)} = \sum_{u=-N/2}^{N/2-1} \sum_{v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i}{m_0}(u\frac{-4k\ell}{RN} + v\frac{2k}{R})}$$

$$= \sum_{u=-N/2}^{N/2-1} \left(\sum_{v=-N/2}^{N/2-1} I(u,v) e^{-\frac{2\pi i v k}{RN+1}} \right) e^{-2\pi i u \ell \cdot \frac{-2k}{(RN+1)\cdot N}}.$$

The above identity shows that $\hat{I}$ on Ω_R^1 can be obtained by performing 1D-FFT on the extension of I along direction v and then applying the frFT along direction u. More precisely, let $\tilde{I}(u, \cdot) := E_{RN+1,N} I(u, \cdot), -N/2 \le u \le N/2 - 1$ be the symmetric zero padding of I and let $\tilde{I}_1(u, \cdot) := F_1 \tilde{I}(u, \cdot)$ be the 1D FFT of $\tilde{I}$ along direction v. Also, let $\tilde{I}_2(\cdot, k) := E_{N+1,N} \tilde{I}_1(\cdot, k)$ for $k = -\frac{RN}{2}, \ldots, \frac{RN}{2}$ be the symmetric zero padding of $\tilde{I}_1$ along direction u. Note that, $\tilde{I}_1$ is then of size $(RN + 1) \times N$ and $\tilde{I}_2$ is of size $(RN + 1) \times (N + 1)$.Then the above identity can be written as

$$\hat{I}(\omega_x, \omega_y) = \sum_{u=-N/2}^{N/2-1} \tilde{I}_1(u, k) e^{-2\pi i u \ell \cdot \frac{-k}{(RN+1) \cdot N/2}} = \sum_{u=-N/2}^{N/2} \tilde{I}_2(u, k) e^{-2\pi i u \ell \cdot \frac{-2k}{(RN+1) \cdot N}}$$
$$= (F_{N+1}^{\alpha_k} \tilde{I}_2(\cdot, k))(\ell),$$

where $\alpha_k = -\frac{k}{(RN+1)N/2}$ is the fraction in the frFT.

Thus, the pseudo-polar Fourier transform $\hat{I}(\omega_x, \omega_y), (\omega_x, \omega_y) \in \Omega_R^1$ – similarly for Ω_R^2 – for arbitrary oversampling rate R can be computed by the steps described in Algorithm 1.

Algorithm 1 Fast Pseudo-Polar Fourier Transform of I on Ω_R^1

(a) **Input**: Image I of size $N \times N$.

(b) **Output**: The pseudo-polar Fourier transform $\hat{I}_{\Omega_R^1}$.

(c) **PPFT**: Image $I(u, v), -N/2 \le u, v \le N/2 - 1$.

1: $\tilde{I}(u, \cdot) \leftarrow E_{RN+1,N} I(u, \cdot), -N/2 \le u \le N/2 - 1$. Symmetrically padding the image I along direction v to obtain an image $\tilde{I}$ of size $(RN + 1) \times N$.

2: $\tilde{I}_1(u, \cdot) \leftarrow F_1 \tilde{I}(u, \cdot), -N/2 \le u \le N/2 - 1$. For each vector of $\tilde{I}$ along direction v, perform the 1D-FFT along direction v to get $\tilde{I}_1$.

3: $\tilde{I}_2(\cdot, k) \leftarrow E_{N+1,N} \tilde{I}_1(\cdot, k), -RN/2 \le k \le RN/2$. Symmetrically padding the image $\tilde{I}_1$ along direction u to get image $\tilde{I}_2$ whose size along direction u is $N + 1$.

4: $\hat{I}_{\Omega_R^1} \leftarrow F_{N+1}^{\alpha_k} \tilde{I}_2(\cdot, k), -RN/2 \le k \le RN/2$. Perform the fractional Fourier transform with respect to a fraction $\alpha_k = -\frac{k}{(RN+1)N/2}$ along direction u.

Since the 1D-FFT and 1D-frFT require only $O(N \log N)$ operations for a vector of size N, the total complexity of the pseudo-polar Fourier transform is $O(N^2 \log N)$ for an image of size $N \times N$. It should be emphasized that our Algorithm 1 is more simple and efficient even in case $R = 2$ than the algorithm described in [1], in which F_2 is applied to $\tilde{I}$ followed by the application of F_1^{-1} to the resulting image in order to obtain $\tilde{I}_1$. The algorithm in [1] is slightly more redundant while ours is more efficient by combining these two steps to one step. Moreover, the fractional Fourier transform along direction u is only of size $N + 1$, as compared to size $RN + 1$ in [1].

We would like to also remark that, for a different choice of constant m_0, one can compute the pseudo-polar Fourier transform also with complexity $O(N^2 \log N)$ for an image of size $N \times N$, in which case the fractional Fourier transform needs to be applied in both directions u and v of the image. For example, for computing $\hat{I}$ on Ω_R^1, the fractional Fourier transform is first applied to I along direction v with a fixed fraction constant $\alpha = \frac{1}{m_0 R/2}$, and then to the another direction u with fractions $\alpha_k = \frac{-k}{m_0 R/2 \cdot N/2}$ depending on k. In this case, it is, of course, slower than the special

choice $m_0 = \frac{2}{R}(RN + 1)$, since frFTs are involved in both directions while the special choice m_0 utilizes FFT.

The next step, which is an application of the weight function, is simply a pointwise multiplication on the pseudo-polar grid, i.e.,

(d) **Weighting**: Let $J : \Omega_R \to \mathbb{C}$ be the pseudo-polar image of I and $w : \Omega_R \to \mathbb{R}^+$ be any suitable weight function on Ω_R. Compute the point-wise multiplication $J_w(\omega_x, \omega_y) = J(\omega_x, \omega_y) \cdot \sqrt{w(\omega_x, \omega_y)}$, $(\omega_x, \omega_y) \in \Omega_R$.

For the windowing, a sequence of subbands $\{\varphi_0^{\iota_0} : \iota_0\} \cup \{\sigma_{j,s,0}^\iota : j, s, \iota\}$ is computed, whose frequency tiling covers the pseudo-polar grid. As discussed in Subsections 2.3 and 5.1, these subbands are constructed using Meyer scaling functions and wavelets. The weighted image is then windowed by each subband, followed by an application of the 2D-iFFT to each windowed subimage. This results in a sequence of coefficients – the digital shearlet coefficients – structured block by block:

(e) **Subband Windowing**: Compute the digital shearlet coefficients by windowing J_w with respect to the digital shearlets $\{\varphi_n^{\iota_0} : \iota_0, n\} \cup \{\sigma_{j,s,m}^\iota : j, s, m, \iota\}$ defined in Definition 2.2.

For the detailed step-by-step description of the algorithm for the FDST used in `ShearLab`, we refer to Algorithm A.1 (see Appendix).

5.3. Adjoint FDST. The algorithm for the adjoint shearlet transform can be straightforwardly derived from the FDST, and the main idea was already described in Section 3. It should be noted that the adjoint fractional Fourier transform for a vector $c \in \mathbb{C}^{N+1}$ with respect to a constant $\alpha \in \mathbb{C}$ is given by $F_{N+1}^{-\alpha} c$. Moreover, for $m > N$ the adjoint operator $E_{m,N}^\star$ for the padding operator $E_{m,N}$ is given by $(E_{m,N}^\star c)(k) = c(k)$, $k = -N/2, \ldots, N/2 - 1$ for a vector $c \in \mathbb{C}^m$. The adjoint shearlet transform can be computed with a complexity $O(N^2 \log N)$ similar to the FDST, since it is obtained simply by 'running the FDST backwards'.

For the detailed step-by-step description of the algorithm for the adjoint FDST used in `ShearLab`, we refer to Algorithm A.2.

5.4. Inverse FDST. The main idea to use the conjugate gradient method was already described in Section 3, and for a detailed step-by-step description of the algorithm for the inverse FDST used in `ShearLab`, we refer to Algorithm A.3.

6. Quality Measures for Algorithmic Realization. To ensure and also prove that our implementation satisfies the previously proposed desiderata, we will now define quality measures for each of those and in the sequel provide numerical results on how accurate our implementation satisfies these. It is moreover our hope that these measures shall serve as comparison measures for future implementations. Although some measures are stated in 'shearlet language', most are applicable to any directional transform based on parabolic scaling.

In the following, P shall denote the pseudo-polar Fourier transform defined by Algorithm 1), w shall denote the weighting applied to the values on the pseudo-polar grid, W shall be the windowing with additional 2D-iFFT, S shall denote the FDST defined by Algorithm A.1, and $S^\star$ its adjoint defined by Algorithm A.2. We will further use the notation $G_A J$ for the solution of a matrix problem $AI = J$ using conjugate gradient method with residual error set to be 10^{-6}.

[D1] ALGEBRAIC EXACTNESS.

Comments: We require the transform to be the precise implementation of a theory for digital data on a pseudo-polar grid. In addition, to ensure numerical accuracy, we

provide the following test, which provides a quantitative measure for the closeness of the windows to form a tight frame.

Measure: Generate a sequence of 5 random images $I_1, \ldots, I_5$ (the integer 5 is chosen for the purpose of fixing the number of images to enable precise comparison) on the pseudo-polar grid for $N = 512$ and $R = 8$ with standard normally distributed entries. Our quality measure will then be the Monte Carlo estimate for the operator norm $\|W^\star W - \mathrm{Id}\|_{op}$ given by

$$M_{alg} = \max_{i=1,\ldots,5} \frac{\|W^\star W I_i - I_i\|_2}{\|I_i\|_2}.$$

[D2] ISOMETRY OF PSEUDO-POLAR FOURIER TRANSFORM.

Comments: To ensure isometry of the utilized pseudo-polar Fourier transform, we introduced a careful weighting of the pseudo-polar grid. The following test will now measure the closeness to being an isometry. We expect to see a trade-off between the oversampling rate and the closeness to being an isometry. Since the measure shall however serve as a common ground to compare different algorithms based on parabolic scaling, we do not take the oversampling rate into account in the proposed measure. Instead, we would like to remind the reader that this rate will instead affect the measure for speed [D6]. In the sequel, we will now provide three different measures, each being designed to test a different aspect.

Measure:
- _Closeness to tightness._ Generate a sequence of 5 random images $I_1, \ldots, I_5$ of size 512×512 with standard normally distributed entries. Our quality measure will then be the Monte Carlo estimate for the operator norm $\|P^\star w P - \mathrm{Id}\|_{op}$ given by

$$M_{isom_1} = \max_{i=1,\ldots,5} \frac{\|P^\star w P I_i - I_i\|_2}{\|I_i\|_2}.$$

- _Quality of preconditioning._ Our quality measure will be the spread of the eigenvalues of the Gram operator $P^\star w P$ given by

$$M_{isom_2} = \frac{\lambda_{\max}(P^\star w P)}{\lambda_{\min}(P^\star w P)}.$$

- _Invertibility._ Our quality measure will be the Monte Carlo estimate for the invertibility of the operator $\sqrt{w} P$ using conjugate gradient method $G_{\sqrt{w} P}$ given by

$$M_{isom_3} = \max_{i=1,\ldots,5} \frac{\|G_{\sqrt{w} P} \sqrt{w} P I_i - I_i\|_2}{\|I_i\|_2}.$$

[D3] TIGHT FRAME PROPERTY.

Comments: We now combine [D1] and [D2] to allow comparison with other transforms, since those singleton tests might not be possible for any transform due to a different inner structure.

Measure: Generate a sequence of 5 random images $I_1, \ldots, I_5$ of size 512×512 with standard normally distributed entries.

- *Adjoint transform.* Our quality measure will be the Monte Carlo estimate for the operator norm $\|S^\star S - \mathrm{Id}\|_{op}$ given by

$$M_{tight_1} = \max_{i=1,\ldots,5} \frac{\|S^\star S I_i - I_i\|_2}{\|I_i\|_2}.$$

- *Inverse transform.* Our quality measure will be the Monte Carlo estimate for the operator norm $\|G_{\sqrt{w}P} W^\star S - \mathrm{Id}\|_{op}$ given by

$$M_{tight_2} = \max_{i=1,\ldots,5} \frac{\|G_{\sqrt{w}P} W^\star S I_i - I_i\|_2}{\|I_i\|_2}.$$

[D4] SPACE-FREQUENCY-LOCALIZATION.

Comments: The purpose of this test is to provide quantitative measures for the degree to which the windows – in our case the digital shearlets – are localized in both space and frequency. For this, we numerically measure mathematically precise notions of both decay and smoothness.

Measure: Let I be a shearlet in a 512×512 image centered at the origin $(257, 257)$ with slope $s = 0$ at scale $j = 3$, which in our case is the shearlet $\sigma_{3,0,0}^{11} + \sigma_{3,0,0}^{12}$. Our quality measure will then be four-fold:

- *Decay in spatial domain.* Compute the decay rates $d_1, \ldots, d_{512}$ along lines parallel to the y-axis starting from the line $[257, :]$ and the decay rates $d_{512}, \ldots, d_{1024}$ with x and y interchanged in the following way: Consider exemplarily the line $[257 : 512, 1]$. First compute the smallest monotone majorant $M(x, 1)$, $x = 257, \ldots, 512$ – note that we could have also chosen a different 'envelope' – for the curve $|I(x, 1)|$, $x = 257, \ldots, 512$. Then the decay rate along the line $[257 : 512, 1]$ is defined to be the average slope of the line, which is a least square fit to the curve $\log(M(x, 1))$, $x = 257, \ldots, 512$. Based on these decay rates, we choose our measure to be the average of the decay rates given by

$$M_{decay_1} = \frac{1}{1024} \sum_{i=1,\ldots,1024} d_i.$$

- *Decay in frequency domain.* To check whether the Fourier transform of I is compactly supported and to check its decay rate, let $\hat{I}$ be the 2D-FFT of I and compute the decay rates d_i, $i = 1, \ldots, 1024$ as before. Then we define the following two measures:
  - ◇ *Compactly supportedness.*

$$M_{supp} = \frac{\max_{|u|,|v|\leq 3} |\hat{I}(u, v)|}{\max_{u,v} |\hat{I}(u, v)|}.$$

  - ◇ *Decay rate.*

$$M_{decay_2} = \frac{1}{1024} \sum_{i=1,\ldots,512} d_i.$$

- *Smoothness in spatial domain.* Smoothness will be measured by the average of local Hölder regularity. For this, for each $(u_0, v_0) \in \{1, \ldots, 512\}^2$, compute

$M(u,v) = |I(u,v) - I(u_0, v_0)|,\ 0 < \max\{|u - u_0|, |v - v_0|\} \le 4$. Then the local Hölder regularity α_{u_0,v_0} is the least square fit to the curve $(u,v) \mapsto \log(|M(u,v)|)$. The measure is given by

$$M_{smooth_1} = \frac{1}{512^2} \sum_{u,v} \alpha_{u,v}.$$

- *Smoothness in frequency domain.* Smoothness will again be measured by the average of local Hölder regularity. Proceed as for measuring the smoothness in spatial domain now applied to $\hat{I}$, the 2D-FFT of I, to derive the local Hölder regularity $\alpha_{u,v}$ for each $(u,v) \in \{1, \ldots, 512\}^2$. The measure is then given by

$$M_{smooth_2} = \frac{1}{512^2} \sum_{u,v} \alpha_{u,v}.$$

[D5] TRUE SHEAR INVARIANCE.

Comments: This test shall provide a measure for how close the transform is to being shear invariant. In our case, the theory gives

$$\left\langle 2^{3j/2} \psi(S_k^{-1} A_{4^j} \cdot -m), f(S_s \cdot) \right\rangle = \left\langle 2^{3j/2} \psi(S_{k+2^j s}^{-1} A_{4^j} \cdot -m), f \right\rangle,$$

and we expect to see this behavior in the shearlet coefficient as discussed in Subsection 4.2.

Measure: Let I be an 256×256 image with an edge through the origin $(129, 129)$ of slope 0. Fix $s = 1/2$, generate an image $I_s := I(S_s \cdot)$, and let $j \in \{1, 2, 3, 4\}$ be a scale. Note that $2^j s \in \mathbb{Z}$ for each of $j \in \{1, 2, 3, 4\}$. Our quality measure will then be the curve

$$M_{shear,j} = \max_{-2^j < k, k+2^j s < 2^j} \frac{\|C_{j,k}(SI_s) - C_{j,k+2^j s}(SI)\|_2}{\|I\|_2}, \qquad \text{scale } j = 1, 2, 3, 4,$$

where $C_{j,k}$ is the set of coefficients – in our case the shearlet coefficient – at scale j and at shear index k.

[D6] SPEED.

Comments: When testing the speed, not only the asymptotic behavior will be measured but also the involved constants. It should further be mentioned that for practical purposes it is usually sufficient to analyze the speed up to a size of $N = 512$.

Measure: Generate a sequence of 5 random images I_i, $i = 5, \ldots, 9$ of size $2^i \times 2^i$ with standard normally distributed entries. Let s_i be the speed of the transform S applied to I_i. Our hypothesis is that the speed behaves like $s_i = c \cdot (2^{2i})^d$; 2^{2i} being the size of the input. Let now $\tilde{d}_a$ be the average slope of the line, which is a least square fit to the curve $i \mapsto \log(s_i)$. Let also f_i be the 2D-FFT applied to I_i, $i = 5, \ldots, 9$. Our quality measure will then be three-fold:

- *Complexity.*

$$M_{speed_1} = \frac{\tilde{d}_a}{2 \log 2}.$$

- *The constant.*

$$M_{speed_2} = \frac{1}{5} \sum_{i=5}^{9} \frac{s_i}{\left(2^{2i}\right)^{M_{speed_1}}}.$$

- *Comparison with 2D-FFT.*

$$M_{speed_3} = \frac{1}{5} \sum_{i=5}^{9} \frac{s_i}{f_i}.$$

[D7] GEOMETRIC EXACTNESS.

Comments: Geometric objects such as edges should be as precise as possible be resembled by the coefficients in the sense that analyzing the decay of the coefficient should detect such features. For the FDST, these properties were theoretically analyzed in Subsection 4.1. Our model will be an image containing one line of a particular slope.

Measure: Let $I_1, \ldots, I_8$ be 256×256 images of an edge through the origin $(129, 129)$ and of slope $[-1, -0.5, 0, 0.5, 1]$ and the transpose of the middle three, and let $c_{i,j}$ be the associated shearlet coefficients for image I_i at scale j. Our quality measure will be two-fold:

- *Decay of significant coefficients.* Consider the curve

$$\frac{1}{8} \sum_{i=1}^{8} \max |c_{i,j}(\text{of shearlets aligned with the line})|, \qquad \text{scale } j,$$

  let d be the average slope of the line, which is a least square fit to the logarithm of this curve, and define

$$M_{geo_1} = d.$$

- *Decay of insignificant coefficients.* Consider the curve

$$\frac{1}{8} \sum_{i=1}^{8} \max |c_{i,j}(\text{of all other shearlets})|, \qquad \text{scale } j,$$

  let d be the average slope of the line, which is a least square fit to the logarithm of this curve, and define

$$M_{geo_2} = d.$$

[D8] ROBUSTNESS.

Comments: Two different types of robustness will be analyzed which we believe are the most common impacts on a sequence of transform coefficients. We wish to mention that we certainly also could have considered additional manipulations of the coefficients such as deletions; but to provide sufficiently many tests balanced with a reasonable testing time, we decided to restrict to those two.

Measure:
- *Thresholding.* Let I be the regular sampling of a Gaussian function with mean 0 and variance 256 on $\{-128, ..., 127\}^2$ generating an 256×256-image. The quality measure for $k = 1, 2$ will be the curve

$$M_{thres_{k,p_k}} = \frac{\|G_{\sqrt{w}P} W^\star \, \text{thres}_{k,p_k} \, SI - I\|_2}{\|I\|_2},$$

  where

– thres$_{1,p_1}$ discards $100 \cdot (1 - 2^{-p_1})$ percent of the coefficients with $p_1 = [2 : 2 : 10]$),
– thres$_{2,p_2}$ sets all those coefficients to zero with absolute values below the threshold $m/2^{p_2}$ with m being the maximal absolute value of all coefficients with $p_2 = [0.001 : 0.01 : 0.05]$).

• *Quantization.* Let I be the regular sampling of a Gaussian function with mean 0 and variance 256 on $\{-128, ..., 127\}^2$ generating an 256×256-image. The quality measure will be the curve

$$M_{quant,q} = \frac{\|G_{\sqrt{w}P} W^\star \text{ quant}_q SI - I\|_2}{\|I\|_2}, \qquad q = [8 : -0.5 : 6],$$

where $\text{quant}_q(c) = \text{round}(c/(m/2^q)) \cdot (m/2^q)$ and m being the maximal absolute value of all coefficients.

7. Numerical Evidence. In this section, we provide numerical results for the tests [D1]–[D8] detailed in Section 6 of our present implementation, which we consider as having reached a mature state after careful tuning the parameters depending on these performance measures. The associated code `ShearLab-PPFT-1.0` can be downloaded from `www.ShearLab.org`.

7.1. Results for Tests [D1]–[D3]. Table 7.1 presents the performance with respect to the quantitative measures in [D1]–[D3].

TABLE 7.1
Results for [D1]–[D3]

M_{alg}	M_{isom_1}	M_{isom_2}	M_{isom_3}	M_{tight_1}	M_{tight_2}
6.6E-16	9.3E-4	1.834	3.3E-7	9.9E-4	3.8E-7

The quantity $M_{alg} \approx$ 6.6E-16 confirms that the $\mathcal{DSH}$ defined in Definition 2.2 is indeed up to machine precision a tight frame.

The slight tightness deficiency of $M_{tight_1} \approx$ 9.9E-4 (also $M_{isom_1} \approx$ 9.3E-4) mainly results from the isometry deficiency of the weighting. However, for practical purposes this transform can be still considered to be an isometry allowing the utilization of the adjoint as inverse transform. Progress on the choice of weights will further improve this measure. Observe though that there is a trade-off between the sophistication of the weights, the running time of S, and the smoothness of the shearlets.

Further, note that the condition number ($M_{isom_2} \approx 1.834$) of the Gram matrix is quite close to 1, which – in case an even higher accurate inverse than the adjoint is required – allows us to employ the conjugate gradient method very efficiently for computing the inverse of the FDST ($M_{isom_3} \approx$ 3.3E-7 and $M_{tight_2} \approx$ 3.8E-7).

7.2. Space-Frequency-Localization Test [D4]. The reference shearlet I required for [D4] is illustrated in Figure 7.1 in both the spatial and frequency domain.

Table 7.2 presents the space-frequency-localization measures M_{decay_1}, M_{decay_2}, and M_{supp} for analyzing the decay in spatial and frequency domains as well as the measures M_{smooth_1} and M_{smooth_2} for smoothness in spatial and frequency domains.

The measurements indicate that the shearlet illustrated in Figure 7.1 decays slower in spatial domain ($M_{decay_1} \approx$ -1.920) than in the frequency domain ($M_{decay_2} \approx$ -3.257) in terms of the average decay rates. In terms of average local Hölder smoothness, the shearlet in spatial domain is smoother than in the frequency domain proven
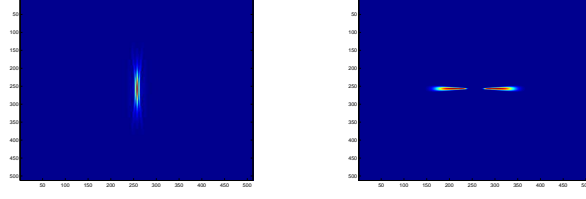
FIG. 7.1. *Graphs of the reference shearlet for test [D4] centered at origin at scale 3 and shear parameter 0 in both spatial and frequency domain with $N = 512$.*

TABLE 7.2
Results for [D4]

M_{decay_1}	M_{supp}	M_{decay_2}	M_{smooth_1}	M_{smooth_2}
-1.920	5.5E-5	-3.257	1.319	0.734

by the fact that $M_{smooth_1} \approx 1.319 > 0.734 \approx M_{smooth_2}$. The very small value $M_{supp} \approx$ 5.5E-5 – although it is not zero due to round off errors – indicates that our shearlet is indeed compactly supported.

7.3. Shear Invariance Test [D5]. Test [D5] requires a reference image I and a sheared version $I_s = I(S_s \cdot)$, which is illustrated in Figure 7.2 for $s = 0.5$ alongside with their pseudo-polar Fourier transforms $\hat{I}$ and $\hat{I}_s$, respectively.
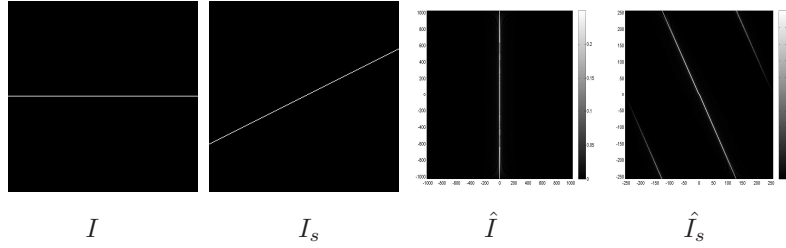

I $\qquad\qquad$ I_s $\qquad\qquad$ $\hat{I}$ $\qquad\qquad$ $\hat{I}_s$

FIG. 7.2. *Graphs of the reference image I and its sheared version I_s for $s = 0.5$ in spatial domain and $\hat{I}$ and $\hat{I}_s$ in pseudo-polar domain.*

The shear invariance measurements are presented in Table 7.3. This table shows that the FDST is indeed almost shear invariant. A closer inspection shows that $M_{shear,1}$ and $M_{shear,2}$ are relatively small compared to the measurements with respect to finer scales $M_{shear,3}$ and $M_{shear,4}$. The reason for this is the aliasing effect which shifts some energy to the high frequency part near the boundary away from the edge in the frequency domain, see also the graph of $\hat{I}_s$ in Figure 7.2.

7.4. Speed Test [D6]. The measurements for speed performance, i.e., for the complexity M_{speed_1}, the constant M_{speed_2}, and the comparison with 2D-FFT M_{speed_3}, are presented in Table 7.4.

The complexity constant $M_{speed_1} \approx 1.156$ confirms that the order of complexity of our FDST algorithm is indeed close to linear (compare $O(N^2 \log N)$ in theory). The constant of comparison with 2D-FFT is $M_{speed_3} \approx 280$, which seems significantly slower than the usually 2D-FFT. However, we should notice that the 2D-FFT is applied directly to the image of size $N \times N$ while the FDST employs fractional

<div align="center">

TABLE 7.3
Results for [D5]

</div>

Scale	1	2	3	4
$M_{shear,j}$	1.6E-5	1.8E-4	0.002	0.003

<div align="center">

TABLE 7.4
Results for [D6]-[D7]

</div>

M_{speed_1}	M_{speed_2}	M_{speed_3}	M_{geo_1}	M_{geo_2}
1.156	9.3E-6	280.560	-1.358	-2.032

Fourier transforms and subband windowing on a oversampling pseudo-polar grid of size $2 \times (RN + 1) \times (N + 1)$ with oversampling rate $R = 8$. Taking into account that the fractional Fourier transform is about 5 times slower than the FFT and that the redundancy (about $4R$) comes from subband windowing with oversampling rate $R = 8$, we conclude that the constant M_{speed_3} is reasonable and our implementation is indeed comparable with the 2D-FFT.

7.5. Geometric Exactness Test [D7]. Test [D7] requires reference images with edges at various slopes which are plotted in Figure 7.3.
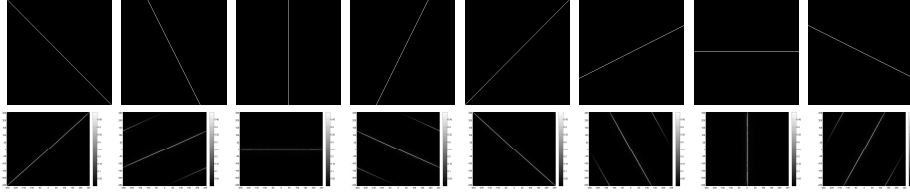

<div align="center">

FIG. 7.3. *The graphs of edges in both spatial and frequency domains. Top 8: Edges in spatial domain. Bottom 8: Edges in frequency domain.*

</div>

The associated geometric exactness measurements M_{geo_1} and M_{geo_2} are presented in Table 7.4. These two measurements show the differences between the decay rates of the shearlet coefficients aligned with edges (significant shearlet coefficients) and the shearlet coefficients not aligned with the edges (insignificant shearlet coefficients). As theoretically proven in Subsection 4.1, and hence expected for the implementation, the insignificant shearlet coefficients ($M_{geo_2} \approx$ -2.032) decay much faster than the significant shearlet coefficients with increasing scale j.

7.6. Robustness Test [D8]. Table 7.6 presents the measurements for the robustness test [D8].

These results show that even if $100(1 - 2^{-10}) \approx 99.9\%$ of the shearlet coefficients are discarded, the original image is still well approximated by the reconstructed image ($M_{thres_{1,p_1}} \approx 0.007$). Thus the number of the significant coefficients is relatively small compared to the total number of shearlet coefficients. The second row indicates that knowledge of the shearlet coefficients with absolute value greater than $m(1-1/2^{0.001})$ – hence about 0.1% of coefficients – is sufficient for precise reconstruction ($M_{thres_{1,p_2}} \approx$ 0.005). The quantization test $M_{quant,q}$ attests the FDST high resilience against even quite coarse quantization.

TABLE 7.5
Results for [D8]

$M_{thres_{1,p_1}}$	1.5E-8	7.2E-8	2.5E-5	0.001	0.007
$M_{thres_{2,p_2}}$	0.005	0.039	0.078	0.113	0.154
$M_{quant,q}$	0.034	0.047	0.057	0.071	0.109

8. Conclusions. We described a natural digitization of the continuous shearlet transform based on the pseudo-polar Fourier transform. For this, we first introduced a digital shearlet tight frame, which is an exact digitization of the classical band-limited system defined in the continuum domain. This shows that the shearlet framework provides a unified treatment of the continuum and digital realm. We then described a fast digital shearlet transform (FDST) whose main ingredients are a weighted pseudo-polar Fourier transform to achieve an isometric mapping into the pseudo-polar domain and a windowing based on the digital shearlet tight frame. We discussed several choices of weighting functions to achieve almost isometry so that inverse digital shearlet transform can be obtained by using the adjoint transform. Various properties of the FDST are proven, and details of its implementation are provided. We further define several measures that quantify the performance of the FDST, thereby justifying the term 'rational design'. These measures also provide a common ground to compare the performance of different algorithms based on parabolic scaling. The FDST as well as the performance measures are publicly available in the software package `ShearLab`, see `www.ShearLab.org`.

Appendix A. Algorithms.

In this appendix, we provide pseudo-codes in Matlab style for algorithms of FDST (Algorithm A.1), of the adjoint FDST (Algorithm A.2), and the inverse FDST (Algorithm A.3). For the required notation, we refer to Section 5.

REFERENCES

[1] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Israeli, and Y. Shkolnisky, *A framework for discrete integral transformations I – the pseudo-polar Fourier transform*, SIAM J. Sci. Comput. **30** (2008), 764–784.
[2] D. H. Bailey and P. N. Swarztrauber, *The fractional Fourier transform and applications*, SIAM Rev. **33** (1991), 389–404.
[3] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms*, Multiscale Model. Simul. **5** (2006), 861–899.
[4] E. J. Candès and D. L. Donoho, *Ridgelets: a key to higher-dimensional intermittency?*, Phil. Trans. R. Soc. Lond. A. **357** (1999), 2495–2509.
[5] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.
[6] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197.
[7] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: II. Discretization of frames*, Appl. Comput. Harmon. Anal. **19** (2005), 198–222.
[8] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.
[9] M. N. Do and M. Vetterli, *The contourlet transform: An efficient directional multiresolution image representation*, IEEE Trans. Image Process. **14** (2005), 2091–2106.
[10] D. L. Donoho and G. Kutyniok, *Microlocal analysis of the geometric separation problem*, preprint.
[11] D. L. Donoho, A. Maleki, M. Shahram, V. Stodden, and I. Ur-Rahman, *Fifteen years of*

Algorithm A.1 Fast Digital Shearlet Transform (FDST)

(a) **Input:** Image $I = \{[I]_{u,v} : -N/2 \le u, v \le N/2 - 1\}$ of size $N \times N$, the oversampling rate R, and the precomputed weight matrix w of size $2 \times (RN + 1) \times (N + 1)$.

(b) **Output:** Digital shearlet coefficients $C_I = \{c^{\iota_0} : \iota_0\} \cup \{c^\iota_{j,s} : j, s, \iota\}$. Each $c^\iota_{j,s}$ is a matrix of size $\mathcal{L}^1_j \times \mathcal{L}^2_{j,s}$, see Definition 2.2.

(c) **PPFT:** The pseudo-polar Fourier transform (see also Algorithm 1).

1: $J \leftarrow zeros(2, RN + 1, N + 1)$. $m \leftarrow RN + 1$. `//J is the grid` $J = J_{\Omega^1_R} \cup J_{\Omega^2_R}$.
2: **for** $sector = 1$ to 2 **do**
3: 　 **if** $sector = 2$ **then** $I = I^T$ **end if**
4: 　 **for** $v = -N/2$ to $N/2 - 1$ **do**
5: 　　 $q \leftarrow [I]_{:,v}$. , $q \leftarrow E_{RN+1,N}q$, $[J]_{sector,:,v} \leftarrow F_1 q$.
　　　 `//symmetric zero-padding and FFT along direction` u.
6: 　 **end for**
7: 　 **for** $k = -RN/2$ to $RN/2$ **do**
8: 　　 $q \leftarrow [J]_{sector,k,:}$, $\alpha \leftarrow -\frac{k}{mN/2}$, $q \leftarrow F^\alpha_{N+1}q$, $[J]_{sector,k,:} \leftarrow q$.
　　　 `//frFT along direction` v.
9: 　 **end for**
10: **end for**

(d) **Weighting:** Weighting on the pseudo-polar grid with precomputed weight w.
11: $J \leftarrow J.*\sqrt{w}$. `//.* is the point-wise multiplication`.

(e) **Subband Windowing:** Subband windowing with $\mathcal{DSH}$.
12: $L \leftarrow -\lceil \log_4(R/4) \rceil$, $H \leftarrow \lceil \log_4 N \rceil$, $sMax \leftarrow 2 \cdot 2^H + 1$.
13: $C_I \leftarrow cell\{4, H - L + 2, sMax\}$.
14: **for** $\iota = 11, 12, 21, 22$ **do**
15: 　 **for** $j = L$ to H **do**
16: 　　 **if** $j < 0$ **then** $tile \leftarrow 0$ **else** $tile \leftarrow 2^j$ **end if**
17: 　　 **for** $s = -tile$ to $tile$ **do**
18: 　　　 Let $c^\iota_{j,s}$ be a submatrix of J of size $\mathcal{L}^1_j \times \mathcal{L}^2_{j,s}$ with respect to the support of the digital shearlet $\sigma^\iota_{j,s,0}$.
19: 　　　 $c^\iota_{j,s} \leftarrow c^\iota_{j,s}.*\overline{\sigma^\iota_{j,s,0}}$, $c^\iota_{j,s} \leftarrow F^{-1}_2 c^\iota_{j,s}$. `//Windowing with 2D iFFT`.
20: 　　　 $C_I\{\iota, j, s\} \leftarrow c^\iota_{j,s}$.
21: 　　 **end for**
22: 　 **end for**
23: **end for**
24: Let φ^1_0, φ^2_0 be the shearlets associated with the low-frequency part. Let $c^i, i = 1, 2$ denote the submatrix of J with respect to the support of φ^i_0, $i = 1, 2$.
25: $c^i \leftarrow c^i.*\overline{\varphi^i_0}$, $i = 1, 2$.
26: $C_I\{i, L - 1, 0\} \leftarrow c^i$, $i = 1, 2$.

---

reproducible research in computational harmonic analysis, Comput. Sci. Eng. **11** (2009), 8–18.

[12] 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.

[13] H. Feichtinger, K. Gröchenig, and T. Strohmer, *Efficient numerical methods in non-uniform sampling theory,* Numer. Math. **69** (1995), 423–440.

[14] K. Guo, G. Kutyniok, and D. Labate, *Sparse multidimensional representations using anisotropic dilation and shear operators,* Wavelets and Splines (Athens, GA, 2005), Nashboro

Algorithm A.2 Fast Adjoint Digital Shearlet Transform (Adjoint FDST)

(a) **Input:** Digital shearlet coefficients $C_I = \{c^{\iota 0} : \iota\} \cup \{c^{\iota}_{j,s} : j, s, \iota\}$, where each $c^{\iota}_{j,s}$ is a matrix of size $\mathcal{L}^1_j \times \mathcal{L}^2_{j,s}$ (see Definition 2.2), the oversampling rate R, and the precomputed weight matrix w of size $2 \times (RN+1) \times (N+1)$.

(b) **Output:** An image $I = \{[I]_{u,v} : -N/2 \leq u, v \leq N/2 - 1\}$ of size $N \times N$.

(c) **Adjoint Subband Windowing:** Subband windowing by $\mathcal{DSH}$.
1: $L \leftarrow -\lceil \log_4(R/4) \rceil$, $H \leftarrow \lceil \log_4 N \rceil$, $J \leftarrow zeros(2, RN+1, N+1)$.
2: $c^i \leftarrow C_I\{i, L-1, 0\}$, $i = 1, 2$.
3: $c^i \leftarrow c^i .* \varphi^i_0$, $i = 1, 2$.
4: Assign c^i to J with respect to the support of φ^i_0, $i = 1, 2$.
5: **for** $\iota = 11, 12, 21, 22$ **do**
6:    **for** $j = L$ to H **do**
7:      **if** $j < 0$ **then** $tile \leftarrow 0$ **else** $tile \leftarrow 2^j$ **end if**
8:      **for** $s = -tile$ to $tile$ **do**
9:        $c^{\iota}_{j,s} \leftarrow C_I\{\iota, j, s\}$.
10:        $c^{\iota}_{j,s} \leftarrow F_2 c^{\iota}_{j,s}$. //the 2D-FFT.
11:        $c^{\iota}_{j,s} \leftarrow c^{\iota}_{j,s} .* \sigma^{\iota}_{j,s,0}$.
12:        Assign $c_{j,s}$ to J with respect to the support of $\sigma^{\iota}_{j,s,0}$.
13:      **end for**
14:    **end for**
15: **end for**

(d) **Weighting:** Weighting with the precomputed weight w.
16: $J \leftarrow J .* \sqrt{w}$. //.* is the point-wise multiplication.

(e) **Adjoint PPFT:** The adjoint pseudo-polar Fourier transform.
17: $I \leftarrow zeros(N, N)$, $I_0 \leftarrow zeros(N, N)$, $m \leftarrow RN+1$, $J_1 \leftarrow zeros(m, N)$.
18: **for** $sector = 1$ to 2 **do**
19:    **for** $k = -RN/2$ to $RN/2$ **do**
20:      $q \leftarrow [J]_{sector,k,:}$. $\alpha \leftarrow -\frac{k}{mN/2}$. $q \leftarrow F^{-\alpha}_{N+1} q$.
21:      $[J_1]_{k,:} \leftarrow E^{\star}_{N+1,N} q$.
22:    **end for**
23:    **for** $v = -N/2$ to $N/2 - 1$ **do**
24:      $q \leftarrow [J_1]_{:,v}$. $q \leftarrow F_1 q$, $q \leftarrow E^{\star}_{m,N} q$.
25:      $[I_0]_{:,v} \leftarrow q$.
26:    **end for**
27:    $I = I + I_0$.
28: **end for**

Press, Nashville, TN (2006), 189–201.

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

[16] K. Gröchenig, *Reconstruction algorithms in irregular sampling*, Math. Comp. **59**(199) (1992), 181–194.

[17] K. Gröchenig, *A discrete theory of irregular sampling*, Linear Algebra Appl. **193** (1993), 129–150.

[18] K. Gröchenig and T. Strohmer, *Numerical and theoretical aspects of non-uniform sampling of band-limited images*, in F. Marvasti, editor, Nonuniform Sampling: Theory and Applications, chapter 6, page 283–324. Kluwer, 2001.

[19] B. Han, G. Kutyniok, and Z. Shen, *A unitary extension principle for Shearlet Systems*, preprint.

Algorithm A.3 Inverse Fast Digital Shearlet Transform (Inverse FDST)

(a) **Input:** Digital shearlet coefficients $C_I = \{c^{\iota_0} : \iota_0\} \cup \{c_{j,s}^{\iota} : j, s, \iota\}$, the oversampling rate R, the precomputed weight matrix w of size $2 \times (RN+1) \times (N+1)$, an initial guess I_0, a precision parameter ε, and a maximal iteration number $itMax$.

(b) **Output:** An image $I = \{[I]_{u,v} : -N/2 \le u, v \le N/2 - 1\}$ of size $N \times N$.

(c) **Initialization:** Perform the adjoint shearlet windowing on C_I by using Algorithm 2 to obtain an image b, say.

(d) **CG Iteration:** Let $A = P^\star w P$ generated by Algorithm 1, Algorithm A.1(e), and Algorithm A.2(e). We apply the CG method to solve the linear system $AI = b$.

1: $r_0 \leftarrow b - AI_0$, $p_0 \leftarrow r_0$, $k \leftarrow 0$.
2: **while** $\|r_k\|_2 > \varepsilon$ and $k < itMax$ **do**
3:     $\alpha_k \leftarrow \frac{r_k^T r_k}{p_k^T A p_k}$.
4:     $I_{k+1} \leftarrow I_k + \alpha_k p_k$.
5:     $r_{k+1} \leftarrow r_k - \alpha_k A p_k$.
6:     $\beta_k \leftarrow \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$.
7:     $p_{k+1} \leftarrow r_{k+1} + \beta_k p_k$.
8:     $k \leftarrow k + 1$.
9: **end while**
10: $I \leftarrow I_k$.

[20] F. J. Herrmann and G. Hennenfent, *Non-parametric seismic data recovery with curvelet frames*, Geophys. J. Int. **173** (2008), 233-248.
[21] E. Hewitt and K.A. Ross, *Abstract Harmonic Analysis I, II*, Springer-Verlag, Berlin/ Heidelberg/New York, 1963.
[22] P. Kittipoom, G. Kutyniok, and W.-Q Lim, *Construction of compactly supported shearlet frames*, preprint.
[23] G. Kutyniok, *Sparsity equivalence of anisotropic decompositions*, preprint.
[24] G. Kutyniok and D. Labate, *Resolution of the wavefront Set using continuous shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.
[25] G. Kutyniok and W.-Q Lim, *Compactly supported shearlets are optimally sparse*, preprint.
[26] G. Kutyniok, J. Lemvig, and W.-Q Lim, *Compactly supported shearlets*, Approximation Theory XIII (San Antonio, TX, 2010), Springer, to appear.
[27] G. Kutyniok and T. Sauer, *Adaptive directional subdivision schemes and shearlet multiresolution analysis*, SIAM J. Math. Anal. **41** (2009), 1436–1471.
[28] W.-Q Lim, *The discrete shearlet transform: A new directional transform and compactly supported shearlet frames*, IEEE Trans. Imag. Proc. **19** (2010), 1166–1180.
[29] J.-L. Starck, E. Candès, and D. L. Donoho, *Astronomical image representation by the curvelet tansform*, Astronomy and Astrophysics **398** (2003), 785–800.