Appl. Comput. Harmon. Anal. 39 (2015) 300–338

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Smooth affine shear tight frames with MRA structure

Bin Han [a,1], Xiaosheng Zhuang [b,*,2]

[a] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada*

[b] *Department of Mathematics, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong*

ARTICLE INFO

Article history:
Received 19 March 2014
Received in revised form 25 August 2014
Accepted 20 September 2014
Available online 30 September 2014
Communicated by Christopher Heil

MSC:
42C40
41A05
42C15
65T60

Keywords:
Affine shear tight frames
Filter banks
Directional tight framelets
Directional multiscale representation systems
Affine systems
Cone-adapted
Smooth shearlets
Image denoising

ABSTRACT

Finding efficient directional representations is one of the most challenging and extensively sought problems in mathematics. Representation using shearlets recently receives a lot of attention due to their desirable properties in both theory and applications. Using the framework of frequency-based affine systems as developed in [16], in this paper we introduce and systematically study affine shear tight frames which include all known shearlet tight frames as special cases. Our results in this paper resolve several important questions on shearlets. We provide a complete characterization for an affine shear tight frame and then use it to construct smooth directional affine shear tight frames with all their generators in the Schwartz class. Though multiresolution analysis (MRA) together with filter banks is the foundation and key features of wavelet analysis for the fast numerical implementation of a wavelet transform, most papers on shearlets do not concern the underlying filter bank structure and its connection to MRA. In order to study affine shear tight frames with MRA structure, following the lines developed in [16], we introduce the notion of a sequence of affine shear tight frames and then we provide a complete characterization for such a sequence. Based on our characterizations, we present two different approaches, i.e., non-stationary and quasi-stationary, for the construction of sequences of directional affine shear tight frames with MRA structure such that all their generators are smooth (in the Schwartz class) and they have underlying filter banks. Consequently, their associated transforms can be efficiently implemented using filter banks and are very similar to the standard fast wavelet transform. Moreover, we provide concrete examples of directional affine shear tight frames with filter banks and apply them to the image denoising problem. Our numerical experiments show that our constructed directional affine shear tight frames perform better than known directional multiscale representation systems such as curvelets and shearlets for the image denoising problem.

* Corresponding author.
E-mail addresses: bhan@ualberta.ca (B. Han), xzhuang7@cityu.edu.hk (X. Zhuang).

[1] Research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC Canada) under Grant 2014-05865.

[2] Research was supported by Research Grants Council of Hong Kong (Project No. CityU 108913).

http://dx.doi.org/10.1016/j.acha.2014.09.005

1. Introduction and motivation

In the era of information, everyday and everywhere, huge amount of information is acquired, processed, stored, and transmitted in the form of high-dimensional digital data through Internet, TVs, cell phones, satellites, and various other modern communication technologies. One of the main goals in today's scientific research is the efficient representation and extraction of information in high-dimensional data. It is well known that high-dimensional data usually exhibit anisotropic phenomena due to data clustering of various types of structures. For example, cosmological data normally consist of many morphological distinct objects concentrated near lower-dimensional structures such as points (stars), filaments, and sheets (nebulae). The anisotropic features of high-dimensional data thus encode a large portion of significant information. Mathematical representation systems that are capable of capturing such anisotropic features are therefore undoubtedly the key for the efficient representation of high-dimensional data.

During the past decade, directional multiscale representation systems have become more and more popular due to their abilities of resolving anisotropic features in high-dimensional data, see [1,2,12,16,19,21,28,32] and many references therein. Our focus in this paper is on investigation and construction of a general type of directional multiscale representation systems: *affine shear tight frames*. Such a type of directional multiscale representation systems has many desirable properties including directionality, multiresolution analysis (MRA), smooth generators, etc. Moreover, the affine shear systems have an underlying filter banks associated with the directional affine (wavelet) systems as considered in [16].

Before proceeding, let us first introduce necessary notation and definitions. Let U be a $d \times d$ real-valued invertible matrix. Throughout the paper we shall use the following notation:

$$f_{U;\mathsf{k},\mathsf{n}}(x) = f_{[\![U;\mathsf{k},\mathsf{n}]\!]}(x) = [\![U;\mathsf{k},\mathsf{n}]\!]f(x) := |\det U|^{1/2} f(Ux-\mathsf{k})e^{-i\mathsf{n}\cdot Ux}, \quad \mathsf{k},\mathsf{n},x \in \mathbb{R}^d.$$

Here U, k, and n refer to dilation, translation, and modulation, respectively. We shall adopt the convention that $f_{U;\mathsf{k}} := f_{U;\mathsf{k},0}$ and $f_{\mathsf{k},\mathsf{n}} := f_{\mathsf{I}_d;\mathsf{k},\mathsf{n}}$ with I_d being the $d \times d$ identity matrix. Note that such a notation $f_{U;\mathsf{k},\mathsf{n}}$ is consistent with the usual notation $\psi_{j,\mathsf{k}}$ for wavelets in 1D, since $\psi_{2^j;\mathsf{k}} = 2^{j/2}\psi(2^j\cdot-\mathsf{k})$.

Though all the discussion and results in this paper can be carried over to any dimensions $\mathbb{R}^d$ with $d \ge 2$, for simplicity of presentation, we restrict ourselves to the two-dimensional case only, which is the most important case in the area of directional multiscale representations. We shall use the following matrices throughout this paper:

$$\mathsf{E} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad S^\tau := \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \qquad S_\tau := \begin{bmatrix} 1 & 0 \\ \tau & 1 \end{bmatrix}, \qquad \mathsf{A}_\lambda := \begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda \end{bmatrix},$$
$$\mathsf{B}_\lambda := (\mathsf{A}_\lambda)^{-\mathsf{T}} = \begin{bmatrix} \lambda^{-2} & 0 \\ 0 & \lambda^{-1} \end{bmatrix}, \tag{1.1}$$

where $\tau \in \mathbb{R}$ and $\lambda > 1$. S^τ and S_τ are the shear operations while A_λ is the dilation matrix. Define $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and define $\delta : \mathbb{Z}^d \to \mathbb{R}$ to be the Kronecker/Dirac sequence such that $\delta(0) = 1$ and $\delta(\mathsf{k}) = 0$ for all $\mathsf{k} \in \mathbb{Z}^d \backslash \{0\}$.

An affine shear system is obtained by applying shear, dilation, and translation to generators at different scales. Note that $f(S^\tau\cdot)$ could be highly tilted when τ is very large for a compactly supported function f. To balance the shear operation, one usually considers cone-adapted systems [8,10,13,22]. A cone-adapted system usually consists of three subsystems: one subsystem covers the low frequency region, one subsystem covers the horizontal cone $\{\xi = (\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2/\xi_1| \le 1\}$, and one subsystem covers the vertical cone $\{\xi = (\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1/\xi_2| \le 1\}$ in the frequency plane. Throughout the paper, ξ is used as a one- or two-dimensional variable for the frequency domain with $\xi = (\xi_1, \xi_2)$ if $\xi \in \mathbb{R}^2$. The vertical-cone subsystem could be constructed to be the 'flipped' version of the horizontal-cone subsystem. More precisely, a function

$\varphi \in L_2(\mathbb{R}^2)$ serves as the generator for the low frequency region, a function $\psi \in L_2(\mathbb{R}^2)$ generates an affine system covering certain region of the horizontal cone in the frequency domain, and $\{\psi^{j,\ell} \in L_2(\mathbb{R}^2) : |\ell| = r_j + 1, \dots, s_j\}$ is a set of generators at the scale level j that generates elements along the seamlines (i.e., diagonal directions $\{\xi \in \mathbb{R}^2 : \xi_2/\xi_1 = \pm 1\}$) to serve the purpose of tightness of the system. Note that $\psi^{j,\ell}, |\ell| = r_j + 1, \dots, s_j$ may not come from a single generator. Define Ψ_j to be the set of generators in $L_2(\mathbb{R}^2)$ as

$$\Psi_j := \{\psi(S^{-\ell}\cdot) : \ell = -r_j, \dots, r_j\} \cup \{\psi^{j,\ell}(S^{-\ell}\cdot) : |\ell| = r_j + 1, \dots, s_j\}, \tag{1.2}$$

where r_j and s_j are nonnegative integers. An *affine shear system* (with the initial scale $J = 0$) is then defined to be

$$\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^\infty) = \{\varphi(\cdot - \mathsf{k}) : \mathsf{k} \in \mathbb{Z}^2\} \cup \{h_{\mathsf{A}_\lambda^j;\mathsf{k}}, h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}} \,:\, \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\}_{j=0}^\infty. \tag{1.3}$$

For a function f defined on $\mathbb{R}^2$, observe that $f_{\mathsf{E};0}(x,y) = f(y,x)$; that is, $f_{\mathsf{E};0}$ is the 'flipped' version of f along the line $y = x$. Note that the system $\{h_{\mathsf{A}_\lambda^j;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\}$ is for the high frequency region at the scale level j with respect to the horizontal cone, while its 'flipped' version $\{h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \mathbf{\Psi}_j\}$ is for the high frequency region at the scale level j with respect to the vertical cone in the frequency plane.

1.1. Related work

In 1D, it is well known that wavelet representation systems provide optimally sparse representations for functions $f \in L_2(\mathbb{R})$ that are smooth except for finitely many discontinuity 'jumps' [5]. In high dimensions, wavelet representation systems could be obtained by using tensor product of 1D wavelets. However, tensor product real-valued wavelets usually lack directionality since they only favor certain directions such as the horizontal and vertical directions. Though directionality of tensor product real-valued wavelets can be improved by using complex wavelets [32] or complex tight framelets [17,19], the limitation of directionality selectivity is intrinsic in any tensor product approach and therefore, tensor product wavelets or framelets fail to provide optimally sparse approximation for 2D piecewise smooth functions with singularities along a closed smooth curve (anisotropic features). To achieve flexible directionality selectivity, additional operation other than dilation and translation is needed.

For a function $f \in L_1(\mathbb{R}^d)$, the Fourier transform $\widehat{f}$ of f in this paper is defined to be

$$\widehat{f}(\xi) = \mathscr{F}f(\xi) := \int_{\mathbb{R}^d} f(x)e^{-ix\cdot\xi}dx, \quad \xi \in \mathbb{R}^d,$$

which can be extended to square-integrable functions in $L_2(\mathbb{R}^d)$ and tempered distributions through duality. Note that the Plancherel identity holds in $L_2(\mathbb{R}^d)$: $\langle f,g\rangle = \frac{1}{(2\pi)^d}\langle \widehat{f}, \widehat{g}\rangle$ for $f,g \in L_2(\mathbb{R}^d)$, where $\langle f,g\rangle := \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx$. We also define $\|f\|_2^2 := \langle f,f\rangle$. Note that $\widehat{f_{U;\mathsf{k}}} = \widehat{f}_{U^{-\mathsf{T}};0,\mathsf{k}}$.

Directional tight framelets in [14,16], directly built from the frequency plane, achieve directionality by separating the frequency plane into annulus at different scales and further splitting each annulus into different wedge shapes. More precisely, in the frequency domain, considering the polar coordinate (r,θ) (i.e., $(x,y) = (r\cos\theta, r\sin\theta)$), one first constructs a pair $\{\boldsymbol{\eta}(r), \boldsymbol{\zeta}(r)\}$ of 1D scaling and wavelet functions in the frequency domain such that $|\boldsymbol{\eta}|^2 + \sum_{j\in\mathbb{N}_0} |\boldsymbol{\zeta}(2^{-j}\cdot)|^2 = 1$. Then, a 2D scaling function φ is defined by $\widehat{\varphi}(r,\theta) := \boldsymbol{\eta}(r)$, while the 2D radial wavelet function ψ is defined by $\widehat{\psi}(r,\theta) := \boldsymbol{\zeta}(r)$. The function $\widehat{\psi}(2^{-j}\cdot)$ is supported on an annulus $\{(r,\theta) : 2^j c_1 \le r \le 2^j c_2, \theta \in [0,2\pi)\}$. Obviously, ψ is an isotropic function. But directionality can be easily achieved by splitting $\widehat{\psi}$ in the angular direction θ with a smooth partition of unity $\alpha_{j,\ell}(\theta)$ for $[0,2\pi)$: $\sum_{\ell=1}^{s_j} |\alpha_{j,\ell}(\theta)|^2 = 1$, $\theta \in [0,2\pi)$. Generators at the scale level j is then given by

$\widehat{\psi^{j,\ell}}(r,\theta) = \boldsymbol{\zeta}(r)\alpha_{j,\ell}(\theta), \ell = 1,\dots,s_j$. The directional tight framelet systems are then obtained by applying the isotropic dilation $\mathsf{M} = 2\mathsf{I}_2$ and translation to the generators, which result in wavelet atoms of the form $\psi^{j,\ell}_{\mathsf{M}^j;\mathsf{k}}$ and the whole system is a tight frame for $L_2(\mathbb{R}^2)$ with all its generators in the Schwartz class.

Although directional tight framelets can easily achieve directionality, yet they still use the isotropic dilation matrices. The system is thus too 'dense' to provide optimally sparse approximation for 2D piecewise C^2 functions with singularity along a closed C^2 curve. By using the parabolic dilation $\mathsf{A} = \mathrm{diag}(2, \sqrt{2})$ instead of an isotropic dilation, the curvelets introduced in [2] not only can achieve directionality selectivity, but also provide optimally sparse approximation for 2D piecewise C^2 functions away from a closed C^2 curve; see [2,10,25,26] for more details on the optimally sparse approximation. The curvelet atom is of the form $\psi^{j,\ell}_{\mathsf{A}^j R_{\theta_{j,\ell}};\mathsf{k}}$ with $R_{\theta_{j,\ell}}$ being a rotation operation determined by the angle $\theta_{j,\ell}$. In other words, each generator $\psi^{j,\ell}$ is attached with a dilation matrix $\mathsf{M}_{j,\ell} := \mathsf{A}^j R_{\theta_{j,\ell}}$ that is determined by both scaling and rotation.

The curvelets use parabolic scaling and rotation and can achieve both directionality and optimally sparse approximation. However, the rotation operation R_θ destroys the preservation of the integer lattice $\mathbb{Z}^2$ since $R_\theta\mathbb{Z}^2$ is not necessarily an integer lattice, yet the integer lattice preservation is a very much desired property in applications. Shearlets, introduced in [7,8,10], replace rotation R_θ by shear S_ℓ. The shear operator not only preserves the integer lattice $S_\ell\mathbb{Z}^2 = \mathbb{Z}^2$, but also enables a shearlet system with only a few generators; that is, $\psi^{j,\ell}$ could come from the shear versions of several generators (even one single generator in the case of non-cone-adapted shearlets [9]). Let $\mathsf{A}_1 := \mathrm{diag}(4, 2)$ and $\mathsf{A}_2 := \mathrm{diag}(2, 4)$. A *cone-adapted shearlet system* in [8,10] is generated by three generators φ (for the low frequency region), ψ^1 (for the horizontal cone in the frequency plane), and $\psi^2 := \psi^1(\mathsf{E}\cdot)$ (for the vertical cone in the frequency plane), through shear, parabolic scaling, and translation:

$$\begin{aligned}\mathsf{CSH}\big(\varphi; \{\psi^1,\psi^2\}\big) = {}& \big\{\varphi(\cdot-\mathsf{k}) : \mathsf{k}\in\mathbb{Z}^2\big\}\\ &\cup \big\{2^{3j/2}\psi^1\big(S^\ell\mathsf{A}_1^j\cdot-\mathsf{k}\big) : \ell=-2^j,\dots,2^j, \mathsf{k}\in\mathbb{Z}^2, j\in\mathbb{N}_0\big\}\\ &\cup \big\{2^{3j/2}\psi^2\big(S_\ell\mathsf{A}_2^j\cdot-\mathsf{k}\big) : \ell=-2^j,\dots,2^j, \mathsf{k}\in\mathbb{Z}^2, j\in\mathbb{N}_0\big\}.\end{aligned} \tag{1.4}$$

It is obvious that the above shearlet system is indeed a special case of the affine shear systems defined in (1.3) by noting that $2^{3j/2}\psi(S^\ell\mathsf{A}_1^j\cdot-\mathsf{k}) = 2^{3j/2}\psi(S^\ell(\mathsf{A}_1^j\cdot-S^{-\ell}\mathsf{k})) = \mathring{\psi}_{\mathsf{A}_1^j;S^{-\ell}\mathsf{k}}$ with $\mathring{\psi} := \psi(S^\ell\cdot)$. The system defined above in (1.4) is in general not a tight frame for $L_2(\mathbb{R}^2)$. In the case of bandlimited generators, such a system can be modified into a tight frame for $L_2(\mathbb{R}^2)$ by using projection techniques [10], which cut the seamline elements $\psi^1(S^\ell\mathsf{A}_1^j\cdot-\mathsf{k}), \psi^2(S_\ell\mathsf{A}_2^j\cdot-\mathsf{k})$ with $\ell = \pm 2^j$ into half pieces in the frequency domain and then restrict them strictly in each cone. Such projection techniques will result in non-smooth shearlets in the frequency domain along the seamlines: $\psi^{1,\pm}(S^{\pm 2^j}\mathsf{A}_1^j\cdot-\mathsf{k}), \psi^{2,\pm}(S_{\pm 2^j}\mathsf{A}_2^j\cdot-\mathsf{k})$.

The non-smoothness of the seamline elements breaks down the arguments in the proof of the optimally sparse approximation for 2D piecewise C^2 functions with singularities along a closed C^2 curve in [10], in which at least twice differentiability is needed for the shearlet atoms in the frequency domain. Recently, Guo and Labate in [13] proposed another type of shearlet-like construction. The idea is still the frequency splitting; but this time for the rectangular strip from the Fourier transform $\widehat{\varphi}$ of the Meyer 2D tensor product scaling function. The splitting is applied to $\widehat{\psi^j} := \sqrt{|\widehat{\varphi}(2^{-2j-2}\cdot)|^2 - |\widehat{\varphi}(2^{-2j}\cdot)|^2}$. A gluing procedure is applied to the two pieces along the seamlines coming from different cones. With appropriate construction, the gluing procedure is smooth and the system in [13] consists of smooth shearlet-like atoms. However, due to the inconsistency of the two cones, a different dilation matrix is needed for the glued shearlet-like atom. We shall discuss the connections of such systems to our affine shear systems in more details in Subsection 4.4.

Though there are various constructions of shearlets available in the literature [8,10,13,22], several key problems remain unresolved. In particular, the following three issues:

Q1) Existence of smooth shearlets. The cone-adapted shearlet system is obtained by applying shear, parabolic scaling, and translation to a few generators. To achieve tightness of the system, the shearlet atoms along the seamlines need to be cut into half pieces. One way to achieve smoothness is by using the gluing procedure as in [13]. However, the system no longer has a full shear structure and is not affine-like. Are there affine shear tight frames using one or a few generators?

Q2) Shearlets with MRA structure. The cone-adapted shearlets achieve directionality by using a parabolic dilation A_λ (in fact it essentially uses two parabolic dilations: $\mathsf{A}_\lambda = \operatorname{diag}(\lambda^2, \lambda)$ for the horizontal cone, and $\mathsf{E}\mathsf{A}_\lambda\mathsf{E} = \operatorname{diag}(\lambda, \lambda^2)$ for the vertical cone) and the shear matrices S^ℓ, S_ℓ while try to keep the generators $\psi^{j,\ell}$ at all scales to be the same. In essence, directionality is achieved in a shearlet (or curvelet) system by using infinitely many dilation matrices so that the initial direction of the generator ψ is dilated and sheared (or rotated) to other directions. This is the main difficulty to build a shearlet system having a multiresolution structure where only a single dilation matrix is employed. It is shown in [20] that there is no traditional shearlet MRA $\{\mathscr{V}_j\}_{j\in\mathbb{Z}}$ with scaling function φ having nice decay property, where $\mathscr{V}_j = \overline{\operatorname{span}}\{\varphi_{S^\ell \mathsf{A}_\lambda^j;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell \in I_j\}$ for some index set I_j. In this case, the space $\mathscr{V}_j$ uses many (possibly infinitely many) dilation matrices. Are there MRA structures in certain setting for shearlet systems?

Q3) Filter bank association. Once we have an MRA for a shear system, it is then natural to ask whether there also exists an associated filter bank system for the shear system. [18,27] have studied the filter bank system with shear operation directly in the discrete setting and provide characterization for such a shear filter bank system. However, it is still not clear whether a filter bank system exists and can be naturally induced from the constructed shear system.

Recently, smooth shearlet-like tight frames have been constructed in [13] using Meyer wavelets with filters. The availability of filters in such shearlet systems in [13] indeed facilitates the computation of coefficients in a shearlet representation. However, to have a fast discrete transform similar to the traditional fast wavelet transform, one must have a sequence of affine shear tight frames with MRA structure and filter banks at every scale level [15,16]. A fast wavelet transform simply transforms between two sets of coefficients in the representations under a sequence of wavelet bases at two consecutive scale levels. Detailed discussion will be given in Subsection 4.4 about the connections and differences of our constructions in this paper with other constructions in [8,10,13].

1.2. Our contributions

In this paper, since shear operation has many nice properties in both theory (optimally sparse approximation, rich group structures, etc., see [23,25]) and applications (edge detection, inpainting, separation, etc., see [11,12,21,24]), we shall focus on the construction of directional multiscale representation systems with shear operation: affine shear systems. Along the way, we will focus on the above issues as discussed in Q1–Q3.

For smoothness, we show that by using one inner smooth generator ψ and only a few smooth boundary generators $\psi^{j,\ell}$ (at most 8 boundary generators in total for each scale level j and they are actually generated by only 2 generators through shear and 'flip' for the non-stationary construction), we can indeed construct smooth affine shear tight frames with all generators in the Schwartz class. In addition, in this paper, we study sequences of affine shear systems. We show that a sequence of affine shear tight frames naturally induces an MRA structure. We would like to point out here that almost all existing approaches and constructions of shearlets [8,10,13] study only one shear system, while it is of fundamental importance to investigate sequences of shear systems as discussed in [15,16].

We propose two approaches for the construction of sequences of smooth affine shear tight frames. One is non-stationary construction and the other is quasi-stationary construction. The function φ^j for the non-stationary construction is different at different scale levels j, while the quasi-stationary construction has a fixed scaling function $\varphi^j = \varphi$. These two approaches actually share the similar idea of frequency splitting as that for the construction of directional tight framelets: at the scale level j, a smooth 2D wavelet function $\boldsymbol{\omega}^j = (|\widehat{\varphi^{j+1}}(\lambda^{-2}\cdot)|^2 - |\widehat{\varphi^j}|^2)^{1/2}$ is constructed; then a smooth partition of unity $\boldsymbol{\gamma}_{j,\ell}, \ell = 1, \dots, s_j$ for $\mathbb{R}^2\backslash\{0\}$ such that $\sum_{\ell=1}^{s_j} |\boldsymbol{\gamma}_{j,\ell}|^2 = 1$ is created using shear operations for two cones instead of rotation for the case of directional tight framelets [16] or curvelets [2]; eventually, generators $\widehat{\psi^{j,\ell}}$ in the frequency domain at the scale level j are obtained by applying $\boldsymbol{\gamma}_{j,\ell}$ to $\boldsymbol{\omega}^j$.

By carefully designing the function $\boldsymbol{\omega}^j$, we show that we can indeed generate a smooth affine shear tight frame (or a sequence of affine shear tight frames), which contains a subsystem (or a sequence of subsystems) that is generated by only one generator. In fact, for the non-stationary case, we will see that $\psi^{j,\ell} = \psi$ for all ℓ except those ℓ with respect to seamline elements (at most 8 in total and they can be generated by only 2 elements). In other words, the shear operations in the non-stationary construction can reach arbitrarily close to the seamlines. For the quasi-stationary construction, we will see that $\psi^{j,\ell} = \psi$ for a total number of ℓ that is proportional to λ^j. In this case, the shear operators in each cone are restricted inside an area with a fixed opening angle. We shall discuss these two types of constructions in Section 4 with more details.

The non-stationary construction and quasi-stationary construction induce two types of MRA structure: non-stationary MRA and stationary MRA. Both of these two types of MRAs are the traditional wavelet MRAs in the sense that the space $\mathscr{V}_j$ is generated by the function (φ or φ^j) using a fixed dilation matrix $\mathsf{M} = \lambda^2\mathsf{I}_2$. On the other hand, the space $\mathscr{W}_j$ is generated by ψ and $\psi^{j,\ell}$ using many dilation matrices determined by shears and parabolic scalings. We show that such types of constructions have a very close relation with the directional tight framelets developed in [14,16]. By a simple modification, we show that the construction of directional tight framelets developed in [14,16] using tensor product on the polar coordinate can be easily adapted to the setting of Cartesian coordinate under the cone-adapted setting. For the directional tight framelets, it is natural and easy to build a directional tight frame with MRA structure and with an underlying filter bank. We show that certain affine shear tight frames can be regarded as a subsystem of certain directional tight framelets. Therefore, such affine shear tight frames have an inherited MRA structure and filter banks from the corresponding directional tight framelets. This observation implies that the transform of such affine shear tight frames can be implemented through the filter banks of their corresponding directional tight framelets.

1.3. Contents

The structure of this paper is as follows. In Section 2, we shall provide a characterization of an affine shear system to be a tight frame in $L_2(\mathbb{R}^2)$. Based on the characterization, simple characterization conditions can be obtained for affine shear systems with nonnegative generators in the frequency domain. Then, we shall present a toy example of bandlimited affine shear tight frames generated by Shannon-like functions (characteristic functions in the frequency domain). In Section 3, since sequences of affine shear systems play a very important role in our study of the MRA structure of affine shear systems, we shall characterize a sequence of affine shear systems to be a sequence of affine shear tight frames for $L_2(\mathbb{R}^2)$. Correspondingly, simple characterization conditions on sequences of affine shear tight frames with nonnegative generators in the frequency domain shall be given. Based on the characterization results, in Section 4, we provide details for the construction of smooth affine shear tight frames with all generators in the Schwartz class. Two approaches shall be introduced, one is the non-stationary construction and the other is the quasi-stationary construction. The connection of our construction of affine shear systems to other existing shear systems shall also be addressed. In Section 5, we shall investigate the relation between our affine shear systems and the directional tight framelets in [14,16]. By modifying the generators for directional tight framelets, we shall construct cone-adapted directional tight framelets, with which a natural filter bank is associated. We shall show that for A_λ with an integer $\lambda > 1$, an affine shear tight frame is in fact a subsystem of a cone-adapted directional tight framelet and therefore an affine shear system has also an inherited filter bank. In Section 6 we shall discuss how to construct a particular family of smooth quasi-stationary affine shear tight frames with MRA structure through the construction of directional tight framelet filter banks. Numerical implementation of our affine shear tight frames, its application to image denoising, as well as performance comparison to curvelets and shearlets will be discussed in Section 6. Some extension and discussion shall be given in Section 7. Some proofs are postponed to Section 8.

2. Affine shear tight frames

Affine systems and their properties have been studied by many researchers, e.g., see [4,6,14–16,31]. In this section we introduce and characterize affine shear tight frames. Based on the characterization, we show that simple characterization conditions could be obtained for affine shear tight frames with generators being nonnegative in the frequency domain. To prepare our study of smooth affine shear tight frames in later sections, we shall present a toy example of bandlimited affine shear tight frames at the end of this section.

For $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$ given as in (1.3) with Ψ_j being given as in (1.2), we define the following functions:

$$\begin{aligned}
&\mathcal{I}_{\varphi}^{\mathsf{k}}(\xi) := \overline{\widehat{\varphi}(\xi)}\widehat{\varphi}(\xi+2\pi\mathsf{k}), \quad \mathsf{k}\in\mathbb{Z}^2,\ \xi\in\mathbb{R}^2;\\
&\mathcal{I}_{\Psi_j}^{\mathsf{k}}(\xi) := \sum_{\ell=-s_j}^{s_j} \overline{\widehat{\psi^{j,\ell}}(S_\ell\xi)}\widehat{\psi^{j,\ell}}\big(S_\ell(\xi+2\pi\mathsf{k})\big), \quad \mathsf{k}\in\mathbb{Z}^2,\ \xi\in\mathbb{R}^2,\ \psi^{j,\ell}=\psi \text{ for } |\ell|\le r_j;\\
&\mathcal{I}_{\varphi}^{\mathsf{k}}(\xi) = \mathcal{I}_{\Psi_j}^{\mathsf{k}}(\xi) := 0, \quad \mathsf{k}\in\mathbb{R}^2\backslash\mathbb{Z}^2,\ \xi\in\mathbb{R}^2.
\end{aligned} \tag{2.1}$$

We say that $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$ is an *affine shear tight frame* for $L_2(\mathbb{R}^2)$ if all generators $\{\varphi\}\cup\{\Psi_j\}_{j=0}^{\infty}\subseteq L_2(\mathbb{R}^2)$ and

$$\|f\|_2^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2}\big|\langle f,\varphi(\cdot-\mathsf{k})\rangle\big|^2 + \sum_{j=0}^{\infty}\sum_{h\in\Psi_j}\sum_{\mathsf{k}\in\mathbb{Z}^2}\big(\big|\langle f,h_{\mathsf{A}_\lambda^j;\mathsf{k}}\rangle\big|^2 + \big|\langle f,h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}}\rangle\big|^2\big) \quad \forall f\in L_2\big(\mathbb{R}^2\big). \tag{2.2}$$

The analysis of $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$ often takes place in the frequency domain. Since we shall apply the results from [16], following [15,16], we define *a frequency-based affine shear system* to be

$$\mathsf{FAS}\big(\widehat{\varphi};\{\widehat{\Psi_j}\}_{j=0}^{\infty}\big) = \big\{\widehat{\varphi}_{0,\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2\big\}\cup\big\{\mathbf{h}_{\mathsf{B}_\lambda^j;0,\mathsf{k}},\mathbf{h}_{\mathsf{B}_\lambda^j\mathsf{E};0,\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2,\mathbf{h}\in\widehat{\Psi_j}\big\}_{j=0}^{\infty},$$

where $\widehat{\Psi_j} := \{\widehat{h} : h\in\Psi_j\}$. Observe that $\widehat{f_{U;\mathsf{k}}} = \widehat{f}_{U^{-\mathsf{T}};0,\mathsf{k}}$. Within the framework of tempered distributions, it is straightforward to see that $\mathsf{FAS}(\widehat{\varphi};\{\widehat{\Psi_j}\}_{j=0}^{\infty})$ is just the image of $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$ under the Fourier transform. The word *frequency-based* here simply means that all discussions take place in the frequency domain and it is not a synonym at all for the word *bandlimited* (i.e., compactly supported in the frequency domain). As argued in [15,16], it is more convenient and important to study the frequency-based system $\mathsf{FAS}(\widehat{\varphi};\{\widehat{\Psi_j}\}_{j=0}^{\infty})$ than the spatially-defined system $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$. Since we are only interested in affine shear tight frames in this paper, due to the Plancherel identity $\langle f,g\rangle = \frac{1}{(2\pi)^2}\langle\widehat{f},\widehat{g}\rangle$ for $f,g\in L_2(\mathbb{R}^2)$, it is straightforward to check [15,16] that $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^{\infty})$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ if and only if $\mathsf{FAS}(\widehat{\varphi};\{\widehat{\Psi_j}\}_{j=0}^{\infty})$ is a frequency-based affine shear tight frame for $L_2(\mathbb{R}^2)$, that is, $\{\widehat{\varphi}\}\cup\{\widehat{\Psi_j}\}_{j=0}^{\infty}\subseteq L_2(\mathbb{R}^2)$ and

$$(2\pi)^2\|\mathbf{f}\|_2^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2}\big|\langle\mathbf{f},\widehat{\varphi}_{0,\mathsf{k}}\rangle\big|^2 + \sum_{j=0}^{\infty}\sum_{\mathbf{h}\in\widehat{\Psi_j}}\sum_{\mathsf{k}\in\mathbb{Z}^2}\big(\big|\langle\mathbf{f},\mathbf{h}_{\mathsf{B}_\lambda^j;0,\mathsf{k}}\rangle\big|^2 + \big|\langle\mathbf{f},\mathbf{h}_{\mathsf{B}_\lambda^j\mathsf{E};0,\mathsf{k}}\rangle\big|^2\big) \quad \forall\mathbf{f}\in L_2\big(\mathbb{R}^2\big).$$

For the convenience of the reader, in this paper we state all results in the spatial domain and try to avoid the direct appearance of frequency-based systems. However, to better understand our analysis and proofs in this paper, it is quite helpful to keep in mind the close relations of an affine shear system $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ with the frequency-based affine shear system $\mathsf{FAS}(\widehat{\varphi}; \{\widehat{\Psi_j}\}_{j=0}^{\infty})$.

We now characterize the system in (1.3) to be an affine shear tight frame. We have the following characterization.

Theorem 1. *Let $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ be defined as in (1.3). Define $\Lambda := \bigcup_{j=0}^{\infty}([\mathsf{A}_\lambda^j\mathbb{Z}^2] \cup [\mathsf{E}\mathsf{A}_\lambda^j\mathbb{Z}^2])$. Then $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ if and only if*

$$\mathcal{I}_\varphi^0(\xi) + \sum_{j=0}^{\infty}\left[\mathcal{I}_{\Psi_j}^0\left(\mathsf{B}_\lambda^j\xi\right) + \mathcal{I}_{\Psi_j}^0\left(\mathsf{B}_\lambda^j\mathsf{E}\xi\right)\right] = 1, \quad a.e.\ \xi \in \mathbb{R}^2 \tag{2.3}$$

and

$$\mathcal{I}_\varphi^{\mathsf{k}}(\xi) + \sum_{j=0}^{\infty}\left[\mathcal{I}_{\Psi_j}^{\mathsf{B}_\lambda^j\mathsf{k}}\left(\mathsf{B}_\lambda^j\xi\right) + \mathcal{I}_{\Psi_j}^{\mathsf{B}_\lambda^j\mathsf{E}\mathsf{k}}\left(\mathsf{B}_\lambda^j\mathsf{E}\xi\right)\right] = 0, \quad a.e.\ \xi \in \mathbb{R}^2, \mathsf{k} \in \Lambda\backslash\{0\}, \tag{2.4}$$

where the sum in (2.3) converges absolutely and the infinite sum in (2.4) is finite for almost every $\xi \in \mathbb{R}^2$.

Proof. Since $\lambda > 1$, the set $B_r(0) \cap \Lambda$ is finite for any ball $B_r(0)$ with radius $r > 0$. Hence, Λ has no accumulation point. Moreover,

$$\left\{j \in \mathbb{N} \cup \{0\} \ : \ \mathsf{B}_\lambda^j\mathsf{k} \in \mathbb{Z}^2 \text{ or } \mathsf{B}_\lambda^j\mathsf{E}\mathsf{k} \in \mathbb{Z}^2\right\} \text{ is a finite set for every } \mathsf{k} \in \Lambda\backslash\{0\}, \tag{2.5}$$

since $\lim_{j\to\infty}\mathsf{B}_\lambda^j\mathsf{k} = 0$ and $\lim_{j\to\infty}\mathsf{B}_\lambda^j\mathsf{E}\mathsf{k} = 0$. Now the claim follows directly from [16, Theorem 11 and Corollary 12]. □

When all generators $\varphi, \psi, \psi^{j,\ell}$ are nonnegative in the frequency domain; that is $\widehat{\varphi} \geq 0$, $\widehat{\psi} \geq 0$, and $\widehat{\psi^{j,\ell}} \geq 0$ for all j, ℓ, the characterization in Theorem 1 becomes

Corollary 1. *Let $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ be defined as in (1.3). Suppose*

$$\widehat{h}(\xi) \geq 0, \quad a.e.\ \xi \in \mathbb{R}^2,\ \forall h \in \{\varphi\} \cup \{\Psi_j\}_{j=0}^{\infty}. \tag{2.6}$$

Then $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ if and only if

$$\left|\widehat{\varphi}(\xi)\right|^2 + \sum_{j=0}^{\infty}\sum_{h\in\Psi_j}\left(\left|\widehat{h}\left(\mathsf{B}_\lambda^j\xi\right)\right|^2 + \left|\widehat{h}\left(\mathsf{B}_\lambda^j\mathsf{E}\xi\right)\right|^2\right) = 1 \tag{2.7}$$

for a.e. $\xi \in \mathbb{R}^2$ and

$$\widehat{h}(\xi)\widehat{h}(\xi + 2\pi\mathsf{k}) = 0, \quad a.e.\ \xi \in \mathbb{R}^2,\ \forall \mathsf{k} \in \mathbb{Z}^2\backslash\{0\},\ and\ \forall h \in \{\varphi\} \cup \{\Psi_j\}_{j=0}^{\infty}. \tag{2.8}$$

Proof. Obviously, (2.3) is equivalent to (2.7). When all generators are nonnegative in the frequency domain, (2.4) is equivalent to (2.8). Now the claim follows directly from Theorem 1. □

By Corollary 1, we see that when all generators are nonnegative in the frequency domain, condition (2.7) is essentially saying that a partition of unity on the frequency plane is required for the system $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$

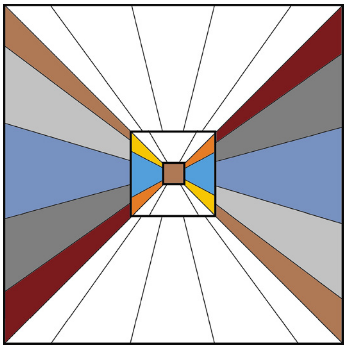

Fig. 1. Frequency tiling of $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ generated by functions in (2.9) and (2.10) with $\lambda = 2$. Inner rectangle: $\widehat{\varphi}$. Middle rectangle: $\widehat{\psi}, \widehat{\psi^{0,-1}}(S_{-1}\cdot), \widehat{\psi^{0,+1}}(S_1\cdot)$ and their flipped versions. Outer rectangle: $\widehat{\psi}(S_\ell \mathsf{B}_2\cdot), \ell = -1, 0, 1, \widehat{\psi^{1,-2}}(S_{-2}\mathsf{B}_2\cdot), \widehat{\psi^{1,+2}}(S_2\mathsf{B}_2\cdot)$ and their flipped versions.

to be a tight frame for $L_2(\mathbb{R}^2)$. Condition (2.8) says that each generator in the frequency domain should not overlap with its 2π-shifted version. In summary, the characterization in Theorem 1 is simplified to a partition of unity condition and a non-overlapping condition.

To prepare for our study of smooth affine shear tight frames in later sections, we next give a simple example of bandlimited affine shear tight frames whose generators are characteristic functions in the frequency domain.

Let $\lambda > 1$ and define $\ell_{\lambda^j} := \lfloor \lambda^j - 1/2 \rfloor + 1$. Choose $0 < \rho \le 1$. Let

$$\begin{aligned} Q &:= \big\{\xi \in \mathbb{R}^2 : -1/2 \le \xi_2/\xi_1 \le 1/2, |\xi_1| \in (\lambda^{-2}\rho\pi, \rho\pi]\big\}, \\ Q_{j,+} &:= \big\{\xi \in \mathbb{R}^2 : -1/2 \le \xi_2/\xi_1 \le \lambda^j - \ell_{\lambda^j}, |\xi_1| \in (\lambda^{-2}\rho\pi, \rho\pi]\big\}, \\ Q_{j,-} &:= \big\{\xi \in \mathbb{R}^2 : -\lambda^j + \ell_{\lambda^j} \le \xi_2/\xi_1 \le 1/2, |\xi_1| \in (\lambda^{-2}\rho\pi, \rho\pi]\big\}. \end{aligned}$$

Define

$$\widehat{\varphi} := \chi_{[-\lambda^{-2}\rho\pi, \lambda^{-2}\rho\pi]^2}, \quad \widehat{\psi} := \chi_Q, \quad \widehat{\psi^{j,\ell_{\lambda^j}}} := \chi_{Q_{j,-}}, \quad \widehat{\psi^{j,-\ell_{\lambda^j}}} := \chi_{Q_{j,+}}. \tag{2.9}$$

Let

$$\Psi_j := \big\{\psi\big(S^{-\ell}\cdot\big) : \ell = -\ell_{\lambda^j} + 1, \dots, \ell_{\lambda^j} - 1\big\} \cup \big\{\psi^{j,\ell_{\lambda^j}}\big(S^{-\ell_{\lambda^j}}\cdot\big), \psi^{j,-\ell_{\lambda^j}}\big(S^{\ell_{\lambda^j}}\cdot\big)\big\}. \tag{2.10}$$

Using Corollary 1, we have the following result whose proof will be given in Section 8.

Corollary 2. *Let $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ be defined as in (1.3) with φ and Ψ_j being given as in (2.9) and (2.10). Then $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$.*

See Fig. 1 for an illustration of $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ with $\lambda = 2$. One of the main goals of this paper is to construct smooth affine shear tight frames that in certain sense can be regarded as the smoothened version (in the frequency domain) of $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ in Corollary 2.

3. Sequences of affine shear tight frames

Most current papers in the literature have investigated only one single affine system. However, to have MRA structure, as argued in [15,16], it is of fundamental importance to study a sequence of affine systems. In order to study the MRA structure of affine shear systems, we next study sequences of affine shear systems.

We first characterize a sequence of affine shear systems to be a sequence of affine shear tight frames for $L_2(\mathbb{R}^2)$. Then, corresponding to Corollary 1, a simple characterization will be given for a sequence of affine shear tight frames with generators being nonnegative in the frequency domain. For $\lambda \neq 0$, we define the following 2×2 matrices

$$\mathsf{M}_\lambda := \lambda^2 \mathsf{I}_2, \qquad \mathsf{N}_\lambda := \mathsf{M}_\lambda^{-\mathsf{T}} = \lambda^{-2}\mathsf{I}_2, \quad \text{and} \quad \mathsf{D}_\lambda := \operatorname{diag}(1, \lambda). \tag{3.1}$$

We shall use M_λ with $\lambda > 1$ as the dilation matrix for the underlying MRA of the affine shear systems in this paper. Let J be an integer. Let $\varphi^j, \psi, \psi^{j,\ell}, |\ell| = r_j + 1, \dots, s_j, j \geq J$ be functions in $L_2(\mathbb{R}^2)$. Let Ψ_j be defined as in (1.2) and $\mathsf{A}_\lambda, \mathsf{B}_\lambda, S^\ell, S_\ell, \mathsf{E}$ be defined as in (1.1). An *affine shear system* $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ (with the initial scale J) is then defined to be

$$\mathsf{AS}_J\big(\varphi^J; \{\Psi_j\}_{j=J}^\infty\big) := \big\{\varphi^J_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\big\} \cup \big\{h_{\mathsf{A}_\lambda^j;\mathsf{k}}, h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\big\}_{j=J}^\infty. \tag{3.2}$$

Considering all integers $J \geq J_0$ for some integer J_0, we then can define a sequence $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$, $J \geq J_0$ of affine shear systems. We denote by $\mathscr{D}(\mathbb{R}^d)$ the linear space of all compactly supported C^∞ (test) functions with the usual topology and recall that $\mathsf{B}_\lambda = (\mathsf{A}_\lambda)^{-\mathsf{T}}$ and $\mathsf{N}_\lambda = (\mathsf{M}_\lambda)^{-\mathsf{T}}$. We have the following characterization for a sequence of affine shear systems $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$, $J \geq J_0$ to be a sequence of affine shear tight frames for $L_2(\mathbb{R}^2)$.

Theorem 2. *Let J_0 be an integer and $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ be defined as in (3.2) with integers $J \geq J_0$. Then the following statements are equivalent to each other.*

(1) *$\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$, i.e., all generators are from $L_2(\mathbb{R}^2)$ and for all $f \in L_2(\mathbb{R}^2)$,*

$$\|f\|_2^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2} \big|\langle f, \varphi^J_{\mathsf{M}_\lambda^J;\mathsf{k}}\rangle\big|^2 + \sum_{j=J}^\infty \sum_{h\in\Psi_j} \sum_{\mathsf{k}\in\mathbb{Z}^2} \big(\big|\langle f, h_{\mathsf{A}_\lambda^j;\mathsf{k}}\rangle\big|^2 + \big|\langle f, h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}}\rangle\big|^2\big) \tag{3.3}$$

for every integer $J \geq J_0$.

(2) *The following identities hold: for all $\widehat{f} \in \mathscr{D}(\mathbb{R}^2)$ and for all integers $j \geq J_0$*

$$\lim_{j\to\infty} \sum_{\mathsf{k}\in\mathbb{Z}^2} \big|\langle f, \varphi^j_{\mathsf{M}_\lambda^j;\mathsf{k}}\rangle\big|^2 = \|f\|_2^2 \tag{3.4}$$

and

$$\sum_{\mathsf{k}\in\mathbb{Z}^2} \big|\langle f, \varphi^{j+1}_{\mathsf{M}_\lambda^{j+1};\mathsf{k}}\rangle\big|^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2} \big|\langle f, \varphi^j_{\mathsf{M}_\lambda^j;\mathsf{k}}\rangle\big|^2 + \sum_{h\in\Psi_j} \sum_{\mathsf{k}\in\mathbb{Z}^2} \big(\big|\langle f, h_{\mathsf{A}_\lambda^j;\mathsf{k}}\rangle\big|^2 + \big|\langle f, h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}}\rangle\big|^2\big). \tag{3.5}$$

(3) *The following identities hold:*

$$\lim_{j\to\infty} \big\langle \big|\widehat{\varphi^j}\big(\mathsf{N}_\lambda^j\cdot\big)\big|^2, \mathbf{h}\big\rangle = \langle 1, \mathbf{h}\rangle \quad \forall \mathbf{h} \in \mathscr{D}\big(\mathbb{R}^2\big) \tag{3.6}$$

and for all integers $j \geq J_0$,

$$\mathcal{I}_{\varphi^j}^{\mathsf{N}_\lambda^j\mathsf{k}}\big(\mathsf{N}_\lambda^j\xi\big) + \big(\mathcal{I}_{\Psi_j}^{\mathsf{B}_\lambda^j\mathsf{k}}\big(\mathsf{B}_\lambda^j\xi\big) + \mathcal{I}_{\Psi_j}^{\mathsf{B}_\lambda^j\mathsf{E}\mathsf{k}}\big(\mathsf{B}_\lambda^j\mathsf{E}\xi\big)\big) = \mathcal{I}_{\varphi^{j+1}}^{\mathsf{N}_\lambda^{j+1}\mathsf{k}}\big(\mathsf{N}_\lambda^{j+1}\xi\big) \tag{3.7}$$

for a.e. $\xi \in \mathbb{R}^2$, $\mathsf{k} \in ([\mathsf{M}_\lambda^j\mathbb{Z}^2] \cup [\mathsf{M}_\lambda^{j+1}\mathbb{Z}^2] \cup [\mathsf{A}_\lambda^j\mathbb{Z}^2] \cup [\mathsf{E}\mathsf{A}_\lambda^j\mathbb{Z}^2])$, where $\mathcal{I}_{\varphi^j}^{\mathsf{k}}, \mathcal{I}_{\Psi_j}^{\mathsf{k}}$ are similarly defined as in (2.1).

Proof. The claim follows directly from [16, Theorem 13 and Corollary 12]. Since this result plays a central role in this paper, for the convenience of the reader, we provide a proof here by following the lines developed in [16, Theorem 13].

Note that by our assumption on M_λ and A_λ, it is easy to show that

$$\{j \in \mathbb{Z} : j \ge J_0, [\mathsf{N}_\lambda^j B_c(0)] \cap \mathbb{Z}^2 \ne \{0\}\} \quad \text{is a finite set for every } c \in [1,\infty). \tag{3.8}$$

(1)⇒(2). Consider (3.3) with two consecutive J and $J+1$ with $J \ge J_0$. Then the difference gives (3.5). Now by (3.5), it is easy to deduce that,

$$\sum_{\mathsf{k}\in\mathbb{Z}^2} |\langle f, \varphi^{J'}_{\mathsf{M}_\lambda^{J'};\mathsf{k}}\rangle|^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2} |\langle f, \varphi^{J}_{\mathsf{M}_\lambda^{J};\mathsf{k}}\rangle|^2 + \sum_{j=J}^{J'-1} \sum_{h\in\Psi_j} \sum_{\mathsf{k}\in\mathbb{Z}^2} \left(|\langle f, h_{\mathsf{A}_\lambda^j;\mathsf{k}}\rangle|^2 + |\langle f, h_{\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}}\rangle|^2\right) \quad \forall J' \ge J. \tag{3.9}$$

Therefore, by (3.3) and letting $J' \to \infty$, we see that (3.4) holds.

(2)⇒(1). By (3.5), we deduce that (3.9) holds. Letting $J' \to \infty$ and in view of (3.4), we conclude that (3.3) holds.

(2)⇔(3). By [16, Lemma 10], we can show that (3.5) is equivalent to

$$\int_{\mathbb{R}^2} \sum_{\mathsf{k}\in\Lambda_j} \widehat{f}(\xi)\overline{\widehat{f}(\xi+2\pi\mathsf{k})}\left(\left[\mathcal{I}^{\mathsf{N}_\lambda^j\mathsf{k}}_{\varphi^j}(\mathsf{N}_\lambda^j\xi) + \mathcal{I}^{\mathsf{B}_\lambda^j\mathsf{k}}_{\Psi_j}(\mathsf{B}_\lambda^j\xi) + \mathcal{I}^{\mathsf{B}_\lambda^j\mathsf{E}\mathsf{k}}_{\Psi_j}(\mathsf{B}_\lambda^j\mathsf{E}\xi)\right] - \mathcal{I}^{\mathsf{N}_\lambda^{j+1}\mathsf{k}}_{\varphi^{j+1}}(\mathsf{N}_\lambda^{j+1}\xi)\right)d\xi = 0, \tag{3.10}$$

where $\Lambda_j = [\mathsf{M}_\lambda^j\mathbb{Z}^2] \cup [\mathsf{M}_\lambda^{j+1}\mathbb{Z}^2] \cup [\mathsf{A}_\lambda^j\mathbb{Z}^2] \cup [\mathsf{E}\mathsf{A}_\lambda^j\mathbb{Z}^2]$. Since $\mathsf{M}_\lambda = \lambda^2\mathsf{I}_2$ and $\mathsf{A}_\lambda = \mathrm{diag}(\lambda^2, \lambda)$ with $\lambda > 1$, we see that the lattice Λ_j is discrete. By the same argument as in the proof of [16, Theorem 13], we see that (3.10) is equivalent to (3.7).

By [16, Lemma 10] and the Plancherel identity $\langle f, g\rangle = \frac{1}{(2\pi)^2}\langle \widehat{f}, \widehat{g}\rangle$ for $f, g \in L_2(\mathbb{R}^2)$, we see that (3.4) is equivalent to

$$\lim_{j\to\infty} \int_{\mathbb{R}^2} \sum_{\mathsf{k}\in[\mathsf{M}_\lambda^j\mathbb{Z}^2]} \widehat{f}(\xi)\overline{\widehat{f}(\xi+2\pi\mathsf{k})}\mathcal{I}^{\mathsf{N}_\lambda^j\mathsf{k}}_{\varphi^j}(\mathsf{N}_\lambda^j\xi) = \|\widehat{f}\|_2^2 \quad \forall \widehat{f} \in \mathscr{D}(\mathbb{R}^2). \tag{3.11}$$

Since $\widehat{f} \in \mathscr{D}(\mathbb{R}^2)$ is compactly supported, there exists $c > 0$ such that $\widehat{f}(\xi)\overline{\widehat{f}(\xi+2\pi\mathsf{k})} = 0$ for all $\xi \in \mathbb{R}^2$ and $|\mathsf{k}| \ge c$. By (3.8), there exists $J'' \ge J_0$ such that $\widehat{f}(\xi)\overline{\widehat{f}(\xi+2\pi\mathsf{k})} = 0$ for all $\xi \in \mathbb{R}^2$, $\mathsf{k} \in [\mathsf{M}_\lambda^j\mathbb{Z}^2]\backslash\{0\}$, and $j \ge J''$. Consequently, for $j \ge J''$, (3.11) becomes

$$\lim_{j\to\infty} \int_{\mathbb{R}^2} |\widehat{f}(\xi)|^2 \mathcal{I}^0_{\varphi^j}(\mathsf{N}_\lambda^j\xi) = \|\widehat{f}\|_2^2 \quad \forall \widehat{f} \in \mathscr{D}(\mathbb{R}^2),$$

which is equivalent to (3.6). □

As argued in [15,16], the relation in (3.5) is critical for a fast transform algorithm. If all elements $\widehat{\varphi^j}, \widehat{\psi}, \widehat{\psi^{j,\ell}}$ are nonnegative, we have the following simple characterization (also see [16, Corollary 18]):

Corollary 3. *Let J_0 be an integer and $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ be defined as in (3.2) with $J \ge J_0$. Suppose that*

$$\widehat{h} \ge 0 \quad \textit{for all } h \in \{\varphi^j : j \ge J_0\} \cup \{\Psi_j\}_{j=J_0}^\infty. \tag{3.12}$$

Then, for all integers $J \ge J_0$, $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ if and only if

$$\widehat{h}(\xi)\overline{\widehat{h}(\xi+2\pi\mathsf{k})}=0,\quad a.e., \xi\in\mathbb{R}^2, \mathsf{k}\in\mathbb{Z}^2\backslash\{0\} \text{ and } h\in\{\varphi^j : j\ge J_0\}\cup\{\Psi_j\}_{j=J_0}^{\infty}, \tag{3.13}$$

$$\big|\widehat{\varphi^{j+1}}(\mathsf{N}_\lambda^{j+1}\xi)\big|^2=\big|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\xi)\big|^2+\sum_{h\in\Psi_j}\big(\big|\widehat{h}(\mathsf{B}_\lambda^j\xi)\big|^2+\big|\widehat{h}(\mathsf{B}_\lambda^j\mathsf{E}\xi)\big|^2\big),\quad a.e., \xi\in\mathbb{R}^2,\ j\ge J_0, \tag{3.14}$$

and (3.6) *holds.*

Proof. When (3.12) holds, by item (3) of Theorem 2, for $\mathsf{k}\in\mathbb{Z}^2\backslash\{0\}$, (3.7) is equivalent to (3.13). For $\mathsf{k}=0$, (3.7) is equivalent to (3.14). Together with the condition (3.6) and by item (3) of Theorem 2, the claim follows from the equivalence between item (1) and item (3) of Theorem 2. □

The condition in (3.6) can be further simplified as in the following lemma.

Lemma 1. *Suppose that there exist two positive numbers c and C such that*

$$\big|\widehat{\varphi^j}(\xi)\big|\le C,\quad a.e.\ \xi\in[-c,c]^2 \text{ and } \forall j\ge J_0. \tag{3.15}$$

Assume that $\mathbf{g}(\xi):=\lim_{j\to\infty}|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\xi)|^2$ exists for almost every $\xi\in\mathbb{R}^2$. Then (3.6) *holds if and only if $\mathbf{g}(\xi)=1$, a.e. $\xi\in\mathbb{R}^2$.*

Proof. Given $\mathbf{h}\in\mathscr{D}(\mathbb{R}^2)$. Since $\mathbf{h}$ has compact support and $\mathsf{N}_\lambda^{-1}=\mathsf{M}_\lambda$ is expansive, there exists $J\in\mathbb{N}$ such that $|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\xi)|^2|\mathbf{h}(\xi)|\le C^2|\mathbf{h}(\xi)|$ for all $j\ge J$ and $\xi\in\mathbb{R}^2$. Since $\mathbf{h}\in L_1(\mathbb{R}^2)$, by Lebesgue Dominated Convergence Theorem, we have $\lim_{j\to\infty}\langle|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\cdot)|^2,\mathbf{h}\rangle=\langle\lim_{j\to\infty}|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\cdot)|^2,\mathbf{h}\rangle=\langle\mathbf{g},\mathbf{h}\rangle$. Now it is trivial to see that (3.6) holds if and only if $\langle\mathbf{g},\mathbf{h}\rangle=\langle 1,\mathbf{h}\rangle$ for all $\mathbf{h}\in\mathscr{D}(\mathbb{R}^2)$, which is equivalent to $\mathbf{g}(\xi)=1$ for almost every $\xi\in\mathbb{R}^2$. □

Consider the toy example in Corollary 2. Define $\varphi^j:=\varphi$ and $\Psi_j:=\{\psi(S^{-\ell}\cdot):\ell=-\ell_{\lambda^j}+1,\dots,\ell_{\lambda^j}-1\}\cup\{\psi^{j,\pm\ell_{\lambda^j}}(S^{\mp\ell_{\lambda^j}}\cdot)\}$ with $\psi,\psi^{j,\pm\ell_{\lambda^j}}$ being constructed as in Corollary 2. Then condition (3.6) holds by Lemma 1 since φ^j satisfies (3.15) and $\mathbf{g}(\xi)=\lim_{j\to\infty}|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\xi)|^2=1$ a.e., $\xi\in\mathbb{R}^2$. Condition (3.13) directly follows from the proof of Corollary 2 (see Section 8). Condition (3.14) holds by our construction. Therefore, by Corollary 3, $\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^{\infty})$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for any integer $J\ge 0$.

A sequence of affine shear tight frames naturally induces an MRA structure $\{\mathscr{V}_j\}_{j=J_0}^{\infty}$ with $\mathscr{V}_j:=\overline{\operatorname{span}}\{\varphi^j(\mathsf{M}_\lambda^j\cdot-\mathsf{k}):\mathsf{k}\in\mathbb{Z}^2\}$. But so far, the generators in the above toy example and its induced sequence of systems are discontinuous in the frequency domain. In the next section, we shall focus on the construction of smooth affine shear tight frames for $L_2(\mathbb{R}^2)$ in the Schwartz class that have many desirable properties. We shall show that not only our systems can have smooth generators, but also have shear structure and more importantly, an MRA structure could be deduced from such type of systems.

4. Construction of smooth affine shear tight frames

In this section we shall provide two types of constructions of smooth affine shear tight frames: one is non-stationary construction and the other is quasi-stationary construction. Both these two types of constructions use the idea of normalization in the frequency domain. In essence, we first construct a smooth affine shear frame for $L_2(\mathbb{R}^2)$ and then a normalization procedure is applied to such a frame. The non-stationary construction uses different functions φ^j for different scale levels j, while the quasi-stationary construction employs a single function φ for every scale level. We first need some auxiliary results and then provide details on the two types of constructions.

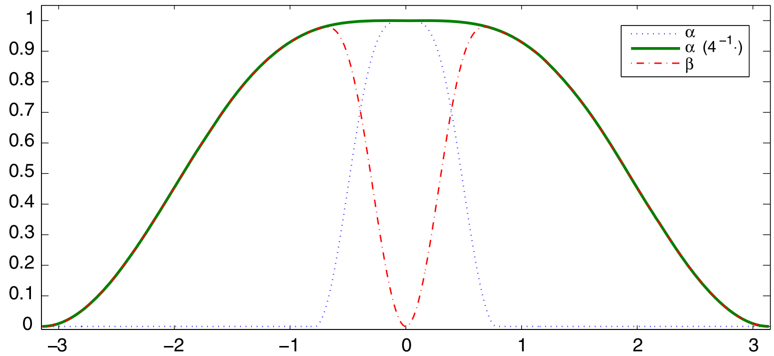

Fig. 2. Graphs of $\boldsymbol{\alpha}_{\lambda,t,\rho}$ (dotted line), $\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2}\cdot)$ (solid line), and $\boldsymbol{\beta}_{\lambda,t,\rho}$ (dashed-dot line) for $\lambda = 2$ and $\rho = t = 1$. Note that $\boldsymbol{\beta}_{\lambda,t,\rho}$ overlaps with $\boldsymbol{\alpha}(\lambda^{-2}\cdot)$ for $\xi \ge \lambda^{-2}\rho\pi$.

4.1. Auxiliary results

We shall use a function $\nu \in C^\infty(\mathbb{R})$ such that $\nu(x) = 0$ for $x \le -1$, $\nu(x) = 1$ for $x \ge 1$, and $|\nu(x)|^2 + |\nu(-x)|^2 = 1$ for all $x \in \mathbb{R}$. There are many choices of such functions. For example, as in [14], we define $f(x) := e^{-1/x^2}$ for $x > 0$ and $f(x) := 0$ for $x \le 0$, and let $g(x) := \int_{-1}^x f(1+t)f(1-t)dt$. Define

$$\nu(x) := \frac{g(x)}{\sqrt{|g(x)|^2 + |g(-x)|^2}}, \quad x \in \mathbb{R}. \tag{4.1}$$

Then $\nu \in C^\infty(\mathbb{R})$ is a desired function. Using such a function ν, we now construct our building blocks $\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho}$ of Meyer-type scaling and wavelet functions with $\lambda > 1$, $0 < t \le 1$, and $0 < \rho \le \lambda^2$ as follows (see Fig. 2):

$$\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi) := \begin{cases} \nu(\frac{\xi+c}{\varepsilon_1}) & \text{if } \xi < -c+\varepsilon, \\ 1 & \text{if } -c+\varepsilon \le \xi \le c-\varepsilon, \\ \nu(\frac{-\xi+c}{\varepsilon}) & \text{if } \xi > c-\varepsilon, \end{cases}$$
$$\boldsymbol{\beta}_{\lambda,t,\rho}(\xi) := \left(\left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2}\xi)\right|^2 - \left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)\right|^2\right)^{1/2}, \tag{4.2}$$

where $c = \lambda^{-2}(1-t/2)\rho\pi$ and $\varepsilon = \lambda^{-2}t\rho\pi/2$. Then $\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho} \in C_c^\infty(\mathbb{R})$, where $C_c^\infty(\mathbb{R})$ denotes the linear space consisting of all compactly supported functions in $C^\infty(\mathbb{R})$. Moreover,

$$\operatorname{supp} \boldsymbol{\alpha}_{\lambda,t,\rho} = \left[-\lambda^{-2}\rho\pi, \lambda^{-2}\rho\pi\right] \quad \text{and} \quad \operatorname{supp} \boldsymbol{\beta}_{\lambda,t,\rho} = \left[-\rho\pi, -\lambda^{-2}(1-t)\rho\pi\right] \cup \left[\lambda^{-2}(1-t)\rho\pi, \rho\pi\right].$$

Furthermore, define a 2π-periodic function $\boldsymbol{\mu}_{\lambda,t,\rho}$ and $\boldsymbol{\upsilon}_{\lambda,t,\rho}$ as follows:

$$\boldsymbol{\mu}_{\lambda,t,\rho}(\xi) := \begin{cases} \frac{\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^2\xi)}{\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)} & \text{if } |\xi| \le \lambda^{-2}\rho\pi, \\ 0 & \text{if } \lambda^{-2}\rho\pi < |\xi| \le \pi, \end{cases}$$
$$\boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi) := \begin{cases} \frac{\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\xi)}{\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)} & \text{if } \lambda^{-4}(1-t)\rho\pi \le |\xi| \le \lambda^{-2}\rho\pi, \\ \mathbf{g}_{\lambda,t,\rho}(\xi) & \text{if } \xi \in [-\pi,\pi) \setminus \operatorname{supp} \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\cdot), \end{cases} \tag{4.3}$$

where $\mathbf{g}_{\lambda,t,\rho}$ is a function in $C^\infty(\mathbb{T})$ such that $[\frac{d^n}{d\xi^n}\mathbf{g}_{\lambda,t,\rho}(\xi)]|_{\xi=\pm\lambda^{-2}\rho\pi} = \delta(n)$ for all $n \in \mathbb{N}_0$. The purpose of $\mathbf{g}_{\lambda,t,\rho}$ is to make the function $\boldsymbol{\upsilon}_{\lambda,t,\rho}$ smooth. Such a $\mathbf{g}_{\lambda,t,\rho}$ exists. In fact, noting that $\frac{\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\xi)}{\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)} = 1$ for $|\xi| \ge \lambda^{-4}\rho\pi$ and $\frac{\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\xi)}{\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)} = 0$ for $|\xi| \le \lambda^{-4}(1-t)\rho\pi$, we can simply define $\mathbf{g}_{\lambda,t,\rho}$ to be $\mathbf{g}_{\lambda,t,\rho}(\xi) := 1$

for $\lambda^{-4}\rho\pi \le |\xi| \le \pi$ and $\mathbf{g}_{\lambda,t,\rho}(\xi) := 0$ for $|\xi| \le \lambda^{-4}(1-t)\rho\pi$. In this case, $\mathbf{g}_{\lambda,t,\rho}$ extends periodically as a constant 1 near the boundary of $\mathbb{T}$. If $\lambda^{-2}\rho < 1$, then another way to make $\boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi)$ smooth is by defining $\mathbf{g}_{\lambda,t,\rho}$ to be $\mathbf{g}_{\lambda,t,\rho}(\xi) := 1$ for $\lambda^{-4}\rho\pi \le |\xi| \le \lambda^{-2}\rho\pi$, and $\mathbf{g}_{\lambda,t,\rho}(\xi) := 0$ for $|\xi| \le \lambda^{-4}(1-t)\rho\pi$ or $\lambda^{-2}\rho_0\pi \le |\xi| \le \pi$ with ρ_0 being a positive constant such that $\lambda^{-2}\rho < \lambda^{-2}\rho_0 < 1$, which can be achieved by using smoothing kernel. We have the following result (see Section 8 for its proof).

Proposition 1. *Let $\lambda > 1$, $0 < t \le 1$, and $0 < \rho \le \lambda^2$. Let $\boldsymbol{\alpha}_{\lambda,t,\rho}$, $\boldsymbol{\beta}_{\lambda,t,\rho}$, and $\boldsymbol{\mu}_{\lambda,t,\rho}, \boldsymbol{\upsilon}_{\lambda,t,\rho}$ be defined as in (4.2) and (4.3), respectively. Then $\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho} \in C_c^\infty(\mathbb{R})$ and $\boldsymbol{\mu}_{\lambda,t,\rho}, \boldsymbol{\upsilon}_{\lambda,t,\rho} \in C^\infty(\mathbb{T})$. Moreover,*

$$\left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)\right|^2 + \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\xi)\right|^2 = \left|\boldsymbol{\alpha}_{\lambda,t,\rho}\left(\lambda^{-2}\xi\right)\right|^2, \quad \xi \in \mathbb{R},$$

and

$$\boldsymbol{\alpha}_{\lambda,t,\rho}\left(\lambda^2\xi\right) = \boldsymbol{\mu}_{\lambda,t,\rho}(\xi)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi), \qquad \boldsymbol{\beta}_{\lambda,t,\rho}\left(\lambda^2\xi\right) = \boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi), \quad \xi \in \mathbb{R}.$$

The functions $\boldsymbol{\alpha}_{\lambda,t,\rho}$ and $\boldsymbol{\beta}_{\lambda,t,\rho}$ shall be used for the horizontal direction. We next define 'bump' function $\boldsymbol{\gamma}_\varepsilon$ for splitting pieces along the vertical direction. Roughly speaking, the core generator for our affine shear systems in the frequency domain looks like $\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1)$, which is a wedge shape generator. Application of parabolic scaling, shear, and translation operations to such a generator induces our affine shear systems. Further technical treatments are then applied on such systems to achieve tightness; see next subsections for details.

In what follows, ε shall be fixed as a constant such that $0 < \varepsilon \le 1/2$. Define a function $\boldsymbol{\gamma}_\varepsilon$ to be

$$\boldsymbol{\gamma}_\varepsilon(x) = \begin{cases} 1 & \text{if } |x| \le 1/2 - \varepsilon, \\ \nu\left(\frac{-|x|+1/2}{\varepsilon}\right) & \text{if } 1/2 - \varepsilon \le |x| \le 1/2 + \varepsilon, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

Then it is easy to check that $\gamma_\varepsilon \in C_c^\infty(\mathbb{R})$ and $\sum_{\ell\in\mathbb{Z}} |\boldsymbol{\gamma}_\varepsilon(\cdot + \ell)|^2 = 1$.

For $\lambda > 1$, define $\ell_\lambda := \lfloor \lambda - (1/2 + \varepsilon) \rfloor + 1 = \lfloor \lambda + (1/2 - \varepsilon) \rfloor$. Define the corner pieces $\gamma^{\pm}_{\lambda,\varepsilon,\varepsilon_0}$ by

$$\begin{aligned} \boldsymbol{\gamma}^{+}_{\lambda,\varepsilon,\varepsilon_0}(\lambda x - \ell_\lambda) &:= \begin{cases} \boldsymbol{\gamma}_\varepsilon(\lambda x - \ell_\lambda) & \text{if } \lambda^{-1}(\ell_\lambda - 1/2 - \varepsilon) \le x \le \lambda^{-1}(\ell_\lambda - 1/2 + \varepsilon), \\ \nu\left(1 + \frac{\lambda^2}{\varepsilon_0}(1 - x)\right) & \text{if } \lambda^{-1}(\ell_\lambda - 1/2 + \varepsilon) \le x \le 1 + \frac{2\varepsilon_0}{\lambda^2}, \end{cases} \\ \boldsymbol{\gamma}^{-}_{\lambda,\varepsilon,\varepsilon_0}(\lambda x + \ell_\lambda) &:= \begin{cases} \boldsymbol{\gamma}_\varepsilon(\lambda x + \ell_\lambda) & \text{if } \lambda^{-1}(-\ell_\lambda + 1/2 - \varepsilon) \le x \le \lambda^{-1}(-\ell_\lambda + 1/2 + \varepsilon), \\ \nu\left(1 + \frac{\lambda^2}{\varepsilon_0}(1 + x)\right) & \text{if } -1 - \frac{2\varepsilon_0}{\lambda^2} \le x \le \lambda^{-1}(-\ell_\lambda + 1/2 + \varepsilon). \end{cases} \end{aligned} \tag{4.5}$$

That is,

$$\begin{aligned} \boldsymbol{\gamma}^{+}_{\lambda,\varepsilon,\varepsilon_0}(x) &= \begin{cases} \boldsymbol{\gamma}_\varepsilon(x) & \text{if } -1/2 - \varepsilon \le x \le -1/2 + \varepsilon, \\ \nu\left(1 + \frac{\lambda^2}{\varepsilon_0} - \frac{\lambda}{\varepsilon_0}(x + \ell_\lambda)\right) & \text{if } -1/2 + \varepsilon \le x \le \lambda(1 + 2\varepsilon_0/\lambda^2) - \ell_\lambda, \end{cases} \\ \boldsymbol{\gamma}^{-}_{\lambda,\varepsilon,\varepsilon_0}(x) &= \begin{cases} \boldsymbol{\gamma}_\varepsilon(x) & \text{if } 1/2 - \varepsilon \le x \le 1/2 + \varepsilon, \\ \nu\left(1 + \frac{\lambda^2}{\varepsilon_0} + \frac{\lambda}{\varepsilon_0}(x - \ell_\lambda)\right) & \text{if } -\lambda(1 + \varepsilon_0/\lambda^2) + \ell_\lambda \le x \le 1/2 - \varepsilon. \end{cases} \end{aligned} \tag{4.6}$$

Here $\varepsilon_0 > 0$ is a parameter to control the overlap of corner pieces around the seamlines. Note that $\gamma^{\pm}_{\lambda,\varepsilon,\varepsilon_0}$ are also C_c^∞ functions. Then, for $\lambda \ge 1$,

$$\left(\sum_{\ell=-\ell_\lambda+1}^{\ell_\lambda-1} \left|\boldsymbol{\gamma}_\varepsilon(\lambda x + \ell)\right|^2\right) + \left|\boldsymbol{\gamma}^{+}_{\lambda,\varepsilon,\varepsilon_0}(\lambda x - \ell_\lambda)\right|^2 + \left|\boldsymbol{\gamma}^{-}_{\lambda,\varepsilon,\varepsilon_0}(\lambda x + \ell_\lambda)\right|^2 = 1 \quad \forall |x| \le 1 \tag{4.7}$$

and

$$\sum_{\ell=-\ell_\lambda}^{\ell_\lambda} \left|\gamma_\varepsilon(\lambda x+\ell)\right|^2 = 1 \quad \forall |x| \le \frac{\ell_\lambda + 1/2-\varepsilon}{\lambda}. \tag{4.8}$$

Accordingly, we next define two functions $\mathbf{\Gamma}^j$ and $\mathbf{\Gamma}_j$, which will be used for frequency splitting along the shear directions. We have the following result (see Section 8 for its proof).

Proposition 2. *Let $j \in \mathbb{N}_0$. Define*

$$\begin{aligned}\mathbf{\Gamma}^j(\xi) := &\left[\sum_{\ell=-\ell_{\lambda^j}+1}^{\ell_{\lambda^j}-1} \left(\left|\gamma_\varepsilon\left(\lambda^j\xi_2/\xi_1+\ell\right)\right|^2 + \left|\gamma_\varepsilon\left(\lambda^j\xi_1/\xi_2+\ell\right)\right|^2\right)\right] + \left|\gamma^+_{\lambda^j,\varepsilon,\varepsilon_0}\left(\lambda^j\xi_2/\xi_1-\ell_{\lambda^j}\right)\right|^2 \\ &+ \left|\gamma^-_{\lambda^j,\varepsilon,\varepsilon_0}\left(\lambda^j\xi_2/\xi_1+\ell_{\lambda^j}\right)\right|^2 + \left|\gamma^+_{\lambda^j,\varepsilon,\varepsilon_0}\left(\lambda^j\xi_1/\xi_2-\ell_{\lambda^j}\right)\right|^2 + \left|\gamma^-_{\lambda^j,\varepsilon,\varepsilon_0}\left(\lambda^j\xi_1/\xi_2+\ell_{\lambda^j}\right)\right|^2\end{aligned} \tag{4.9}$$

and

$$\mathbf{\Gamma}_j(\xi) := \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\gamma_\varepsilon\left(\lambda^j\xi_2/\xi_1+\ell\right)\right|^2 + \left|\gamma_\varepsilon\left(\lambda^j\xi_1/\xi_2+\ell\right)\right|^2\right). \tag{4.10}$$

Then $\mathbf{\Gamma}^j, \mathbf{\Gamma}_j \in C^\infty(\mathbb{R}^2\backslash\{0\})$ have the following properties.

(i) $1 \le \mathbf{\Gamma}^j(\xi) \le 2$, $\mathbf{\Gamma}^j(\mathsf{E}\xi) = \mathbf{\Gamma}^j(\xi)$, *and* $\mathbf{\Gamma}^j(t\xi) = \mathbf{\Gamma}^j(\xi)$ *for all* $t \ne 0$ *and* $\xi \ne 0$.
(ii) $0 < \mathbf{\Gamma}_j(\xi) \le 2$, $\mathbf{\Gamma}_j(\mathsf{E}\xi) = \mathbf{\Gamma}_j(\xi)$, *and* $\mathbf{\Gamma}_j(t\xi) = \mathbf{\Gamma}_j(\xi)$ *for all* $t \ne 0$ *and* $\xi \ne 0$.
(iii) $\mathbf{\Gamma}^j$ *and* $\mathbf{\Gamma}_j$ *satisfy*

$$\mathbf{\Gamma}^j(\xi) = 1, \quad \xi \in \left\{\xi \in \mathbb{R}^2\backslash\{0\} : \max\{|\xi_2/\xi_1|, |\xi_1/\xi_2|\} \le \frac{\lambda^{2j}}{\lambda^{2j}+2\varepsilon_0}\right\}, \tag{4.11}$$

and

$$\mathbf{\Gamma}_j(\xi) = 1, \quad \xi \in \left\{\xi \in \mathbb{R}^2 : \max\{|\xi_2/\xi_1|, |\xi_1/\xi_2|\} \le \frac{\lambda^j}{\ell_{\lambda^j}+1/2+\varepsilon}\right\}. \tag{4.12}$$

Equations (4.9) and (4.10) will be used to construct two types of smooth affine shear tight frames. One is non-stationary construction with φ^j changing at different scale levels and the other is quasi-stationary construction with φ being the same for all scale levels. We next discuss the details of these two types of constructions.

4.2. Non-stationary construction

We first discuss the non-stationary construction. For such a type of construction, the shear operations could reach arbitrarily close to the seamlines when j goes to infinity. The idea of constructing such a smooth affine shear tight frame in the non-stationary setting is simple. We first construct an affine shear frame from only a few generators and then apply normalization to such a frame to obtain a tight frame.

More precisely, we fix $\lambda > 1$, $0 < t \le 1$, $0 < \rho \le 1$, and $0 < \varepsilon \le 1/2$ as parameters. Below, we shall omit the dependency of $\varphi, \boldsymbol{\eta}, \boldsymbol{\zeta}, \mathbf{\Theta}^J$, etc., on the parameters $\lambda, t, \rho, \varepsilon$ for simplicity of presentation. Let $\mathsf{A}_\lambda, \mathsf{B}_\lambda, \mathsf{M}_\lambda, \mathsf{N}_\lambda, \boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho}$, and $\boldsymbol{\gamma}_\varepsilon, \boldsymbol{\gamma}^\pm_{\lambda,\varepsilon,\varepsilon_0}, \ell_\lambda$ be defined as in (1.1), (3.1), (4.2), (4.4), (4.6). Define

$$\begin{aligned}\boldsymbol{\eta}(\xi_1,\xi_2) &:= \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1), \quad (\xi_1,\xi_2) \in \mathbb{R}^2, \\ \boldsymbol{\zeta}(\xi_1,\xi_2) &:= \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1), \quad (\xi_1,\xi_2) \in \mathbb{R}^2\backslash\{0\},\end{aligned} \tag{4.13}$$

as well as the corner pieces

$$\begin{aligned}\boldsymbol{\eta}^{j,\pm\ell_{\lambda^j}}(\xi_1,\xi_2) &:= \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}(\xi_2/\xi_1), \quad (\xi_1,\xi_2)\in\mathbb{R}^2,\\ \boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}}(\xi_1,\xi_2) &:= \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}(\xi_2/\xi_1), \quad (\xi_1,\xi_2)\in\mathbb{R}^2\backslash\{0\}.\end{aligned} \tag{4.14}$$

For $\xi = 0$, $\boldsymbol{\zeta}(0) := 0$ and $\boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}}(0) := 0$. Since the support of $\boldsymbol{\beta}_{\lambda,t,\rho}$ is away from the origin, we have $\boldsymbol{\zeta}, \boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}} \in C_c^\infty(\mathbb{R}^2)$. Let

$$\widehat{\varphi}(\xi) := \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_2), \quad \xi\in\mathbb{R}^2.$$

Then, $\widehat{\varphi}$ is also a function in $C_c^\infty(\mathbb{R}^2)$ hence $\varphi\in C^\infty(\mathbb{R}^2)$.

For a nonnegative integer J_0, define

$$\boldsymbol{\Theta}^{J_0}(\xi) := \left|\widehat{\varphi}(\mathsf{N}_\lambda^{J_0}\xi)\right|^2 + \sum_{j=J_0}^{\infty}\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}}\left[\left|\boldsymbol{\zeta}^{j,\ell}(S_\ell\mathsf{B}_\lambda^j\xi)\right|^2 + \left|\boldsymbol{\zeta}^{j,\ell}(S_\ell\mathsf{B}_\lambda^j\mathsf{E}\xi)\right|^2\right] \tag{4.15}$$

for $\xi\in\mathbb{R}^2$, where for $|\ell|<\ell_{\lambda^j}$, $\boldsymbol{\eta}^{j,\ell} = \boldsymbol{\eta}$ and $\boldsymbol{\zeta}^{j,\ell} = \boldsymbol{\zeta}$, respectively. We have the following result concerning the function $\boldsymbol{\Theta}^{J_0}$ (see Section 8 for its proof).

Proposition 3. *Let $\lambda>1$, $0<\varepsilon\le 1/2$, $0<t\le 1$, and $0<\rho\le 1$. Let J_0 be a nonnegative integer and $\boldsymbol{\Theta}^{J_0}$ be defined as in (4.15). Choose ε_0 such that $0<\varepsilon_0<\frac{1}{2}\lambda^{J_0-1}$. Then $\boldsymbol{\Theta}^{J_0}$ has the following properties:*

(i) $\boldsymbol{\Theta}^{J_0}\in C^\infty(\mathbb{R}^2)$, $\boldsymbol{\Theta}^{J_0} = \boldsymbol{\Theta}^{J_0}(\mathsf{E}\cdot)$, *and* $0<\boldsymbol{\Theta}^{J_0}\le 2$.
(ii) $\boldsymbol{\Theta}^{J_0}(\xi) = \boldsymbol{\Theta}^{J_0}(\mathsf{E}\xi) = 1 \quad \forall\xi\in\cup_{j=J_0+1}^{\infty}\cup_{\ell=-\ell_{\lambda^j}+2}^{\ell_{\lambda^j}-2}[(S_\ell\mathsf{B}_\lambda^j)^{-1}\operatorname{supp}\boldsymbol{\zeta}^{j,\ell}]$.

The function $\boldsymbol{\Theta}^{J_0}$ will be used for the normalization of the frame generated by $\boldsymbol{\zeta}^{j,\ell}$.

Since $0<\boldsymbol{\Theta}^{J_0}\le 2$, we can take the square root of $\boldsymbol{\Theta}^{J_0}$, which is still a smooth function. Moreover, $1/\sqrt{\boldsymbol{\Theta}^{J_0}}$ is also a smooth function. Define $\widehat{\varphi^{J_0}} := \frac{\widehat{\varphi}}{\sqrt{\boldsymbol{\Theta}(\mathsf{M}_\lambda^{J_0}\cdot)}}$ and

$$\boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{N}_\lambda^j\xi) := \frac{\left(\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}}(|\boldsymbol{\zeta}^{j,\ell}(S_\ell\mathsf{B}_\lambda^j\xi)|^2 + |\boldsymbol{\zeta}^{j,\ell}(S_\ell\mathsf{B}_\lambda^j\mathsf{E}\xi)|^2)\right)^{1/2}}{\sqrt{\boldsymbol{\Theta}^{J_0}(\xi)}}, \quad j\ge J_0. \tag{4.16}$$

Define φ^{j+1} to be

$$\widehat{\varphi^{j+1}}(\mathsf{N}_\lambda^{j+1}\xi) := \left(\left|\widehat{\varphi^j}(\mathsf{N}_\lambda^j\xi)\right|^2 + \left|\boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{N}_\lambda^j\xi)\right|^2\right)^{1/2}. \tag{4.17}$$

Now, we split the function $\boldsymbol{\omega}^j_{\lambda,t,\rho}$ as follows. Recall that $\mathsf{D}_\lambda := \operatorname{diag}(1,\lambda)$ as in (3.1). For $\xi\ne 0$, define

$$\widehat{\psi^{j,\ell}}(\xi) := \boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{D}_\lambda^{-j}S_{-\ell}\xi)\frac{\gamma_\varepsilon(\xi_2/\xi_1)}{\boldsymbol{\Gamma}^j((S_\ell\mathsf{B}_\lambda^j)^{-1}\xi)}, \quad \ell = -\ell_{\lambda^j}+1,\dots,\ell_{\lambda^j}-1, \tag{4.18}$$

and

$$\widehat{\psi^{j,\pm\ell_{\lambda^j}}}(\xi) := \boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{D}_\lambda^{-j}S_{\mp\ell_{\lambda^j}}\xi)\frac{\gamma^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}(\xi_2/\xi_1)}{\boldsymbol{\Gamma}^j((S_{\pm\ell_{\lambda^j}}\mathsf{B}_\lambda^j)^{-1}\xi)}. \tag{4.19}$$

For $\xi = 0$, we define $\widehat{\psi^{j,\ell}}(0) := 0$. Since the support of $\boldsymbol{\omega}^j_{\lambda,t,\rho}$ is away from the origin and in view of the properties of $\boldsymbol{\Gamma}^j$, we deduce that $\widehat{\psi^{j,\ell}} \in C_c^\infty(\mathbb{R}^2)$ and hence $\psi^{j,\ell}$ is function in $C^\infty(\mathbb{R}^2)$. Let

$$\Psi_j := \{\psi^{j,\ell}(S^{-\ell}\cdot) : \ell = -\ell_{\lambda_j}, \dots, \ell_{\lambda_j}\} \tag{4.20}$$

with $\psi^{j,\ell}$ being given as in (4.18) and (4.19). The (non-stationary) *affine shear system* $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ is then defined as follows:

$$\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty) := \{\varphi^J_{\mathsf{M}^J_\lambda;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\} \cup \{h_{\mathsf{A}^j_\lambda;\mathsf{k}}, h_{\mathsf{A}^j_\lambda \mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\}_{j=J}^\infty. \tag{4.21}$$

Explicitly, we have,

$$\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty) = \{\varphi^J_{\mathsf{M}^J_\lambda;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\} \cup \{\psi^{j,\ell}_{S^{-\ell}\mathsf{A}^j_\lambda;\mathsf{k}}, \psi^{j,\ell}_{S^{-\ell}\mathsf{A}^j_\lambda \mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\}_{j=J}^\infty. \tag{4.22}$$

With the property of $\boldsymbol{\Theta}^{J_0}$ in item (ii) of Proposition 3, we can show that the system defined in (4.21) can have shear structure for elements inside each cone. Moreover, with the scale j going to infinity, the shear operation could reach the seamline arbitrarily close. Indeed, we have the following result.

Theorem 3. *Let $\lambda > 1$, $0 < \varepsilon \le 1/2$, $0 < t \le 1$, and $0 < \rho \le 1$ such that $1/\rho - 1/2 - \varepsilon > 0$. Let J_0 be a nonnegative integer. Choose $\varepsilon_0 > 0$ such that $\varepsilon_0 < \min\{\frac{\lambda^{J_0-1}}{2}, \lambda^{2J_0}(\frac{\lambda^2}{2\rho} - 1/2), (1/\rho - 1/2 - \varepsilon)\lambda^{J_0}\}$. Then the system $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ defined as in (4.21) with φ^j and Ψ_j being given as in (4.17) and (4.20), respectively, is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge J_0$. All elements in $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ have compactly supported Fourier transforms in $C_c^\infty(\mathbb{R}^2)$. Moreover, let $\psi := \mathscr{F}^{-1}\boldsymbol{\zeta}$. We have*

$$\{\psi(S^{-\ell}\cdot) : |\ell| < \ell_{\lambda^j} - 1\} \subseteq \boldsymbol{\Psi}_j, \quad j \ge J_0 + 1,$$

and

$$\{\psi_{S^{-\ell}\mathsf{A}^j_\lambda;\mathsf{k}}, \psi_{S^{-\ell}\mathsf{A}^j_\lambda \mathsf{E};\mathsf{k}} : j \ge J, \mathsf{k} \in \mathbb{Z}^2, |\ell| < \ell_{\lambda^j} - 1\} \subseteq \underset{J}{\mathsf{AS}}(\varphi^J; \{\Psi_j\}_{j=J}^\infty), \quad J \ge J_0 + 1.$$

Proof. By the property of $\boldsymbol{\Theta}^{J_0}$ in Proposition 3, we see that $\boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{N}^j_\lambda \xi) = \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)$ for $\xi \in \operatorname{supp} \boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}^j_\lambda \cdot)$ with $|\ell| < \ell_{\lambda^j} - 1$, $j \ge J_0 + 1$, and $\boldsymbol{\omega}^j_{\lambda,t,\rho}(\mathsf{N}^j_\lambda \xi) = \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_2)$ for $\xi \in \operatorname{supp} \boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}^j_\lambda \mathsf{E}\cdot)$ with $|\ell| < \ell_{\lambda^j} - 1$ and $j \ge J_0 + 1$. Hence, it is easily seen that for $j \ge J_0 + 1$,

$$\widehat{\psi^{j,\ell}}(\xi) = \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1) = \boldsymbol{\zeta}(\xi) = \widehat{\psi}(\xi), \quad |\ell| < \ell_{\lambda^j} - 1.$$

For $j \ge J_0 + 1$, we observe that $\Psi_j = \{\psi(S^{-\ell}\cdot) : \ell = -\ell_{\lambda^j} + 2, \dots, \ell_{\lambda^j} - 2\} \cup \{\psi^{j,\ell}(S^{-\ell}\cdot) : |\ell| = \ell_{\lambda^j} - 1, \ell_{\lambda^j}\}$.

By our construction, (3.14) and (3.6) hold. Moreover, all generators are nonnegative. Noting that $\operatorname{supp} \boldsymbol{\alpha}_{\lambda,t,\rho} = [-\lambda^{-2}\rho\pi, \lambda^{-2}\rho\pi]$, $\operatorname{supp} \boldsymbol{\beta}_{\lambda,t,\rho} = [-\rho\pi, -\lambda^{-2}\rho\pi] \cup [\lambda^{-2}\rho\pi, \rho\pi]$, and $\operatorname{supp} \gamma_\varepsilon = [-1/2-\varepsilon, 1/2+\varepsilon]$, together with $\rho \le 1$ and $0 < \varepsilon \le 1/2$, we see that $\operatorname{supp} \widehat{\psi^{j,\ell}} \subseteq [-\rho\pi, \rho\pi]^2 \subseteq [-\pi, \pi]^2$ for $|\ell| \le \ell_{\lambda^j} - 1$. Hence, we have $\widehat{\psi^{j,\ell}}(\xi)\widehat{\psi^{j,\ell}}(\xi + 2\pi\mathsf{k}) = 0$, a.e., $\xi \in \mathbb{R}^2$ and $\mathsf{k} \in \mathbb{Z}^2 \backslash \{0\}$ for $|\ell| \le \ell_{\lambda^j} - 1$. For $\widehat{\psi^{j,-\ell_{\lambda^j}}}$ we have

$$\operatorname{supp} \widehat{\psi^{j,-\ell_{\lambda^j}}} \subseteq \{\xi \in \mathbb{R}^2 : \xi_1 \in [-\rho\pi, \rho\pi], -1/2 - \varepsilon \le \xi_2/\xi_1 \le \lambda^j(1 + 2\varepsilon_0/\lambda^{2j}) - \ell_{\lambda^j}\}.$$

Since $2\varepsilon_0 \le \lambda^{J_0}(2/\rho - 1 - 2\varepsilon)$, we have,

$$\begin{aligned}(\lambda^j(1 + 2\varepsilon_0/\lambda^{2j}) - \ell_{\lambda^j} + 1/2 + \varepsilon) &\le (\lambda^j(1 + 2\varepsilon_0/\lambda^{2j}) - (\lambda^j + 1/2 - \varepsilon) + 1 + 1/2 + \varepsilon) \\ &\le \frac{2\varepsilon_0}{\lambda^j} + 1 + 2\varepsilon \le 2/\rho.\end{aligned}$$

This implies the support of $\widehat{\psi^{j,-\ell_{\lambda^j}}}(\xi_1,\cdot)$ is of length less than 2π for any $\xi_1 \in [-\rho\pi, \rho\pi]$. Similar property holds for supp $\widehat{\psi^{j,+\ell_{\lambda^j}}}$. Hence, we conclude that $\widehat{\psi^{j,\pm\ell_{\lambda^j}}}(\xi)\widehat{\psi^{j,\pm\ell_{\lambda^j}}}(\xi+2\pi\mathsf{k}) = 0$, a.e., $\xi \in \mathbb{R}^2$ and $\mathsf{k} \in \mathbb{Z}^2\backslash\{0\}$.

By the definition of φ^j and $\varepsilon_0 \le \lambda^{2J_0}(\frac{\lambda^2}{2\rho} - 1/2)$, we have

$$\operatorname{supp}\widehat{\varphi^j} \subseteq \big[-\lambda^{-2}\rho\big(1+2\varepsilon_0/\lambda^{2j}\big)\pi, \lambda^{-2}\rho\big(1+2\varepsilon_0/\lambda^{2j}\big)\pi\big]^2 \subseteq [-\pi,\pi]^2.$$

Hence, we conclude that $\widehat{\varphi^j}(\xi)\widehat{\varphi^j}(\xi+2\pi\mathsf{k}) = 0$ for all $\mathsf{k} \in \mathbb{Z}^2\backslash\{0\}$ and for almost every $\xi \in \mathbb{R}^2$. Therefore, (3.13) holds. By the result of Corollary 3, $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge J_0$. Since all involved auxiliary functions are from $C_c^\infty(\mathbb{R}^2)$, all elements in $\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty)$ have compactly supported Fourier transforms in $C_c^\infty(\mathbb{R}^2)$. $\square$

From Theorem 3 we see that

$$\boldsymbol{\zeta}\big(S_{-\ell_{\lambda^j}+2}\mathsf{B}_\lambda^j\xi\big) = \boldsymbol{\beta}_{\lambda,t,\rho}\big(\lambda^{-2j}\xi_1\big)\gamma_\varepsilon\big(\lambda^j\xi_2/\xi_1 - \ell_{\lambda^j}+2\big), \quad \xi \in \mathbb{R}^2$$

has support satisfying $\xi_2/\xi_1 \le \frac{\ell_{\lambda^j}-2+1/2+\varepsilon}{\lambda^j} \to 1$ as $j \to \infty$. In other words, the shear operation reaches arbitrarily close to the seamlines $\{\xi \in \mathbb{R}^2 : \xi_2/\xi_1 = \pm 1\}$.

4.3. Quasi-stationary construction

Let us next discuss the quasi-stationary construction. The idea is to use the tensor product of functions in 1D to obtain rectangular bands for different scale levels, and then a frequency splitting using $\boldsymbol{\gamma}_\varepsilon$ is applied to produce generators with respect to different shears. More precisely, let $\lambda > 1$, $0 < t \le 1$, and $0 < \rho \le 1$. Consider $\widehat{\varphi}(\xi) := \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_2)$, $\xi = (\xi_1,\xi_2) \in \mathbb{R}^2$ and define

$$\boldsymbol{\omega}_{\lambda,t,\rho}(\xi) := \sqrt{\big|\widehat{\varphi}\big(\lambda^{-2}\xi\big)\big|^2 - \big|\widehat{\varphi}(\xi)\big|^2}, \quad \xi \in \mathbb{R}^2. \tag{4.23}$$

Then $\boldsymbol{\omega}_{\lambda,t,\rho} \in C_c^\infty(\mathbb{R}^2)$. In fact, it is easy to show that if $\widehat{\varphi}(\xi_0) = 0$ or 1, then all the derivatives of $\widehat{\varphi}$ vanish at ξ_0. Now if $\widetilde{\boldsymbol{\omega}}_{\lambda,t,\rho}(\xi) := |\widehat{\varphi}(\lambda^{-2}\xi)|^2 - |\widehat{\varphi}(\xi)|^2$ does not vanish for $\xi = \xi_0$, then it is trivial to see that $\boldsymbol{\omega}_{\lambda,t,\rho} = \sqrt{\widetilde{\boldsymbol{\omega}}_{\lambda,t,\rho}}$ is infinitely differentiable at $\xi = \xi_0$. If $\widetilde{\boldsymbol{\omega}}_{\lambda,t,\rho}(\xi) = 0$ at $\xi = \xi_0$, then we must have $\widehat{\varphi}(\xi_0) = \widehat{\varphi}(\lambda^{-2}\xi_0) = 0$ or $\widehat{\varphi}(\xi_0) = \widehat{\varphi}(\lambda^{-2}\xi_0) = 1$. Then, all the derivatives of $\widetilde{\boldsymbol{\omega}}_{\lambda,t,\rho}$ vanish at ξ_0. By the Taylor expansion, we see that $\boldsymbol{\omega}_{\lambda,t,\rho} = \sqrt{\widetilde{\boldsymbol{\omega}}_{\lambda,t,\rho}}$ must be infinitely differentiable at ξ_0 with all its derivatives at ξ_0 being zero. Therefore, $\boldsymbol{\omega}_{\lambda,t,\rho} \in C_c^\infty(\mathbb{R}^2)$.

In view of the construction of $\widehat{\varphi}$, the refinable structure is clear. We have $\widehat{\varphi}(\lambda^2\xi) = \widehat{a}(\xi)\widehat{\varphi}(\xi)$, $\xi \in \mathbb{R}^2$ with $\widehat{a} = \boldsymbol{\mu}_{\lambda,t,\rho} \otimes \boldsymbol{\mu}_{\lambda,t,\rho}$ being the tensor product of the 1D mask $\boldsymbol{\mu}_{\lambda,t,\rho}$ given in (4.3). Moreover, we have $\boldsymbol{\omega}(\lambda^2\xi) = \widehat{b}(\xi)\widehat{\varphi}(\xi)$ with $\widehat{b} \in C^\infty(\mathbb{T})$ being given by $\widehat{b}(\xi) = (\mathbf{g}(\xi) - |\widehat{a}(\xi)|^2)^{1/2}$ for any smooth function $\mathbf{g} \in C^\infty(\mathbb{T}^2)$ such that $\mathbf{g} = 1$ on the support of $\widehat{\varphi}$.

Note that for simplicity of presentation, we omit the dependency of $\varphi, \psi^{j,\ell}, a, b, \boldsymbol{\Gamma}_j$, etc., on the parameters $\lambda, t, \rho, \varepsilon$.

Since $0 < \boldsymbol{\Gamma}_j \le 2$ and $\boldsymbol{\Gamma}_j$ is in $C^\infty(\mathbb{R}^2\backslash\{0\})$, we have that $\sqrt{\boldsymbol{\Gamma}_j}$ is infinitely differentiable for all $\xi \in \mathbb{R}^2\backslash\{0\}$. Let $\mathsf{A}_\lambda, \mathsf{B}_\lambda, \mathsf{M}_\lambda, \mathsf{N}_\lambda, \mathsf{D}_\lambda$ with $\lambda > 1$ be defined as in (1.1) and (3.1). Let $\Psi_j := \{\psi^{j,\ell}(S^{-\ell}\cdot) : \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}\}$ with

$$\widehat{\psi^{j,\ell}}(\xi) := \boldsymbol{\omega}_{\lambda,t,\rho}\big(\mathsf{D}_\lambda^{-j}S_{-\ell}\xi\big)\frac{\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1)}{\sqrt{\boldsymbol{\Gamma}_j((S_\ell\mathsf{B}_\lambda^j)^{-1}\xi)}} = \boldsymbol{\omega}_{\lambda,t,\rho}\big(\xi_1, \lambda^{-j}(-\xi_1\ell+\xi_2)\big)\frac{\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1)}{\sqrt{\boldsymbol{\Gamma}_j((S_\ell\mathsf{B}_\lambda^j)^{-1}\xi)}} \tag{4.24}$$

for $\xi \in \mathbb{R}^2 \backslash \{0\}$ and $\widehat{\psi^{j,\ell}}(0) := 0$, which gives $\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j \xi) = \boldsymbol{\omega}_{\lambda,t,\rho}(\mathsf{N}_\lambda^j \xi) \frac{\gamma_\varepsilon(\lambda^j \xi_2/\xi_1 + \ell)}{\sqrt{\boldsymbol{\Gamma}_j(\xi)}}$. By the properties of $\boldsymbol{\Gamma}_j$ and that the support of $\boldsymbol{\omega}_{\lambda,t,\rho}$ is away from the origin, we see that $\widehat{\psi^{j,\ell}}$ are compactly supported functions in $C_c^\infty(\mathbb{R}^2)$ and hence $\psi^{j,\ell} \in C^\infty(\mathbb{R}^2)$. We now define a (quasi-stationary) *affine shear system*:

$$\begin{aligned} \mathsf{AS}_J\big(\varphi; \{\Psi_j\}_{j=J}^\infty\big) &:= \big\{\varphi_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\big\} \cup \big\{h_{\mathsf{A}_\lambda^j;\mathsf{k}}, h_{\mathsf{A}_\lambda^j \mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\big\}_{j=J}^\infty \\ &= \big\{\varphi_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\big\} \cup \big\{\psi^{j,\ell}_{S^{-\ell}\mathsf{A}_\lambda^j;\mathsf{k}}, \psi^{j,\ell}_{S^{-\ell}\mathsf{A}_\lambda^j \mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\big\}_{j=J}^\infty. \end{aligned} \tag{4.25}$$

At first glance, such a system does not have shear structure at all due to that the function $\boldsymbol{\omega}_{\lambda,t,\rho}$ is not shear-invariant. However, we shall show that such a system do have certain affine and shear structure in the sense that a sub-system of this system is from shear and dilation of one single generator.

Theorem 4. *Let $\lambda > 1$, $0 < t \le 1$, and $0 < \rho \le 1$. Let $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ be defined as in (4.25) with $\widehat{\varphi} = \boldsymbol{\alpha}_{\lambda,t,\rho} \otimes \boldsymbol{\alpha}_{\lambda,t,\rho}$ and $\psi^{j,\ell}$ being given by (4.24). Then $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge 0$. All elements in $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ have compactly supported Fourier transforms in $C_c^\infty(\mathbb{R}^2)$. Moreover, we have*

$$\big\{\psi\big(S^{-\ell}\cdot\big) : \ell = -r_j, \dots, r_j\big\} \subseteq \Psi_j, \quad j \ge J,$$

where $r_j := \lfloor \lambda^{j-2}(1-t)\rho - (1/2 + \varepsilon)\rfloor$ and $\widehat{\psi}(\xi) := \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\gamma_\varepsilon(\xi_2/\xi_1)$, $\xi \in \mathbb{R}^2$. In other words,

$$\big\{\psi_{S^{-\ell}\mathsf{A}_\lambda^j;\mathsf{k}}, \psi_{S^{-\ell}\mathsf{A}_\lambda^j \mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell = -r_j, \dots, r_j\big\}_{j=J}^\infty \subseteq \underset{J}{\mathsf{AS}}\big(\varphi; \{\Psi_j\}_{j=J}^\infty\big).$$

Proof. By our construction, we have

$$\begin{aligned} &\big|\widehat{\varphi}\big(\mathsf{N}_\lambda^j \xi\big)\big|^2 + \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \big[\big|\widehat{\psi^{j,\ell}}\big(S_\ell \mathsf{B}_\lambda^j \xi\big)\big|^2 + \big|\widehat{\psi^{j,\ell}}\big(S_\ell \mathsf{B}_\lambda^j \mathsf{E}\xi\big)\big|^2\big] \\ &= \big|\widehat{\varphi}\big(\mathsf{N}_\lambda^j \xi\big)\big|^2 + \frac{|\boldsymbol{\omega}_{\lambda,t,\rho}(\mathsf{N}^j \xi)|^2}{\boldsymbol{\Gamma}_j(\xi)} \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \big[\big|\gamma_\varepsilon\big(\lambda^j \xi_2/\xi_1 + \ell\big)\big|^2 + \big|\gamma_\varepsilon\big(\lambda^j \xi_1/\xi_2 + \ell\big)\big|^2\big] \\ &= \big|\widehat{\varphi}\big(\mathsf{N}_\lambda^j \xi\big)\big|^2 + \big|\boldsymbol{\omega}_{\lambda,t,\rho}\big(\mathsf{N}^j \xi\big)\big|^2 = \big|\widehat{\varphi}\big(\mathsf{N}^{j+1} \xi\big)\big|^2, \quad \xi \in \mathbb{R}^2. \end{aligned}$$

Hence, (3.14) holds. By the definition of φ, (3.6) also holds. Note that all generators satisfy $\widehat{\psi^{j,\ell}} \ge 0$ and $\operatorname{supp} \widehat{\psi^{j,\ell}} \subseteq [-\rho\pi, \rho\pi]^2$ with $\rho \le 1$. Hence, (3.13) is true. Now, by Corollary 3, we conclude that $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge 0$. Since all $\widehat{\varphi}, \widehat{\psi^{j,\ell}}$ are compactly supported functions in $C_c^\infty(\mathbb{R}^2)$, all elements in $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ are functions in $C^\infty(\mathbb{R}^2)$.

By the definition of $\boldsymbol{\omega}_{\lambda,t,\rho}$, it is easy to see that

$$\big|\boldsymbol{\omega}_{\lambda,t,\rho}(\xi_1, \xi_2)\big|^2 = \big|\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_2)\big|^2 + \big|\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_2)\big|^2 + \big|\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_2)\big|^2.$$

And for $|\xi_2| \le \lambda^{-2}(1-t)\rho\pi$, we have $\boldsymbol{\omega}_{\lambda,t,\rho}(\xi_1, \xi_2) = \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_2) = \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)$. Consequently, if for all $\xi \in \operatorname{supp} \widehat{\psi^{j,\ell}}_{S_\ell \mathsf{B}_\lambda^j;0,\mathsf{k}}$, we have $|\xi_2| \le \lambda^{2j-2}(1-t)\rho\pi$, then we have

$$\begin{aligned} \widehat{\psi^{j,\ell}}_{S_\ell \mathsf{B}_\lambda^j;0,\mathsf{k}}(\xi) &= \lambda^{-3j/2}\boldsymbol{\omega}_{\lambda,t,\rho}\big(\lambda^{-2j}\xi\big)\gamma_\varepsilon\big(\lambda^j \xi_2/\xi_1 + \ell\big)e^{-i\mathsf{k}\cdot S_\ell \mathsf{B}_\lambda^j \xi} \\ &= \lambda^{-3j/2}\boldsymbol{\beta}_{\lambda,t,\rho}\big(\lambda^{-2j}\xi_1\big)\gamma_\varepsilon\big(\lambda^j \xi_2/\xi_1 + \ell\big)e^{-i\mathsf{k}\cdot S_\ell \mathsf{B}_\lambda^j \xi} \\ &= \widehat{\psi}_{S_\ell \mathsf{B}_\lambda^j;0,\mathsf{k}}(\xi). \end{aligned}$$

Now let us find the range of ℓ such that the above support constrain for $\widehat{\psi^{j,\ell}}_{S_\ell \mathsf{B}^j_\lambda;0,\mathsf{k}}$ holds. At the scale level j, we have

$$\operatorname{supp} \boldsymbol{\omega}_{\lambda,t,\rho}(\lambda^{-2j}\cdot) \subseteq [-\lambda^{2j}\rho\pi, \lambda^{2j}\rho\pi]^2 \backslash [-\lambda^{2j-2}(1-t)\rho\pi, \lambda^{2j-2}(1-t)\rho\pi]^2.$$

Then, the support constrain means that at the scale level j, one needs $|\xi_2/\xi_1| \le \lambda^{-2}(1-t)\rho$. Hence, the support of $\boldsymbol{\gamma}_\varepsilon(\lambda^j\xi_2/\xi_1 + \ell)$ must satisfy

$$-\lambda^{-2}(1-t)\rho \le -\lambda^{-j}(1/2+\varepsilon+\ell) \le \xi_2/\xi_1 \le \lambda^{-j}(1/2+\varepsilon-\ell) \le \lambda^{-2}(1-t)\rho.$$

Consequently, we obtain

$$-\lambda^{j-2}(1-t)\rho + (1/2+\varepsilon) \le \ell \le \lambda^{j-2}(1-t)\rho - (1/2+\varepsilon).$$

That is, $|\ell| \le \lambda^{j-2}(1-t)\rho - (1/2+\varepsilon)$. In summary, letting $r_j := \lfloor \lambda^{j-2}(1-t)\rho - (1/2+\varepsilon)\rfloor$, we have

$$\{\psi(S^{-\ell}\cdot) : \ell = -r_j, \dots, r_j\} \subseteq \Psi_j, \quad j \ge J,$$

and

$$\{\psi_{S^{-\ell}\mathsf{A}^j_\lambda;\mathsf{k}}, \psi_{S^{-\ell}\mathsf{A}^j_\lambda \mathsf{E};\mathsf{k}} : j \ge J, \mathsf{k} \in \mathbb{Z}^2, \ell = -r_j, \dots, r_j\} \subseteq \mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty).$$

This completes the proof. □

Note that when $\ell = -r_j$, the support of $\widehat{\psi}(S_\ell \mathsf{B}^j_\lambda \xi) = \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\gamma_\varepsilon(\lambda^j\xi_2/\xi_1 - r_j)$ satisfies

$$\xi_2/\xi_1 \le \lambda^{-j}(r_j + 1/2+\varepsilon) \le \lambda^{-j}\left(\lfloor \lambda^{j-2}(1-t)\rho - 1/2 - \varepsilon\rfloor + 1/2+\varepsilon\right) \le \lambda^{-2}(1-t)\rho.$$

Hence, by the symmetry property of $\boldsymbol{\Gamma}_j$, we see that the shear operation generates a subsystem of $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=0}^\infty)$ inside the cone area $\{\xi \in \mathbb{R}^2 : \max\{|\xi_2/\xi_1|, |\xi_1/\xi_2|\} \le \lambda^{-2}(1-t)\rho\}$ in the frequency domain.

4.4. Connections to other directional mutliscale representation systems

In this subsection, we shall discuss the connections of our affine shear tight frames to those shearlet systems in [8,10] or shearlet-like systems in [13].

Define corner pieces

$$\boldsymbol{\gamma}^+_\lambda(x) := \begin{cases} \boldsymbol{\gamma}_\varepsilon(x) & \text{if } -1/2-\varepsilon \le x \le -1/2+\varepsilon, \\ 1 & \text{if } -1/2+\varepsilon \le x \le \lambda - \ell_\lambda, \\ 0 & \text{otherwise,} \end{cases}$$
$$\boldsymbol{\gamma}^-_\lambda(x) := \begin{cases} \boldsymbol{\gamma}_\varepsilon(x) & \text{if } 1/2-\varepsilon \le x \le 1/2+\varepsilon, \\ 1 & \text{if } -\lambda + \ell_\lambda \le x \le 1/2-\varepsilon, \\ 0 & \text{otherwise.} \end{cases} \tag{4.26}$$

These are the corner pieces that shall be used to achieve tightness of the system or for gluing two seamline elements together smoothly. Let $\{\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho}, \boldsymbol{\gamma}_\varepsilon, \boldsymbol{\gamma}^\pm_\lambda\}$ be defined as in (4.2), (4.4), and (4.26). Similarly, for the half pieces of the system generated by the characteristic functions as in (2.9), we define $\psi, \psi^{j,\pm\ell_{\lambda^j}}$ by

$$\widehat{\psi}(\xi) := \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1), \qquad \widehat{\psi^{j,\pm\ell_{\lambda^j}}}(\xi) := \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\gamma^{\mp}_{\lambda^j}(\xi_2/\xi_1), \quad \xi \neq 0$$

and $\widehat{\psi}(0) := 0$, $\widehat{\psi^{j,\pm\ell_{\lambda_j}}}(0) := 0$. The scaling function φ is defined to be

$$\varphi := \varphi^1 + \varphi^2 \tag{4.27}$$

with $\widehat{\varphi^1}(\xi) = \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_1)\chi_{\{\xi\in\mathbb{R}^2:|\xi_2/\xi_1|\le 1\}}(\xi)$ and $\widehat{\varphi^2} = \widehat{\varphi^1}(\mathsf{E}\cdot) = \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_2)\chi_{\{\xi\in\mathbb{R}^2:|\xi_1/\xi_2|\le 1\}}(\xi)$, $\xi = (\xi_1,\xi_2) \in \mathbb{R}^2$. Now define

$$\Psi_j := \{\psi(S^{-\ell}\cdot) : \ell = -\ell_{\lambda^j}+1,\dots,\ell_{\lambda^j}-1\} \cup \{\psi^{j,\ell}(S^{-\ell}\cdot) : \ell = \pm\ell_{\lambda^j}\}. \tag{4.28}$$

Note that $\widehat{\psi}$ is smooth while the corner pieces $\widehat{\psi^{j,\pm\ell_{\lambda^j}}}$ are not smooth. We have the following result.

Corollary 4. *Let $\mathsf{A}_\lambda, \mathsf{B}_\lambda, \mathsf{M}_\lambda, \mathsf{N}_\lambda, S_\ell, \mathsf{E}$ be defined as in (1.1) and (3.1) with $\lambda > 1$. Let $0 < t \le 1$, $0 < \rho \le 1$ and $0 < \varepsilon \le 1/2$. Then the system $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ defined as in (3.2) with φ, Ψ_j being given by (4.27), (4.28), respectively, is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge 0$.*

Proof. By the definition of $\boldsymbol{\gamma}_\varepsilon$ and $\boldsymbol{\gamma}^{\pm}_\lambda$, for a fixed $j \ge 0$, it is easy to show that

$$\sum_{\ell=-\ell_{\lambda^j}+1}^{\ell_{\lambda^j}-1} |\boldsymbol{\gamma}_\varepsilon(\lambda^j\xi_2/\xi_1+\ell)|^2 + |\boldsymbol{\gamma}^+_{\lambda^j}(\lambda^j\xi_2/\xi_1-\ell_{\lambda^j})|^2 + |\boldsymbol{\gamma}^-_{\lambda^j}(\lambda^j\xi_2/\xi_1+\ell_{\lambda^j})|^2 = \chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi), \quad \xi \neq 0.$$

Hence, for $\xi = (\xi_1,\xi_2) \in \mathbb{R}^2$, we have

$$\begin{aligned}|\widehat{\varphi^1}(\mathsf{N}^j_\lambda\xi)|^2 + \sum_{h\in\Psi_j}|\widehat{h}(\mathsf{B}^j_\lambda\xi)|^2 &= \left(|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)|^2 + |\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)|^2\right)\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi)\\ &= |\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2j-2}\xi_1)|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi)\\ &= |\widehat{\varphi^1}(\mathsf{N}^{j+1}_\lambda\xi)|^2.\end{aligned}$$

Similarly, we have $|\widehat{\varphi^2}(\mathsf{N}^j_\lambda\xi)|^2 + \sum_{h\in\Psi_j}|\widehat{h}(\mathsf{B}^j_\lambda\mathsf{E}\xi)|^2 = |\widehat{\varphi^2}(\mathsf{N}^{j+1}_\lambda\xi)|^2$. Consequently, we have

$$|\widehat{\varphi}(\mathsf{N}^j_\lambda\xi)|^2 + \sum_{h\in\Psi_j}\left(|\widehat{h}(\mathsf{B}^j_\lambda\xi)|^2 + |\widehat{h}(\mathsf{B}^j_\lambda\mathsf{E}\xi)|^2\right) = |\widehat{\varphi}(\mathsf{N}^{j+1}_\lambda\xi)|^2, \quad \text{a.e. } \xi \in \mathbb{R}^2.$$

Hence (3.14) holds.

Moreover, we have $\widehat{h}(\xi)\overline{\widehat{h}(\xi+2\pi\mathsf{k})} = 0$ for all $h \in \{\varphi\}\cup\{\Psi_j\}_{j=0}^\infty$ and $\mathsf{k} \in \mathbb{Z}^2\backslash\{0\}$. In fact, if $\mathsf{k} = (k_1,k_2) \in \mathbb{Z}^2$ with $k_1 \neq 0$, then $\widehat{h}(\xi)\overline{\widehat{h}(\xi+2\pi\mathsf{k})} = 0$ due to that $\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho}$ are supported on $[-\rho\pi,\rho\pi]$ with $\rho \le 1$. If $k_1 = 0$ but $k_2 \neq 0$, then by $\boldsymbol{\gamma}_\varepsilon((\xi_2+2\pi k_2)/\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1) = \boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1+2\pi k_2/\xi_1)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1) = 0$ for $\xi_1 \in [-\rho\pi,\rho\pi]$, we have $\widehat{h}(\xi)\overline{\widehat{h}(\xi+2\pi\mathsf{k})} = 0$ as well. Hence, (3.13) is satisfied. Obviously, (3.6) is true by our construction of φ.

Therefore, by Corollary 3, $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$ defined in (1.3) with φ, Ψ_j being given by (4.27), (4.28), respectively, is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for all $J \ge 0$. $\square$

Now, it is easy to show that the cone-adapted shearlet system constructed in [10] is indeed the initial system of a sequence of affine shear tight frames. In fact, let $\lambda = 2$, and $\mathsf{A}_1 := \mathsf{A}_\lambda$, $\mathsf{A}_2 := \mathsf{E}\mathsf{A}_1\mathsf{E}$. Let $\psi^1 = \psi$ and $\psi^2 := \psi^1(\mathsf{E}\cdot)$. It is easy to show that

$$\psi^1(S^\ell \mathsf{A}_1^j \cdot -\mathsf{k}) = \psi(S_\ell \mathsf{A}_\lambda^j \cdot -\mathsf{k}) \quad \text{and} \quad \psi^2(S_\ell \mathsf{A}_2^j \cdot -\mathsf{k}) = \psi(S^\ell \mathsf{A}_\lambda^j \mathsf{E} \cdot -\mathsf{E}\mathsf{k}).$$

Noting that $\mathsf{E}\mathbb{Z}^2 = \mathbb{Z}^2$ and the symmetry of the range of ℓ for each scale level j, we see that the cone-adapted shearlet system in (1.4) with modified seamline elements is the affine shear tight frame $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^\infty)$ defined as in (1.3) with φ, Ψ_j being given by (4.27), (4.28), and $\lambda = 2$. Moreover, it is the initial system of the sequence of affine shear tight frames $\mathsf{AS}_J(\varphi; \{\Psi_j\}_{j=J}^\infty)$, $J \in \mathbb{N}_0$ defined as in (3.2) with φ, Ψ_j being given by (4.27), (4.28), respectively.

For the smooth shearlet-like systems constructed in [13], it is also a special case of the following system. Note that $\boldsymbol{\gamma}_\lambda^+, \boldsymbol{\gamma}_\lambda^-$ satisfy

$$\left[\frac{d^n}{dx^n}\boldsymbol{\gamma}_\lambda^\pm(\lambda x \mp \ell_\lambda)\right]\Big|_{x=\pm 1} = \delta(n) \quad \forall n \in \mathbb{N}_0, \tag{4.29}$$

which guarantees the smoothness by gluing the two corner pieces.

Let $\Psi_j := \{\psi^{j,\ell}(S^{-\ell}\cdot) : \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\}$, where for $|\ell| < \ell_{\lambda_j}$,

$$\widehat{\psi^{j,\ell}}(\xi) := \boldsymbol{\omega}_{\lambda,t,\rho}(\mathsf{D}_\lambda^{-j} S_{-\ell}\xi)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1) = \boldsymbol{\omega}_{\lambda,t,\rho}(\xi_1, \lambda^{-j}(-\xi_1\ell + \xi_2))\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1), \quad \xi \in \mathbb{R}^2, \tag{4.30}$$

which gives

$$\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j \xi) = \boldsymbol{\omega}_{\lambda,t,\rho}(\lambda^{-2j}\xi)\boldsymbol{\gamma}_\varepsilon(\lambda^j \xi_2/\xi_1 + \ell);$$

and for those elements on the seamlines, i.e., for $\ell = \pm\ell_{\lambda^j}$ and $j \ge 1$,

$$\widehat{\psi^{j,\pm\ell_{\lambda^j}}}(S_{\pm\ell_{\lambda^j}} \mathsf{B}_\lambda^j/2\xi) := \begin{cases} \boldsymbol{\omega}_{\lambda,t,\rho}(\lambda^{-2j}\xi)\boldsymbol{\gamma}_{\lambda^j}^\mp(\lambda^j\xi_2/\xi_1 \pm \ell_{\lambda^j}) & \text{if } |\xi_2/\xi_1| \le 1, \\ \boldsymbol{\omega}_{\lambda,t,\rho}(\lambda^{-2j}\xi)\boldsymbol{\gamma}_{\lambda^j}^\mp(\lambda^j\xi_1/\xi_2 \pm \ell_{\lambda^j}) & \text{if } |\xi_2/\xi_1| \ge 1. \end{cases}$$

For $j = 0$,

$$\widehat{\psi^{0,\pm 1}}(S_{\pm 1}\xi) := \begin{cases} \boldsymbol{\omega}_{\lambda,t,\rho}(\xi)\boldsymbol{\gamma}_\varepsilon(\xi_2/\xi_1 \pm 1) & \text{if } |\xi_2/\xi_1| \le 1, \\ \boldsymbol{\omega}_{\lambda,t,\rho}(\xi)\boldsymbol{\gamma}_\varepsilon(\xi_1/\xi_2 \pm 1) & \text{if } |\xi_2/\xi_1| \ge 1. \end{cases}$$

Let $\mathsf{A}_\lambda^{j,\ell} := \mathsf{A}_\lambda^j$ for $j \ge 1$ and $\ell < \ell_{\lambda^j}$, $\mathsf{A}_\lambda^{j,\pm\ell_{\lambda^j}} := 2\mathsf{A}_\lambda^j$ for $j \ge 1$, and for $j = 0$, $\mathsf{A}_\lambda^{j,\ell} := \mathsf{I}_2$. Then, we can define the following system:

$$\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^\infty) = \{\varphi(\cdot - \mathsf{k}) : \mathsf{k} \in \mathbb{Z}^2\} \cup \{h_{\mathsf{A}_\lambda^{j,\ell};\mathsf{k}}, h_{\mathsf{A}_\lambda^{j,\ell}\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \Psi_j\}_{j=0}^\infty \tag{4.31}$$

Corollary 5. *$\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^\infty)$ in (4.31) is an affine shear tight frame for $L_2(\mathbb{R}^2)$ and all elements in $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^\infty)$ have compactly supported Fourier transforms in $C_c^\infty(\mathbb{R}^2)$.*

Proof. By our construction, we have

$$\begin{aligned} &|\widehat{\varphi}(\cdot)|^2 + \sum_{j=0}^\infty \sum_{\ell=-\ell_{\lambda^j}+1}^{\ell_{\lambda^j}-1} [|\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j \cdot)|^2 + |\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j \mathsf{E}\cdot)|^2] \\ &+ \sum_{j=0}^\infty \sum_{\ell=\pm\ell_{\lambda^j}} |\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j/2\cdot)|^2 + |\widehat{\psi^{j,\ell}}(S_\ell \mathsf{B}_\lambda^j \mathsf{E}/2\cdot)|^2 = 1, \quad \text{a.e. } \xi \in \mathbb{R}^2. \end{aligned}$$

Moreover, all generators satisfy $\widehat{\psi^{j,\ell}} \ge 0$ and $\operatorname{supp} \widehat{\psi^{j,\ell}} \subseteq [-\pi, \pi]^2$. Note that dilation matrices of the seamline generators $\psi^{j,\pm\ell_{\lambda^j}}$ are $2\mathsf{A}_\lambda^j$ instead of A_λ^j. A simple adaptation of the proof of Theorem 1 gives that $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^\infty)$ is a tight frame for $L_2(\mathbb{R}^2)$. By the definition of $\boldsymbol{\gamma}, \boldsymbol{\gamma}_\lambda^\pm$ in (4.4), (4.26), $\widehat{\psi^{j,\ell}}$ are compactly supported smooth functions. Consequently, all elements in $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^\infty)$ have compactly supported Fourier transforms in $C_c^\infty(\mathbb{R}^2)$. $\square$

We finish this section by making some comments on the connections and differences of our affine shear systems with other shearlet or shearlet-like systems. First, when $\lambda = 2, t = 1-\lambda^{-2}$, and $\rho = 1$, except those seamline elements, $\mathsf{AS}(\varphi;\{\Psi_j\}_{j=0}^\infty)$ defined in (4.31) is essentially the system defined in [13]. Second, the shear subsystem (generated by one single generator through shear, parabolic scaling, and translation) in [13] can have its shear operations reach only up to slope (in absolute value) $\lambda^{-4} = 1/16$. Here, in our construction, the shear subsystem $\{\psi_{S^{-\ell}\mathsf{A}_\lambda^j;\mathsf{k}}, \psi_{S^{-\ell}\mathsf{A}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell = -r_j, \dots, r_j\}_{j=J}^\infty$ as in Theorem 4 can reach up to slope $\lambda^{-2}(1-t)\rho$ in the frequency domain with any $0 < t \le 1, 0 < \rho \le 1$. In other words, we have a shear subsystem covers larger cones (horizontal and vertical) in the frequency domain than those in [13]. Third, the ideas of achieving tightness for our quasi-stationary construction and the construction in [13] are essentially different. Our tightness is achieved by normalizing an affine frame obtained through application of shear, dilation, translation, together with flip operations to a single generator while the tightness in [13] is achieved by a gluing process. Comparing with our quasi-stationary construction, the gluing procedure is somewhat unnatural since one can see that a different dilation matrix $2\mathsf{A}_\lambda^j$ needs to be applied to the gluing elements at the scale level j while all other generators use the dilation matrix A_λ^j (see boundary shearlets in Section 2.1 in [13]). Our affine shear systems, either under our quasi-stationary construction or non-stationary construction, obey the parabolic rule. More importantly, at all scale levels j for each cone, the dilation matrix is fixed as A_λ^j for all generators. Finally, we would like to point out that our quasi-stationary construction is more general and flexible with several control parameters $\lambda, t, \rho, \varepsilon$. Moreover, to our best knowledge, the non-stationary construction in this paper is new and we make a clear and important connection between affine shear tight frames and directional tight framelets, which shall be investigated in the next section.

5. MRA structures and filter banks

By connecting affine shear tight frames to tight framelets in [16], in this section we shall study the MRA structure of sequences of affine shear tight frames constructed in Section 4 and investigate their underlying filter banks.

As discussed in [16], a sequence of tight framelets is closely linked to filter banks and MRA structure. Let us first discuss the connections of affine shear tight frames to a special class of tight framelets.

5.1. Connections to affine tight M_λ-framelets through subsampling

Let $\{\varphi^J\} \cup \{\mathring{\Psi}_j\}_{j=J}^\infty$ be a set of generators with

$$\mathring{\Psi}_j := \{\mathring{\psi}^{j,\ell} : \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\}. \tag{5.1}$$

Using the dilation matrix M_λ for all generators in $\{\varphi^J\} \cup \{\mathring{\psi}^{j,\ell}\}_{j=J}^\infty$ with $J \in \mathbb{N}_0$, we define a sequence of (non-stationary cone-adapted) *affine* M_λ*-framelet systems* by

$$\begin{aligned}\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty) &:= \{\varphi^J_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\} \cup \{h_{\mathsf{M}_\lambda^j;\mathsf{k}}, h_{\mathsf{M}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, h \in \mathring{\Psi}_j\}_{j=J}^\infty \\ &= \{\varphi^J_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2\} \cup \{\mathring{\psi}^{j,\ell}_{\mathsf{M}_\lambda^j;\mathsf{k}}, \mathring{\psi}^{j,\ell}_{\mathsf{M}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^2, \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\}_{j=J}^\infty.\end{aligned} \tag{5.2}$$

Similarly, when $\varphi^j = \varphi$ is fixed across all scale levels $j \in \mathbb{N}_0$, we define a sequence of (quasi-stationary cone-adapted) *affine* M_λ*-framelet systems* by

$$\begin{aligned}\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty) &= \{\varphi_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2\} \cup \{h_{\mathsf{M}_\lambda^j;\mathsf{k}}, h_{\mathsf{M}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2, h\in\mathring{\Psi}_j\}_{j=J}^\infty \\ &= \{\varphi_{\mathsf{M}_\lambda^J;\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2\} \cup \{\mathring{\psi}^{j,\ell}_{\mathsf{M}_\lambda^j;\mathsf{k}}, \mathring{\psi}^{j,\ell}_{\mathsf{M}_\lambda^j\mathsf{E};\mathsf{k}} : \mathsf{k}\in\mathbb{Z}^2, \ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}\}_{j=J}^\infty. \end{aligned} \tag{5.3}$$

We have the following result connecting affine shear systems with affine M_λ-framelet systems.

Theorem 5. *Let $\mathsf{M}_\lambda, \mathsf{N}_\lambda, \mathsf{D}_\lambda$ be defined as in (3.1). Let $\{\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)\}_{J=0}^\infty$ be a sequence of affine shear systems in (4.21) with $\Psi_j = \{\psi^{j,\ell}(S^{-\ell}\cdot) : \ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}\}$. Define $\mathring{\Psi}_j := \{\mathring{\psi}^{j,\ell} : \ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}\}$ with $j\in\mathbb{N}_0$ and*

$$\mathring{\psi}^{j,\ell} := \lambda^{-j}\psi^{j,\ell}(S^{-\ell}\mathsf{D}_\lambda^{-j}\cdot), \quad \ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}. \tag{5.4}$$

If

$$\widehat{\varphi^j}(\xi)\widehat{\varphi^j}(\xi+2\pi\mathsf{k}) = 0, \quad a.e.\ \xi\in\mathbb{R}^2,\ \forall\mathsf{k}\in\mathbb{Z}^2\backslash\{0\},\ j\in\mathbb{N}_0, \tag{5.5}$$

$$\widehat{\psi^{j,\ell}}(\xi)\widehat{\psi^{j,\ell}}(\xi+2\pi\mathsf{k}) = 0, \quad a.e.\ \xi\in\mathbb{R}^2,\ \forall\mathsf{k}\in\mathbb{Z}^2\backslash\{0\},\ j\in\mathbb{N}_0,\ |\ell|\le\ell_{\lambda^j}, \tag{5.6}$$

and

$$\widehat{\mathring{\psi}^{j,\ell}}(\xi)\widehat{\mathring{\psi}^{j,\ell}}(\xi+2\pi\mathsf{k}) = 0, \quad a.e.\ \xi\in\mathbb{R}^2,\ \forall\mathsf{k}\in\mathbb{Z}^2\backslash\{0\},\ j\in\mathbb{N}_0,\ |\ell|\le\ell_{\lambda^j}, \tag{5.7}$$

then $\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for every $J\in\mathbb{N}_0$ if and only if $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ in (5.2) is an affine tight M_λ-framelet for $L_2(\mathbb{R}^2)$ for every $J\in\mathbb{N}_0$, that is, $\{\varphi^j : j\in\mathbb{N}_0\}\cup\{\mathring{\Psi}_j\}_{j=0}^\infty\subseteq L_2(\mathbb{R}^2)$ and for every integer $J\in\mathbb{N}_0$,

$$\|f\|_2^2 = \sum_{\mathsf{k}\in\mathbb{Z}^2}|\langle f,\varphi^J_{\mathsf{M}_\lambda^J;\mathsf{k}}\rangle|^2 + \sum_{j=J}^\infty\sum_{h\in\mathring{\Psi}_j}\sum_{\mathsf{k}\in\mathbb{Z}^2}\left(|\langle f,h_{\mathsf{M}_\lambda^j;\mathsf{k}}\rangle|^2 + |\langle f,h_{\mathsf{M}_\lambda^j\mathsf{E};\mathsf{k}}\rangle|^2\right) \quad \forall f\in L_2(\mathbb{R}^2). \tag{5.8}$$

Proof. Since (5.5) and (5.6) are satisfied, by Theorem 2 (also cf. Corollary 3), $\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^2)$ for every $J\in\mathbb{N}_0$ if and only if (3.6) and (3.14) are satisfied with $J_0=0$. Observe that (5.4) implies $\widehat{\mathring{\psi}^{j,\ell}} = \widehat{\psi^{j,\ell}}(S_\ell\mathsf{D}_\lambda^j\cdot)$. Therefore, by $\mathsf{B}_\lambda^j\mathsf{M}_\lambda^j = \mathsf{D}_\lambda^j$ and $\mathsf{E}\mathsf{M}_\lambda^j = \mathsf{M}_\lambda^j\mathsf{E}$, we see that (3.14) is equivalent to

$$\begin{aligned}|\widehat{\varphi^{j+1}}(\mathsf{N}_\lambda\xi)|^2 &= |\widehat{\varphi^{j+1}}(\xi)|^2 + \sum_{\ell=-s_j}^{s_j}\left(|\widehat{\psi^{j,\ell}}(S_\ell\mathsf{B}_\lambda^j\mathsf{M}_\lambda^j\xi)|^2 + |\widehat{\psi^{j,\ell}}(S_\ell\mathsf{B}_\lambda^j\mathsf{E}\mathsf{M}_\lambda^j\xi)|^2\right) \\ &= |\widehat{\varphi^{j+1}}(\xi)|^2 + \sum_{\ell=-s_j}^{s_j}\left(|\widehat{\mathring{\psi}^{j,\ell}}(\xi)|^2 + |\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{E}\xi)|^2\right)\end{aligned}$$

a.e. $\xi\in\mathbb{R}^2$ and $j\in\mathbb{N}_0$. Hence, by (5.5) and (5.7), the claim follows directly from Theorem 2 and [16, Corollary 18]. □

Immediately, we have the following corollary.

Corollary 6. *Retain all the conditions on $\lambda, t, \rho, \varepsilon, \varepsilon_0$ for $\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)$ in (4.21) of Theorem 3 with $J_0=0$ (respectively, for $\mathsf{AS}_J(\varphi;\{\Psi_j\}_{j=J}^\infty)$ in (4.25) of Theorem 4). Let $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ be defined as in (5.2) (respectively, $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ be defined in (5.3)) with $\mathring{\Psi}_j$ being given as in (5.4). Then $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ (respectively, $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$) is an affine tight M_λ-framelet for $L_2(\mathbb{R}^2)$ for every integer $J\in\mathbb{N}_0$ and*

$$\psi^{j,\ell}_{S^{-\ell}\mathsf{A}^j_\lambda;\mathsf{k}} = \lambda^{j/2}\mathring{\psi}^{j,\ell}_{\mathsf{M}^j_\lambda;\mathsf{D}^j_\lambda S^\ell \mathsf{k}}, \qquad \psi^{j,\ell}_{S^{-\ell}\mathsf{A}^j_\lambda\mathsf{E};\mathsf{k}} = \lambda^{j/2}\mathring{\psi}^{j,\ell}_{\mathsf{M}^j_\lambda\mathsf{E};\mathsf{D}^j_\lambda S^\ell \mathsf{k}}. \tag{5.9}$$

Proof. By (5.4), it is straightforward to check that (5.9) holds. By (5.4), we also have

$$\widehat{\mathring{\psi}^{j,\ell}}(\xi) := \boldsymbol{\omega}^j_{\lambda,t,\rho}(\xi)\frac{\gamma_\varepsilon(\lambda^j\xi_2/\xi_1+\ell)}{\boldsymbol{\Gamma}^j(\xi)}, \quad \xi\neq 0,\ |\ell|\le \ell_{\lambda^j}-1, \tag{5.10}$$

$$\widehat{\mathring{\psi}^{j,\pm\ell_{\lambda^j}}}(\xi) := \boldsymbol{\omega}^j_{\lambda,t,\rho}(\xi)\frac{\gamma^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}(\lambda^j\xi_2/\xi_1\pm\ell_{\lambda^j})}{\boldsymbol{\Gamma}^j(\xi)}, \quad \xi\neq 0, \tag{5.11}$$

and $\widehat{\mathring{\psi}^{j,\ell}}(0) := 0$. Comparing with $\psi^{j,\ell}$ in (4.18) and (4.19), we can easily check that (5.7) holds by a similar argument as in the proof of Theorem 3. By Corollary 3, we see that (5.5) and (5.6) hold. Now the conclusion that $\mathsf{AS}^{\mathsf{M}_\lambda}_J(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ is an affine tight M_λ-framelet for $L_2(\mathbb{R}^2)$ follows from Theorem 5.

The proof for $\mathsf{AS}^{\mathsf{M}_\lambda}_J(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ is essentially the same. □

5.2. The filter bank structure of smooth affine shear tight frames

Since tight framelets are closely related to filter banks ([16]), we next study the filter bank structure of affine tight M_λ-framelets and affine shear tight frames in Corollary 6. For a filter $u=\{u(\mathsf{k})\}_{\mathsf{k}\in\mathbb{Z}^2}:\mathbb{Z}^2\to\mathbb{C}$, we define its Fourier series $\widehat{u}:\mathbb{R}^2\to\mathbb{C}$ to be $\widehat{u}(\xi)=\sum_{\mathsf{k}\in\mathbb{Z}^2}u(\mathsf{k})e^{-i\mathsf{k}\cdot\xi}$, $\xi\in\mathbb{R}^2$. Obviously, $\widehat{u}$ is $2\pi\mathbb{Z}^2$-periodic. For the non-stationary $\mathsf{AS}^{\mathsf{M}_\lambda}_J(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$, by our construction and setting $0<\rho<1$, we can choose ρ_0 and ε to satisfy $0<\rho<\rho_0<1$ and $0<\varepsilon_0<\lambda^2(\rho_0/\rho-1)/2$ so that $\operatorname{supp}\widehat{\varphi^j}(\mathsf{M}_\lambda\cdot)\subseteq\operatorname{supp}\widehat{\varphi^{j+1}}\subseteq[-\rho_0\pi,\rho_0\pi]^2$ and $\operatorname{supp}\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda\cdot)\subseteq\operatorname{supp}\widehat{\varphi^{j+1}}\subseteq[-\rho_0\pi,\rho_0\pi]^2$. Let $a^j, b^{j,\ell}, j\in\mathbb{N}_0$ be filters defined by their Fourier series as follows:

$$\widehat{a^j}(\xi) := \begin{cases}\frac{\widehat{\varphi^j}(\mathsf{M}_\lambda\xi)}{\widehat{\varphi^{j+1}}(\xi)} & \text{if } \xi\in\operatorname{supp}\widehat{\varphi^j}(\mathsf{M}_\lambda\cdot),\\ 0 & \text{if } \xi\in[-\pi,\pi)^2\setminus\operatorname{supp}\widehat{\varphi^j}(\mathsf{M}_\lambda\cdot),\end{cases}$$

$$\widehat{b^{j,\ell}}(\xi) := \widehat{b^j}(\xi)\frac{\gamma_\varepsilon(\lambda^j\xi_2/\xi_1+\ell)}{\sqrt{\boldsymbol{\Gamma}^j(\xi)}}, \quad |\ell|<\ell_{\lambda^j}-1,$$

$$\widehat{b^{j,\pm\ell_{\lambda^j}}}(\xi) := \widehat{b^j}(\xi)\frac{\gamma^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}(\lambda^j\xi_2/\xi_1\pm\ell_{\lambda^j})}{\sqrt{\boldsymbol{\Gamma}^j(\xi)}}, \tag{5.12}$$

where $\widehat{b^j}(\xi)=\sqrt{\mathbf{g}^j(\xi)-|\widehat{a^j}(\xi)|^2}$ for some function $\mathbf{g}^j$ defined on $\mathbb{T}^2$ satisfying $\mathbf{g}^j=1$ on the support of $\widehat{\varphi^{j+1}}$.

Similarly for the quasi-stationary $\mathsf{AS}^{\mathsf{M}_\lambda}_J(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$, we can define a sequence of filter banks. In this case, the low-pass filter $\widehat{a}$ of $2\pi\mathbb{Z}^2$-periodic function for φ is fixed as follows:

$$\widehat{a}(\xi)=\boldsymbol{\mu}_{\lambda,t,\rho}(\xi_1)\boldsymbol{\mu}_{\lambda,t,\rho}(\xi_2), \quad \xi\in[-\pi,\pi)^2 \tag{5.13}$$

with $\boldsymbol{\mu}_{\lambda,t,\rho}$ as in (4.3). Note that $\operatorname{supp}\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda\cdot)\subseteq\operatorname{supp}\widehat{\varphi}$. Define $2\pi\mathbb{Z}^2$-periodic functions $\widehat{b^{j,\ell}}$ for $\widehat{\mathring{\psi}^{j,\ell}}$, $j\in\mathbb{N}_0$ as follows:

$$\widehat{b^{j,\ell}}(\xi) := \widehat{b}(\xi)\frac{\gamma_\varepsilon(\lambda^j\xi_2/\xi_1+\ell)}{\sqrt{\boldsymbol{\Gamma}_j(\xi)}}, \quad |\ell|\le\ell_{\lambda^j},\ \xi\in[-\pi,\pi)^2 \tag{5.14}$$

with $\widehat{b}(\xi) := \sqrt{\mathbf{g}(\xi)-|\widehat{a}(\xi)|^2}$ for some function $\mathbf{g}$ defined on $\mathbb{T}^2$ satisfying $\mathbf{g}=1$ on the support of $\widehat{\varphi}$.

By [16, Corollary 18 and Theorem 17], we have the following result.

Corollary 7. *Retain all the conditions for $\lambda, t, \rho, \varepsilon, \varepsilon_0$ as in Theorem 3 with $J_0 = 0$ and choose $\rho, \rho_0, \varepsilon$ to satisfy $0 < \rho < \rho_0 < 1$ and $0 < \varepsilon_0 < \lambda^2(\rho_0/\rho - 1)$ so that $\operatorname{supp}\widehat{\varphi^j}$ and $\operatorname{supp}\widehat{\psi^{j,\ell}}$ are inside $[-\rho_0\pi, \rho_0\pi]^2$. Let $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J; \{\mathring{\Psi}_j\}_{j=J}^\infty)$, $J \in \mathbb{N}_0$ be defined as in (5.2) with $\mathring{\Psi}_j$ as in (5.4). Let $a^j, b^{j,\ell}$ be defined as in (5.12). Then there exist $\mathbf{g}^j \in C^\infty(\mathbb{T}^2)$, $j \ge J_0$ such that $\widehat{a^j}, \widehat{b^{j,\ell}} \in C^\infty(\mathbb{T}^2)$ for all $j \in \mathbb{N}_0, \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}$, and*

$$\widehat{\varphi^j}(\mathsf{M}_\lambda\xi) = \widehat{a^j}(\xi)\widehat{\varphi^{j+1}}(\xi) \quad \text{and} \quad \widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda\xi) = \widehat{b^{j,\ell}}(\xi)\widehat{\varphi^{j+1}}(\xi), \quad j \in \mathbb{N}_0, \ \text{a.e. } \xi \in \mathbb{R}^2. \tag{5.15}$$

If in addition M_λ is an integer matrix (that is, $|\lambda|^{1/2} \in \mathbb{N}$), then $\{a^j; b^{j,\ell}, b^{j,\ell}(\mathsf{E}\cdot) : \ell = -\ell_{\lambda^j}, \dots, \ell_{\lambda^j}\}$ is a generalized tight M_λ-framelet filter bank, i.e.,

$$\left|\widehat{a^j}(\xi)\right|^2 + \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\widehat{b^{j,\ell}}(\xi)\right|^2 + \left|\widehat{b^{j,\ell}}(\mathsf{E}\xi)\right|^2\right) = 1, \quad \text{a.e. } \xi \in \sigma_{\varphi^{j+1}}, \tag{5.16}$$

and

$$\overline{\widehat{a^j}(\xi)}\widehat{a^j}(\xi + 2\pi\omega) + \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left[\overline{\widehat{b^{j,\ell}}(\xi)}\widehat{b^{j,\ell}}(\xi + 2\pi\omega) + \overline{\widehat{b^{j,\ell}}(\mathsf{E}\xi)}\widehat{b^{j,\ell}}\big(\mathsf{E}(\xi + 2\pi\omega)\big)\right] = 0 \tag{5.17}$$

for a.e. $\xi \in \sigma_{\varphi^{j+1}} \cap (\sigma_{\varphi^{j+1}} - 2\pi\omega)$ and for $\omega \in \Omega_{\mathsf{M}_\lambda}\backslash\{0\}$ with $\Omega_{\mathsf{M}_\lambda} := [\mathsf{M}_\lambda^{-1}\mathbb{Z}^2] \cap [0,1)^2$ and $\sigma_{\varphi^{j+1}} := \{\xi \in \mathbb{R}^2 : \sum_{\mathsf{k}\in\mathbb{Z}^2} |\widehat{\varphi^{j+1}}(\xi + 2\pi\mathsf{k})|^2 \ne 0\}$.

Proof. It follows from Corollary 6 that $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J; \{\mathring{\Psi}_j\}_{j=J}^\infty)$ in (5.2) is an affine tight M_λ-framelet for $L_2(\mathbb{R}^2)$ for every $J \in \mathbb{N}_0$. By the construction of $\widehat{a^j}$, it is easily seen that the first identity in (5.15) holds.

Since $\operatorname{supp}\widehat{\varphi^j}(\mathsf{M}_\lambda\cdot)$ is strictly inside $\operatorname{supp}\widehat{\varphi^{j+1}}$, by the smoothness of $\widehat{\varphi^j}$ and $\widehat{\varphi^{j+1}}$, it is trivial that $\widehat{a^j} \in C^\infty(\mathbb{T}^2)$. We next show that there exist $\mathbf{g}^j \in C^\infty(\mathbb{T}^2)$ such that $\widehat{b^{j,\ell}} \in C^\infty(\mathbb{T}^2)$. Since $\operatorname{supp}\widehat{\varphi^{j+1}}$ and $\operatorname{supp}\widehat{\psi^{j,\ell}}$ are inside $[-\rho_0\pi, \rho_0\pi]^2$, one can construct a function $\mathbf{g}^j \in C^\infty(\mathbb{T}^2)$ such that $\mathbf{g}^j(\xi) = 1$ for $\xi \in [-\rho_0\pi, \rho_0\pi]^2$ and $\mathbf{g}^j(\xi) = 0$ for $\xi \in \mathbb{T}^2\backslash[-\rho_1\pi, \rho_1\pi]^2$ for some ρ_1 such that $0 < \rho_0 < \rho_1 < 1$. Since $\operatorname{supp}\boldsymbol{\omega}_{\lambda,t,\rho}^j(\mathsf{M}_\lambda\cdot) \subseteq \operatorname{supp}\widehat{\varphi^{j+1}}$, we have

$$\begin{aligned}\boldsymbol{\omega}_{\lambda,t,\rho}^j\big(\mathsf{M}_\lambda^{j+1}\xi\big) &= \left(\left|\widehat{\varphi^{j+1}}(\xi)\right|^2 - \left|\widehat{\varphi^j}(\mathsf{M}_\lambda\xi)\right|^2\right)^{1/2} = \left(\left|\widehat{\varphi^{j+1}}(\xi)\right|^2 - \left|\widehat{a^j}(\xi)\widehat{\varphi^{j+1}}(\xi)\right|^2\right)^{1/2}\\ &= \left(1 - \left|\widehat{a^j}(\xi)\right|^2\right)^{1/2}\widehat{\varphi^{j+1}}(\xi) = \left(\mathbf{g}^j(\xi) - \left|\widehat{a^j}(\xi)\right|^2\right)^{1/2}\widehat{\varphi^{j+1}}(\xi).\end{aligned}$$

Obviously, $(\mathbf{g}^j(\xi) - |\widehat{a^j}(\xi)|^2)^{1/2} \in C^\infty(\mathbb{T}^2)$. Then,

$$\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda\xi) = \boldsymbol{\omega}_{\lambda,t,\rho}^j(\mathsf{M}_\lambda\xi)\frac{\gamma_\varepsilon(\lambda^j\xi_2/\xi_1 + \ell)}{\sqrt{\boldsymbol{\Gamma}^j(\xi)}} = \widehat{b^{j,\ell}}(\xi)\widehat{\varphi^{j+1}}(\xi)$$

with $\widehat{b^{j,\ell}}(\xi) = \widehat{b^j}(\xi)\frac{\gamma_\varepsilon(\lambda^j\xi_2/\xi_1+\ell)}{\sqrt{\boldsymbol{\Gamma}^j(\xi)}}$ being a function in $\mathbb{C}^\infty(\mathbb{T}^2)$. Similarly,

$$\widehat{\mathring{\psi}^{j,\pm\ell_{\lambda^j}}}(\mathsf{M}_\lambda\xi) = \boldsymbol{\omega}_{\lambda,t,\rho}^j(\mathsf{M}_\lambda\xi)\frac{\gamma_{\lambda^j,\varepsilon,\varepsilon_0}^{\mp}(\lambda^j\xi_2/\xi_1 \pm \ell)}{\sqrt{\boldsymbol{\Gamma}^j(\xi)}} = \widehat{b^{j,\pm\ell_{\lambda^j}}}(\xi)\widehat{\varphi^{j+1}}(\xi).$$

This proves the second identity in (5.15) and $\widehat{a^j}, \widehat{b^{j,\ell}} \in C^\infty(\mathbb{T}^2)$.

Since M_λ is an integer matrix, by [16, Theorem 17], (5.16) and (5.17) hold. □

The conclusions in Corollary 7 also hold for $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ in Corollary 6 by replacing φ^j, a^j, $\mathbf{g}^j$ with φ, a, $\mathbf{g}$, respectively.

The sequence of systems $\{\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)\}_{J=0}^\infty$ in (4.21) of Theorem 3 with $J_0 = 0$ has two different dilation matrices A_λ and $\mathsf{E}\mathsf{A}_\lambda\mathsf{E}$ for two cones. On the one hand, the functions $\varphi^J, J \in \mathbb{N}_0$ with the dilation matrix M_λ induce an MRA $\{\mathscr{V}_j\}_{j=0}^\infty$ with

$$\mathscr{V}_j := \overline{\operatorname{span}}\{\varphi^j(\mathsf{M}_\lambda^j \cdot -\mathsf{k}) : \mathsf{k} \in \mathbb{Z}^2\}.$$

The function space $\mathscr{V}_j$ is shift-invariant on the lattice $\mathsf{N}_\lambda^j\mathbb{Z}^2$. On the other hand, at the scale level $j \in \mathbb{N}_0$, the wavelet subspace $\mathscr{W}_j$ is given by

$$\mathscr{W}_j := \overline{\operatorname{span}}\{\psi^{j,\ell}(S^\ell\mathsf{A}_\lambda^j \cdot -\mathsf{k}), \psi^{j,\ell}(S^\ell\mathsf{A}_\lambda^j\mathsf{E} \cdot -\mathsf{k}) \ : \ \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}, \mathsf{k} \in \mathbb{Z}^2\}.$$

While $\mathscr{V}_j$ is shift-invariant on the lattice $\mathsf{N}_\lambda^j\mathbb{Z}^2$, $\mathscr{W}_j$ is not shift-invariant on the lattice $\mathsf{N}_\lambda^j\mathbb{Z}^2$. Define $\mathring{\Psi}_j := \{\mathring{\psi}^{j,\ell} \ : \ \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}\}$ with $j \in \mathbb{N}_0$ and $\mathring{\psi}^{j,\ell}$ in (5.4). Define the wavelet subspace $\mathring{\mathscr{W}}_j$ of $\{\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)\}_{J=0}^\infty$ at the scale level j as follows:

$$\mathring{\mathscr{W}}_j := \overline{\operatorname{span}}\{\mathring{\psi}^{j,\ell}(\mathsf{M}_\lambda^j \cdot -\mathsf{k}), \mathring{\psi}^{j,\ell}(\mathsf{M}_\lambda^j\mathsf{E} \cdot -\mathsf{k}) : \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}, \mathsf{k} \in \mathbb{Z}^2\}.$$

It is trivial to see that $\mathring{\mathscr{W}}_j$ is shift-invariant on the lattice $\mathsf{N}_\lambda^j\mathbb{Z}^2$.

By Corollary 6, $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ is a tight M_λ-framelet for $L_2(\mathbb{R}^2)$ for all $J \in \mathbb{N}_0$. By (5.16) in Corollary 7 and a similar relation as in [16, (1.6)], we have $\mathscr{V}_j \subseteq \mathscr{V}_{j+1}$ and $\mathring{\mathscr{W}}_j \subseteq \mathscr{V}_{j+1}$ for all $j \in \mathbb{N}_0$. By the relations in (5.4), we have

$$\mathring{\mathscr{W}}_j := \overline{\operatorname{span}}\{\psi^{j,\ell}(S^\ell\mathsf{A}_\lambda^j \cdot -S^\ell\mathsf{D}_\lambda^{-j}\mathsf{k}), \psi^{j,\ell}(S^\ell\mathsf{A}_\lambda^j\mathsf{E} \cdot -S^\ell\mathsf{D}_\lambda^{-j}\mathsf{k}) \ : \ \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}, \mathsf{k} \in \mathbb{Z}^2\}.$$

When λ is an integer, we have $\mathbb{Z}^2 \subseteq S^\ell\mathsf{D}_\lambda^{-j}\mathbb{Z}^2$ and we see that (5.6) implies (5.7) by $\widehat{\mathring{\psi}^{j,\ell}} = \widehat{\psi^{j,\ell}}(S_\ell\mathsf{D}_\lambda^j\cdot)$. Therefore, for this case, we have $\mathscr{W}_j \subseteq \mathring{\mathscr{W}}_j$ for all $j \in \mathbb{N}_0$, that is, when λ is an integer, the affine shear tight frame $\mathsf{AS}_J(\varphi^J;\{\Psi_j\}_{j=J}^\infty)$ is indeed a (properly re-scaled) subsystem of the affine tight M_λ-framelet $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi^J;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ through subsampling. Since both of these two systems share the same refinable functions φ^j, $\mathscr{V}_j$ of the MRA for these two systems are the same.

6. Numerical implementation and comparison results on image denoising

In this section we first discuss how to construct a particular family of smooth quasi-stationary affine shear tight frames through the construction of directional affine tight framelets and their underlying directional tight framelet filter banks. Then we briefly discuss the numerical implementation of our smooth affine shear tight frames by employing their underlying filter banks, and compare their performance on the image denoising problem to other existing directional multiscale representation systems such as curvelets and shearlets. Our construction and implementation are based on the underlying filter banks of our quasi-stationary smooth affine shear tight frames.

In a nutshell, we shall construct a sequence of directional tight framelet filter banks $\{a; b^{j,\ell}, b^{j,\ell}(\mathsf{E}\cdot) : \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}\}$ with $j \in \mathbb{N}_0$ for decomposition and reconstruction of images. Here, j corresponds to the scale level and $\mathsf{M}_\lambda = \lambda^2\mathsf{I}_2$ is an integer matrix with $\lambda > 1$. As argued in [16] for directional tight framelets, when j increases, the number ℓ_{λ^j} of directions should also increase. With such a sequence of filter banks, we can define φ through $\widehat{\varphi} := \lim_{J\to\infty}\prod_{j=0}^J \widehat{a}(\mathsf{N}_\lambda^{j+1}\cdot)$ and $\mathring{\psi}^{j,\ell}$ by $\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda^\mathsf{T}\xi) = \widehat{b^{j,\ell}}(\xi)\widehat{\varphi}(\xi)$, $\xi \in \mathbb{R}^2$. Then we automatically have $\widehat{\varphi}(\mathsf{M}_\lambda\cdot) = \widehat{a}\widehat{\varphi}$. Let $\mathring{\Psi}_j := \{\mathring{\psi}^{j,\ell} : \ell = -\ell_{\lambda^j}, \ldots, \ell_{\lambda^j}\}$. Define $\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)$ as in (5.3). Now by [16, Theorem 17] (also cf. [16, Corollary 18]), we have a sequence of affine tight M_λ-framelets

$\{\mathsf{AS}_J^{\mathsf{M}_\lambda}(\varphi;\{\mathring{\Psi}_j\}_{j=J}^\infty)\}_{J=0}^\infty$. By (5.4), we can define $\psi^{j,\ell}$ to be $\psi^{j,\ell}=\lambda^j\mathring{\psi}^{j,\ell}(\mathsf{D}_\lambda^j S^\ell\cdot)$. Let $\Psi_j:=\{\psi^{j,\ell}(S^{-\ell}\cdot):\ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}\}$. Assume that $\widehat{a}\ge 0$ and $\widehat{b^{j,\ell}}\ge 0$ for all $j\in\mathbb{N}_0$ and $\ell=-\ell_{\lambda^j},\dots,\ell_{\lambda^j}$. It follows from [16, Corollary 18] that (5.7) holds and $\widehat{\varphi}(\xi)\widehat{\varphi}(\xi+2\pi\mathsf{k})=0$ for all $\mathsf{k}\in\mathbb{Z}^2\backslash\{0\}$. Moreover, all the generators in $\{\varphi\}\cup\{\mathring{\Psi}_j\}_{j=0}^\infty$ have nonnegative Fourier transforms. If $\operatorname{supp}\widehat{a}$ is small enough, then the support of $\widehat{\varphi}$ will be contained inside $[-\rho\pi,\rho\pi]^2$ for sufficiently small $0<\rho<1$. Consequently, due to the identity $\widehat{\mathring{\psi}^{j,\ell}}(\mathsf{M}_\lambda^\mathsf{T}\cdot)=\widehat{b^{j,\ell}}\widehat{\varphi}$, the support of $\widehat{\psi^{j,\ell}}$ will be small enough so that (5.6) holds. Consequently, by Theorem 5, we also have a sequence of affine shear tight frames $\{\mathsf{AS}_J(\varphi;\{\Psi_j\}_{j=J}^\infty)\}_{J=0}^\infty$ as defined in (4.25).

Following the lines developed in [16] for the construction of directional tight framelets, we next give a concrete example for the construction of a sequence of tight framelet filter banks for our affine shear tight frames. For simplicity, we assume $\lambda=\sqrt{2}$ so that $\mathsf{M}_\lambda=2\mathsf{I}_2$ is an integer matrix, though the same construction can be modified for general λ satisfying $|\lambda|^{1/2}\in\mathbb{N}$. For parameters $c_0>0$ and $\epsilon_0>0$ satisfying $c_0+\epsilon_0\le\pi/2$ (for downsampling by 2), we can define a low-pass filter $a:\mathbb{Z}^2\to\mathbb{R}$ by

$$\widehat{a}(\xi)=\chi_{[-c_0,c_0];\epsilon_0,\epsilon_0}(\xi_1)\chi_{[-c_0,c_0];\epsilon_0,\epsilon_0}(\xi_2),\quad \xi\in[-\pi,\pi)^2 \tag{6.1}$$

with the bump function $\chi_{[c_L,c_R];\epsilon_L,\epsilon_R}$ being defined to be

$$\chi_{[c_L,c_R];\epsilon_L,\epsilon_R}(t)=\begin{cases}\nu(\frac{t-c_L}{\epsilon_L}) & \text{if } t<c_L+\epsilon_L,\\ 1 & \text{if } c_L+\epsilon_L\le t\le c_R-\epsilon_R,\\ \nu(\frac{c_R-t}{\epsilon_R}) & \text{if } t>c_R-\epsilon_R.\end{cases}$$

Now at each scale level $j\in\mathbb{N}_0$ and $k_j\in\mathbb{N}_0$, using a similar idea as in [16] for directional tight framelets, we are going to construct directional high-pass filters $b^{j,\ell}:\mathbb{Z}^2\to\mathbb{C}$ such that $\{a;b^{j,\ell},b^{j,\ell}(\mathsf{E}\cdot):\ell=-2^{k_j},\dots,2^{k_j}\}$ forms a tight M_λ-framelet filter bank:

$$\left|\widehat{a}(\xi)\right|^2+\sum_{\ell=-2^{k_j}}^{2^{k_j}}\left(\left|\widehat{b^{j,\ell}}(\xi)\right|^2+\left|\widehat{b^{j,\ell}}(\mathsf{E}\xi)\right|^2\right)=1,\quad \xi\in\mathbb{R}^2, \tag{6.2}$$

$$\overline{\widehat{a}(\xi)}\widehat{a}(\xi+2\pi\omega)+\sum_{\ell=-2^{k_j}}^{2^{k_j}}\left[\overline{\widehat{b^{j,\ell}}(\xi)}\widehat{b^{j,\ell}}(\xi+2\pi\omega)+\overline{\widehat{b^{j,\ell}}(\mathsf{E}\xi)}\widehat{b^{j,\ell}}\big(\mathsf{E}(\xi+2\pi\omega)\big)\right]=0,\quad \xi\in\mathbb{R}^2 \tag{6.3}$$

for all $\omega\in[\mathsf{N}_\lambda\mathbb{Z}^2]\cap[0,1)^2$ with $\mathsf{N}_\lambda=\frac{1}{2}\mathsf{I}_2$ (note that we assumed $\sqrt{\lambda}=2$). The total number of shear directions at this scale level j is $2^{k_j+2}+2$ with $2^{k_j+1}+1$ shear directions for both the horizontal cone and the vertical cone. To this end, we use an auxiliary function $a_1\in L_2(\mathbb{R}^2)$ defined by $\widehat{a_1}(\xi):=\chi_{[-c_1,c_1];\epsilon_1,\epsilon_1}(\xi_1)\chi_{[-c_1,c_1];\epsilon_1,\epsilon_1}(\xi_2)$ (with $c_1=\pi$ and $c_1+\epsilon_1-(c_0-\epsilon_0)\le\pi$ for downsampling at least by 2 for high-pass filter coefficients). Thanks to the property of ν, one can show that

$$\sum_{\mathsf{k}\in\mathbb{Z}^2}\left|\widehat{a_1}(\xi+2\pi\mathsf{k})\right|^2=1\quad\forall\xi\in\mathbb{R}^2. \tag{6.4}$$

We define a function $b\in L_2(\mathbb{R}^2)$ by

$$\widehat{b}(\xi):=\begin{cases}\sqrt{|\widehat{a_1}(\xi)|^2-|\widehat{a}(\xi)|^2} & \text{if } \xi\in\operatorname{supp}\widehat{a_1},\\ 0 & \text{otherwise.}\end{cases} \tag{6.5}$$

Now, we apply the splitting technique to $\widehat{b}$ for the construction of high-pass filters $b^{j,\ell}$. Recall the definition of $\mathbf{\Gamma}_j$ in (4.10), we have

$$\mathbf{\Gamma}_{k_j}(\xi) = \sum_{\ell=-2^{k_j}}^{2^{k_j}} \left(\left|\gamma_\varepsilon\left(2^{k_j}\xi_2/\xi_1+\ell\right)\right|^2 + \left|\gamma_\varepsilon\left(2^{k_j}\xi_1/\xi_2+\ell\right)\right|^2\right), \quad \xi \neq 0 \tag{6.6}$$

for some $0 \le \varepsilon \le 1/2$. Note that $\widehat{b}(\xi)\gamma_\varepsilon(2^{k_j}\xi_2/\xi_1+\ell)/\sqrt{\mathbf{\Gamma}_{k_j}(\xi)}$ is not a $2\pi\mathbb{Z}^2$-periodic function. We define $\widehat{b^{j,\ell}}$, $\ell = -2^{k_j}, \dots, 2^{k_j}$ to be the $2\pi\mathbb{Z}^2$-periodization of $\widehat{b}(\xi)\gamma_\varepsilon(2^{k_j}\xi_2/\xi_1+\ell)/\sqrt{\mathbf{\Gamma}_{k_j}(\xi)}$ as follows:

$$\widehat{b^{j,\ell}}(\xi) := \sum_{\mathsf{k}=(\mathsf{k}_1,\mathsf{k}_2)\in\mathbb{Z}^2} \widehat{b}(\xi+2\pi\mathsf{k})\frac{\gamma_\varepsilon(2^{k_j}(\xi_2+2\pi\mathsf{k}_2)/(\xi_1+2\pi\mathsf{k}_1)+\ell)}{\sqrt{\mathbf{\Gamma}_{k_j}(\xi+2\pi\mathsf{k})}}, \quad \xi \in \mathbb{T}^2\backslash\{0\} \tag{6.7}$$

and $\widehat{b^{j,\ell}}(0) := 0$. Now, in view of (6.4) and (6.6), it is easy to show that $\{a; b^{j,\ell}, b^{j,\ell}(\mathsf{E}\cdot) : \ell = -2^{k_j}, \dots, 2^{k_j}\}$ is a tight M_λ-framelet filter bank such that $\widehat{a} \ge 0$ and $\widehat{b^{j,\ell}} \ge 0$ for all $j \in \mathbb{N}_0$ and $\ell = -2^{k_j}, \dots, 2^{k_j}$. For simplicity, we denote by $B_j := \{b^{j,\ell}, b^{j,\ell}(\mathsf{E}\cdot) : \ell = -2^{k_j}, \dots, 2^{k_j}\}$. Then we have a tight M_λ-framelet filter bank $\{a; B_j\}$ at every scale level $j \in \mathbb{N}_0$. For practical applications, we only apply $\{a; B_j\}$, $j = 0, \dots, J-1$ for some positive integer J, where J is the level of decomposition.

We next discuss the implementation of the forward transform (decomposition). Without loss of generality and for simplicity of presentation, we assume that the image size $N = 2^K$ for some integer $K \ge 0$. Choose a positive integer $J \le K-1$ as the decomposition level and a sequence of nonnegative integers $k_j, j = 0, \dots, J-1$ corresponding to the number of shear directions at the scale level j. Let $\Lambda_{(K,K)} := [0, 2^K-1]^2 \cap \mathbb{Z}^2$ be a Cartesian grid for images of size $2^K \times 2^K$. Then the (normalized) discrete Fourier transform (DFT) $\mathscr{F}_K$ maps an image $u^J : \Lambda_{(K,K)} \to \mathbb{R}$ to a $2\pi\mathbb{Z}^2$-periodic image $\widehat{u^J} : \widehat{\Lambda}_{(K,K)} \to \mathbb{C}$ in the frequency domain, where $\widehat{u^J}(\mathsf{k}) = (\mathscr{F}_K u^J)(\mathsf{k}) := \frac{1}{2^K}\sum_{\mathsf{n}\in\Lambda_K} u^J(\mathsf{n})e^{-i\mathsf{n}\cdot\mathsf{k}/2^K}$ and $\mathsf{k} \in \widehat{\Lambda}_{(K,K)} := \frac{2\pi}{2^K}\Lambda_{(K,K)}$. In what follows, we regard u^J as a $2^K\mathbb{Z}^2$-periodic image; that is, $u^J(\cdot + 2^K\mathsf{k}) = u^J$ for all $\mathsf{k} \in \mathbb{Z}^2$.

At the scale level J, with u^J as our input image, we apply the tight framelet filter bank $\{a; B_{J-1}\}$. For the low-pass coefficients, we first compute the convolution $u^J * a$ given by $[u^J * a](\mathsf{k}) := \sum_{\mathsf{n}} u^J(\mathsf{n})a(\mathsf{n}-\mathsf{k})$, $\mathsf{k} \in \mathbb{Z}^2$, which can be implemented by $\mathscr{F}_K^{-1}(\widehat{u^J} \cdot \widehat{a}|_{\widehat{\Lambda}_{(K,K)}})$, where $\widehat{a}|_{\widehat{\Lambda}_{(K,K)}}$ is the restriction of $\widehat{a}$ on the lattice $\widehat{\Lambda}_K$. Applying the downsampling by 2 operation on $u^J * a$, we obtain the $2^{K-1}\mathbb{Z}^2$-periodic low-pass coefficients $u^{J-1} : \Lambda_{(K-1,K-1)} \to \mathbb{R}$ by $u^{J-1}(\mathsf{k}) = 2u^J(2\mathsf{k})$, $\mathsf{k} \in \mathbb{Z}^2$.

For the high-pass coefficients $c_h^{J-1,\ell} : \Lambda_{(K-1,K-k_{J-1})} \to \mathbb{C}$ and $c_v^{J-1,\ell} : \Lambda_{(K-k_{J-1},K-1)} \to \mathbb{C}$, similar to the case of the low-pass coefficients, we have

$$c_h^{J-1,\ell}(\mathsf{k}_1,\mathsf{k}_2) = \sqrt{2^{k_{J-1}+1}}\left[\mathscr{F}_K^{-1}\left(\widehat{u^J}\cdot\widehat{b^{J-1,\ell}}\big|_{\Lambda_{(K,K)}}\right)\right]\left(2\mathsf{k}_1, 2^{k_{J-1}}\mathsf{k}_2\right), \quad (\mathsf{k}_1,\mathsf{k}_2) \in \Lambda_{(K-1,K-k_{J-1})},$$
$$c_v^{J-1,\ell}(\mathsf{k}_1,\mathsf{k}_2) = \sqrt{2^{k_{J-1}+1}}\left[\mathscr{F}_K^{-1}\left(\widehat{u^J}\cdot\widehat{b^{J-1,\ell}(\mathsf{E}\cdot)}\big|_{\Lambda_{(K,K)}}\right)\right]\left(2^{k_{J-1}}\mathsf{k}_1, 2\mathsf{k}_2\right), \quad (\mathsf{k}_1,\mathsf{k}_2) \in \Lambda_{(K-k_{J-1},K-1)},$$

for $\ell = -2^{k_{J-1}}, \dots, 2^{k_{J-1}}$.

Now u^{J-1} is a $2^{K-1} \times 2^{K-1}$ image. Repeating the above procedure by replacing u^J, $\mathscr{F}_K$, and $\{a; B_{J-1}\}$ with u^{J-1}, $\mathscr{F}_{K-1}$, and $\{a, B_{J-2}\}$, respectively. We can obtain the next scale low-pass coefficients u^{J-2} and high-pass coefficients $c_h^{J-2,\ell}, c_v^{J-2,\ell}$, $\ell = -2^{k_{J-2}}, \dots, 2^{k_{J-2}}$. Repeating the procedure for $J-2, \dots, 0$. Eventually, we have a sequence of coefficients: $\{u^0\} \cup \{c_h^{j,\ell}, c_v^{j,\ell} : \ell = -2^{k_j}, \dots, 2^{k_j}\}_{j=0}^{J-1}$.

For the implementation of the backward transform (reconstruction), it is merely the reverse of the above steps with each operator replaced by its adjoint operator. For example, suppose that we are at the scale level j with $\{u^{J-j}\} \cup \{c_h^{J-j,\ell}, c_v^{J-j,\ell} : \ell = -2^{k_{J-j}}, \dots, 2^{k_{J-j}}\}$ given as above and would like to reconstruct the finer scale image $u^{J-j+1} : \Lambda_{(K-j+1,K-j+1)} \to \mathbb{R}$. We first upsample the low-pass filter image u^{J-j} to be $c_0^{J-j+1} : \Lambda_{(K-j+1,K-j+1)} \to \mathbb{R}$:

$$c_0^{J-j+1}(\mathsf{k}) = \begin{cases} 2u^{J-j}(\mathsf{k}/2) & \text{if } \mathsf{k}/2 \in \Lambda_{(K-j,K-j)}, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, we upsample the high-pass coefficients to obtain $c_{h,0}^{J-j+1,\ell} : \Lambda_{(K-j+1,K-j+1)} \to \mathbb{R}$ and $c_{v,0}^{J-j+1,\ell} : \Lambda_{(K-j+1,K-j+1)} \to \mathbb{R}$ by

$$c_{h,0}^{J-j+1,\ell}(\mathsf{k}_1,\mathsf{k}_2) = \begin{cases} \sqrt{2^{k_{J-j}+1}} c_h^{j,\ell}(\mathsf{k}_1/2, \mathsf{k}_2/2^{k_{J-j}}) & \text{if } (\mathsf{k}_1/2, \mathsf{k}_2/2^{k_{J-j}}) \in \Lambda_{(K-j,K-k_{J-j})}, \\ 0 & \text{otherwise} \end{cases}$$

and

$$c_{v,0}^{J-j+1,\ell}(\mathsf{k}_1,\mathsf{k}_2) = \begin{cases} \sqrt{2^{k_{J-j}+1}} c_v^{j,\ell}(\mathsf{k}_1/2^{k_{J-j}}, \mathsf{k}_2/2) & \text{if } (\mathsf{k}_1/2^{k_{J-j}}, \mathsf{k}_2/2) \in \Lambda_{(K-k_{J-j},K-j)}, \\ 0 & \text{otherwise.} \end{cases}$$

The reconstructed image u^{J-j+1} is then given by

$$\begin{aligned} u^{J-j+1} = \mathscr{F}_{K-j+1}^{-1} \Bigg[& \mathscr{F}_{K-j+1}\big(c_0^{J-j+1}\big) \cdot \widehat{a}\big|_{\Lambda_{(K-j+1,K-j+1)}} \\ & + \sum_{\ell=-2^{k_{J-j}}}^{2^{k_{J-j}}} \mathscr{F}_{K-j+1}\big(c_{h,0}^{J-j+1,\ell}\big) \cdot \widehat{b}^{J-j,\ell}\big|_{\Lambda_{(K-j+1,K-j+1)}} \\ & + \sum_{\ell=-2^{k_{J-j}}}^{2^{k_{J-j}}} \mathscr{F}_{K-j+1}\big(c_{v,0}^{J-j+1,\ell}\big) \cdot \widehat{b}^{J-j,\ell}(\mathsf{E}\cdot)\big|_{\Lambda_{(K-j+1,K-j+1)}} \Bigg]. \end{aligned}$$

In the rest of this section we apply our systems to the image denoising problem. We choose the decomposition level $J = 4$; that is, we decompose an image into 5 scales. The parameters $c_0, \epsilon_0, \epsilon_1$ of a_1, a are given by $c_0 = 33/32$, $\epsilon_0 = 69/128$, $\epsilon_1 = 69/128$, and $\varepsilon = 1/2$ for $\boldsymbol{\gamma}_\varepsilon$. For the finest scale level, we use $k_4 = 4$ (total 16 shear directions, 8 on the horizontal cone and 8 on the vertical cone). For the next three scales, we use $k_3 = k_2 = k_1 = 2$ (8 shear directions), and for the coarsest scale level, we use $k_0 = 1$ (4 shear directions). The redundancy rate of our system is about 5.4.

For image denoising, we employ the local-soft (LS) thresholding method: For each sub-band coefficients $c_h^{j,\ell}, c_v^{j,\ell}$, we first normalize them with respect to the sub-band energy, which can be computed by applying the backward transform to a delta image on the support of $c_h^{j,\ell}$ or $c_v^{j,\ell}$ and then compute the l_2 norm of the reconstructed image. Using a local window of size 9×9 with uniform weight $\frac{1}{81}$ and convolving with the normalized coefficients $c_h^{j,\ell}$, $c_v^{j,\ell}$ centering at position k, we can estimate the local coefficient variance σ_{k} at position k and then use the threshold value $T = \frac{\sigma^2}{(\sigma_{\mathsf{k}}^2-\sigma^2)^{1/2}}$ for soft threshold at position k. As usual, the standard deviation σ of Gaussian noise is assumed to be known. Note that the thresholding procedure does not apply to the low-pass coefficients u^0.

We test two standard images: Lena and Barbara of size 512×512. We first employ symmetric boundary extension (with 32 pixels) on the noisy image to avoid boundary effect. We then apply our forward transform to obtain the coefficients. After performing the local-soft thresholding procedure, we then apply the backward transform to the thresholded coefficients and throw away the extended boundary pixels to obtain the final denoised image.

We compare the performance of our method with two other known directional multiscale representation systems: Curvelets in [2] and compactly supported shearlets in [29,30]. All implementations of these two directional multiscale representation systems using curvelets and shearlets can be downloaded from the corresponding authors' websites. We download each of their packages and run their denoising codes for test images of Lena and Barbara. The peak signal-to-noise ratio (PSNR) is used to measure the performance and is defined to be $\mathrm{PSNR}(u, \mathring{u}) = 10 \log_{10} \frac{255^2}{\mathrm{MSE}(u,\mathring{u})}$, where u is the original clean image and $\mathring{u}$ is the

Table 1
PSNR values for test images of Lena and Barbara. ASTF is our proposed affine shear tight frames with redundancy rate 5.4. CurveLab uses frequency wrapping with redundancy rate 2.8. ShearLab DST and DNST employ compactly supported shearlets which have redundancy rate 40 for DST and 49 for DNST and are implemented by an undecimated transform.

σ	512 × 512 Lena				512 × 512 Barbara			
	ASTF (LS)	CurveLab (Wrap)	ShearLab (DST)	ShearLab (DNST)	ASTF (LS)	CurveLab (Wrap)	ShearLab (DST)	ShearLab (DNST)
5	38.19	35.75(2.44)	38.22(−0.03)	38.01(0.18)	37.40	33.81(3.59)	37.76(−0.36)	37.17(0.23)
10	35.18	33.34(1.84)	35.19(−0.01)	35.35(−0.17)	33.74	29.16(4.58)	33.94(−0.20)	33.62(0.12)
15	33.50	31.96(1.54)	33.41(0.09)	33.72(−0.22)	31.75	26.68(5.07)	31.71(0.04)	31.54(0.21)
20	32.33	30.89(1.44)	32.12(0.21)	32.51(−0.18)	30.36	25.46(4.90)	30.12(0.24)	30.08(0.28)
25	31.40	30.06(1.34)	31.09(0.31)	31.51(−0.11)	29.29	24.84(4.45)	28.90(0.39)	28.93(0.36)
30	30.64	29.32(1.32)	30.25(0.39)	30.68(−0.04)	28.42	24.45(3.97)	27.90(0.52)	27.97(0.45)
35	29.98	28.67(1.31)	29.53(0.45)	29.96(0.02)	27.70	24.14(3.56)	27.07(0.63)	27.18(0.52)
40	29.40	28.13(1.27)	28.92(0.48)	29.32(0.08)	27.08	23.87(3.21)	26.36(0.72)	26.48(0.60)
45	28.90	27.62(1.28)	28.37(0.53)	28.74(0.16)	26.54	23.63(2.91)	25.75(0.79)	25.86(0.68)
50	28.46	27.13(1.33)	27.89(0.57)	28.21(0.25)	26.05	23.40(2.65)	25.22(0.83)	25.31(0.74)

denoised image and $\mathrm{MSE}(u, \mathring{u})$ is the mean squared error $\frac{1}{N^2}\sum_{\mathsf{k}\in[0,N-1]^2} |u(\mathsf{k}) - \mathring{u}(\mathsf{k})|^2$. The unit of PSNR is dB.

The CurveLab package at http://www.curvelab.org has two subpackages: one uses un-equispace FFT and the other uses frequency wrapping. Here we use the frequency wrapping package; detailed information on CurveLab package can be found at [3]. The performance of these two subpackages are very close to each other (less than 0.2 dB differences) and here we choose the one with the frequency wrapping for comparison. The total number of scales is 5. At the finest scale level, the CurveLab uses an isotropic wavelet transform to avoid checkerboard effect. At the scale level 4, 32 (angular) directions are used. At the scale levels 3 and 2, 16 (angular) directions are used. At the coarsest scale level, 8 (angular) directions are used. The redundancy rate of the CurveLab wrapping package is about 2.8.

The ShearLab package at http://www.shearlab.org also has many subpackages for different implementations. Here we choose two subpackages using compactly supported shearlets. One is DST as described in [29] and the other is DNST as described in [30]. The DNST in [30] has the best performance so far in the ShearLab package. For DST, the total number of scales is 5. 10 shear directions are used across all scale levels. The redundancy rate of the DST is 40. For DNST, the total number of scales is 4. 16 shear directions are used for the finest scale levels 4 and 3; while 8 shear directions are used for the other two scale levels. All filters are implemented in an undecimated fashion. The redundancy rate of DNST is 49.

We compare the denoising performance over different noise level σ ranging from 5 to 50 with step size 5. The comparison results are presented in Table 1. The values in the brackets are gain or loss of our method comparing to other methods. From Table 1, we see that our method has significant improvement over CurveLab. We have about 1.51 dB improvement in average for Lena and 3.89 dB improvement in average for Barbara. Comparing our method with DST, we have about 0.30 dB improvement in average for Lena and 0.36 dB improvement in average for Barbara. Moreover, our method has better performance than that of DST for all noise level except when noise level σ is small ($\sigma = 5, 10$). When the noise level is high ($\sigma = 50$), we have 0.57 dB improvement over DST for Lena and 0.83 dB improvement for Barbara. Comparing our method with DNST, the average performance is the same for both our method and DNST for Lena while our method has 0.42 dB improvement in average for Barbara. For Lena, our method performs slightly worst than that of DNST when $\sigma \le 30$ but outperforms DNST when $\sigma > 30$. For Barbara, our method outperforms DNST for all σ. We must point out that, in terms of redundancy, CurveLab has the lowest redundancy rate 2.8 but in general has far worse performance. DNST and DST have better performance over CurveLab, yet their redundancy is significantly higher (40 for DST and 49 for DNST since they are undecimated transforms.) than that of CurveLab. Not only our method has small redundancy rate 5.4, but also the performance of our method is in general better than DNST and DST, especially for Barbara.

7. Discussion and extension

In this paper, we mainly investigate affine shear systems in $L_2(\mathbb{R}^2)$ for the purpose of simplicity of presentation. Our characterization and construction can be easily extended to higher dimensions. In $\mathbb{R}^d$ with $d \ge 2$, the shear operator S^τ with $\tau = (\tau_2, \dots, \tau_d) \in \mathbb{R}^{d-1}$ and A_λ are of the form:

$$S^\tau = \begin{bmatrix} 1 & \tau_2 & \dots & \tau_d \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad \text{and} \quad \mathsf{A}_\lambda = \begin{bmatrix} \lambda^2 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{bmatrix}.$$

Define $S_\tau := (S^\tau)^\mathsf{T}$ and denote E_n to be the elementary matrix corresponding to the coordinate exchange between the first axis and the nth one. For example, $\mathsf{E}_1 = \mathsf{I}_d$ and $\mathsf{E}_2 = \operatorname{diag}(\mathsf{E}, I_{d-2})$. For $d = 2$, we have $\mathsf{E}_2 = \mathsf{E}$. Let Ψ_j be given by

$$\Psi_j := \left\{\psi(S^{-\ell}\cdot) : \ell_n = -r_j^n, \dots, r_j^n, n = 2, \dots, d\right\} \cup \left\{\psi^{j,\ell}(S^{-\ell}\cdot) : |\ell_n| = r_j^n + 1, \dots, s_j^n, n = 2, \dots, d\right\}$$

with $\ell = (\ell_2, \dots, \ell_d) \in \mathbb{Z}^{d-1}$ and $\psi, \psi^{j,\ell}$ being functions in $L_2(\mathbb{R}^d)$. For the low frequency part, it corresponds to a function $\varphi^j \in L_2(\mathbb{R}^d)$. Then an affine shear system in $\mathbb{R}^d$ is defined to be

$$\mathsf{AS}_J(\varphi^J; \{\Psi_j\}_{j=J}^\infty) = \left\{\varphi^J_{\mathsf{M}_\lambda^j;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^d\right\} \cup \left\{h_{\mathsf{A}_\lambda^j \mathsf{E}_n;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^d, n = 1, \dots, d,\ h \in \Psi_j\right\}_{j=J}^\infty. \tag{7.1}$$

All the characterizations for affine shear tight frames and sequences of affine shear tight frames can be carried over to the d-dimensional case for the system defined as in (7.1). Since the essential idea of our smooth non-stationary construction and smooth quasi-stationary construction is frequency splitting (see [14]), our 2D construction thus can be easily extended to any high dimensions once an $\boldsymbol{\omega}^j$ is constructed in a way satisfying $\boldsymbol{\omega}^j = (|\widehat{\varphi^{j+1}}(\mathsf{N}_\lambda\cdot)|^2 - |\widehat{\varphi^j}|^2)^{1/2}$ for both the non-stationary and quasi-stationary construction. Filter banks associated with high-dimensional affine systems can be obtained as well as their connection to cone-adapted high-dimensional directional tight framelets.

Several problems remain open in our study of affine shear tight frames. For example, the existence and construction of affine shear tight frames with compactly supported generators in the spatial domain. If we drop the tightness requirement, there are indeed compactly supported shearlet frames, e.g., see [22]. In view of the connection between affine shear tight frames and cone-adapted directional tight framelets, one might want to consider the existence and construction of cone-adapted directional framelets with compactly supported generators first. Another problem is the existence of affine shear tight frames with only one smooth generator; that is, $\Psi_j := \{\psi(S^{-\ell}\cdot) : \ell = -s_j, \dots, s_j\}$ is from one generator ψ. Considering that the shear operator along the seamlines is not consistent for both cones, our conjecture is that there is even no affine shear tight frame with one single generator that is continuous in the frequency domain. In other words, additional seamline generators seem to be unavoidable when considering cone-adapted construction. When $\lambda > 1$ is an integer, we know that an affine shear tight frame can be regarded as a subsystem of a directional tight framelet through sub-sampling, from which an underlying filter bank exists for the affine shear tight frame. However, when $\lambda > 1$ is not an integer, though an affine shear tight frame is still related to a cone-adapted directional tight framelet via (5.9), the lattice $\mathsf{D}_\lambda^j S^\ell \mathbb{Z}^2$ is no longer an integer lattice, the sub-sampling procedure thus fails and we do not know whether there is still an underlying filter bank for such an affine shear tight frame.

Since our affine shear tight frames includes the smooth tight frame of shearlets in [13] as a special case (when $\lambda = 2$), the property of optimal sparse approximation for cartoon-like functions holds true automatically. Curvelets and shearlets possessing optimal sparse approximation property necessarily should have low redundancy rate and their performance for practical applications such as the image denoising problem could be theoretically optimal after performing the simple soft/hard thresholding on the coefficients. However, by Table 1 for image denoising, curvelet (Wrap) with the smallest redundancy rate has significantly lower performance than shearlet (DNST), which has the highest redundancy rate and is implemented by a fully un-decimated transform similar to an undecimated wavelet transform. Therefore, shearlet (DNST) has even much higher redundancy rate than the directional tight framelets in [16] and affine shear tight frames in this paper. Consequently, due to the high/maximum redundancy rate, shearlet (DNST) cannot have the optimal sparse approximation property. For image denoising, this dilemma between optimal sparse approximation property and redundancy rate remains as a mystery to us. Nevertheless, Table 1 suggests that directional multiscale representations indeed are very important for high-dimensional problems such as image denoising.

Our results reveal that the affine shear tight frames are connected to the affine tight framelets in [16] via subsampling, which shows that there are an underlying filter bank structures for the affine shear tight frames. Moreover, in term of numerical implementation, one only needs to implement the affine tight framelets while the implementation of affine shear tight frames are then automatically follows by subsampling. The comparison results show the advantages of our construction and implementation over several existing directional multiscale representation systems, including the compactly supported shearlets implementation. Although there exist compactly supported shearlet systems, yet they are frames but not tight frames. In terms of implementations, one needs to use iterative methods to perform the inverse transform for compactly supported shearlets. It is still open for the construction of compactly supported affine shear tight frames.

8. Proofs

In this section, we provide proofs of some results in the paper.

8.1. Proofs of results in Section 2

Proof of Corollary 2. Note that for a fixed j, $\widehat{\psi}(S_\ell \mathsf{B}_\lambda^j \cdot) = \chi_{Q_{j,\ell}}$ with

$$Q_{j,\ell} = \{\xi \in \mathbb{R}^2 : \lambda^{-j}(-\ell - 1/2) \le \xi_2/\xi_1 \le \lambda^{-j}(-\ell + 1/2), |\xi_1| \in (\lambda^{2j-2}\rho\pi, \lambda^{2j}\rho\pi]\}$$

and $\widehat{\psi^{j,-\ell_{\lambda^j}}}(S_{-\ell_{\lambda^j}} \mathsf{B}_\lambda^j \cdot) = \chi_{Q_{j,-\ell_{\lambda^j}}}$, $\widehat{\psi^{j,\ell_{\lambda^j}}}(S_{\ell_{\lambda^j}} \mathsf{B}_\lambda^j \cdot) = \chi_{Q_{j,\ell_{\lambda^j}}}$ with

$$\begin{aligned} Q_{j,-\ell_{\lambda^j}} &= \{\xi \in \mathbb{R}^2 : \lambda^{-j}(\ell_{\lambda^j} - 1/2) \le \xi_2/\xi_1 \le 1, |\xi_1| \in (\lambda^{2j-2}\rho\pi, \lambda^{2j}\rho\pi]\}, \\ Q_{j,\ell_{\lambda^j}} &= \{\xi \in \mathbb{R}^2 : -1 \le \xi_2/\xi_1 \le -\lambda^{-j}(\ell_{\lambda^j} - 1/2), |\xi_1| \in (\lambda^{2j-2}\rho\pi, \lambda^{2j}\rho\pi]\}. \end{aligned}$$

Thus, we have,

$$\begin{aligned} \mathcal{I}^0_{\Psi_j}(\mathsf{B}_\lambda^j \cdot) &= \sum_{\ell=-\ell_{\lambda^j}-1}^{\ell_{\lambda^j}-1} \left|\widehat{\psi}(S_\ell \mathsf{B}_\lambda^j \cdot)\right|^2 + \left|\widehat{\psi^{j,-\ell_{\lambda^j}}}(S_{-\ell_{\lambda^j}} \mathsf{B}_\lambda^j \cdot)\right|^2 + \left|\widehat{\psi^{j,\ell_{\lambda^j}}}(S_{\ell_{\lambda^j}} \mathsf{B}_\lambda^j \cdot)\right|^2 \\ &= \chi_{\bigcup_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} Q_{j,\ell}} = \chi_{\{\xi \in \mathbb{R}^2 : -1 \le \xi_2/\xi_1 \le 1, |\xi_1| \in (\lambda^{2j-2}\rho\pi, \lambda^{2j}\rho\pi]\}}. \end{aligned}$$

Similarly, we have

$$\mathcal{I}^0_{\Psi_j}(\mathsf{B}_\lambda^j \mathsf{E} \cdot) = \chi_{\{\xi \in \mathbb{R}^2 : -1 \le \xi_1/\xi_2 \le 1, |\xi_2| \in (\lambda^{2j-2}\rho\pi, \lambda^{2j}\rho\pi]\}}.$$

Consequently, we obtain

$$\mathcal{I}^0_{\varphi}(\xi) + \sum_{j=0}^{\infty}\left[\mathcal{I}^0_{\Psi_j}\left(\mathsf{B}^j_\lambda \xi\right) + \mathcal{I}^0_{\Psi_j}\left(\mathsf{B}^j_\lambda \mathsf{E}\xi\right)\right] = 1, \quad \text{a.e. } \xi \in \mathbb{R}^2.$$

Hence, (2.7) holds.

Since $0 < \rho \le 1$, we have $\operatorname{supp}(\widehat{\varphi}) \subseteq [-\pi,\pi]^2$ and $\operatorname{supp}(\widehat{\psi}) \subseteq [-\pi,\pi]^2$. Hence, $\widehat{\varphi}(\xi)\widehat{\varphi}(\xi + 2\mathsf{k}\pi) = 0$ and $\widehat{\psi}(\xi)\widehat{\psi}(\xi + 2\mathsf{k}\pi) = 0$, a.e. $\xi \in \mathbb{R}^2$ and $\mathsf{k} \in \mathbb{Z}^2\backslash\{0\}$. The case that $\widehat{\psi^{j,\pm\ell_{\lambda j}}}(\xi)\widehat{\psi^{j,\pm\ell_{\lambda j}}}(\xi + 2\mathsf{k}\pi) = 0$ a.e. $\xi \in \mathbb{R}^2$ and $\mathsf{k} \in \mathbb{Z}^2\backslash\{0\}$ can be argued in the same way. Hence, (2.8) holds.

Note that all generators are nonnegative. Therefore, by Corollary 1, $\mathsf{AS}(\varphi; \{\Psi_j\}_{j=0}^{\infty})$ with φ and Ψ_j being given as in (2.9) and (2.10) is an affine shear tight frame for $L_2(\mathbb{R}^2)$. □

8.2. Proofs of results in Section 4.1

Proof of Proposition 1. Explicitly, we have

$$\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi) = \begin{cases} 1 & \text{if } |\xi| \le \lambda^{-2}(1-t)\rho\pi, \\ \nu(\frac{-2\lambda^2|\xi| + (2-t)\rho\pi}{t\rho\pi}) & \text{if } \lambda^{-2}(1-t)\rho\pi < |\xi| \le \lambda^{-2}\rho\pi, \\ 0 & \text{otherwise.} \end{cases} \tag{8.1}$$

Hence, by the smoothness of ν and noting $[\frac{d^n}{d\xi^n}\nu(\xi)]|_{\xi=1} = \delta(n)$ we have $\boldsymbol{\alpha}_{\lambda,t,\rho} \in C_c^{\infty}(\mathbb{R})$.

If $1 - t \ge \lambda^{-2}$, by the definition, $\boldsymbol{\beta}_{\lambda,t,\rho}$ can be written as

$$\boldsymbol{\beta}_{\lambda,t,\rho}(\xi) = \begin{cases} \nu(\frac{2\lambda^2|\xi| - (2-t)\rho\pi}{t\rho\pi}) & \text{if } \lambda^{-2}(1-t)\rho\pi \le |\xi| < \lambda^{-2}\rho\pi, \\ 1 & \text{if } \lambda^{-2}\rho\pi \le |\xi| < (1-t)\rho\pi, \\ \nu(\frac{-2|\xi| + (2-t)\rho\pi}{t\rho\pi}) & \text{if } (1-t)\rho\pi \le |\xi| \le \rho\pi, \\ 0 & \text{otherwise.} \end{cases} \tag{8.2}$$

Again, by the smoothness of ν and $[\frac{d^n}{d\xi^n}\nu(\xi)]|_{\xi=1} = \delta(n)$, we have $\boldsymbol{\beta}_{\lambda,t,\rho} \in C_c^{\infty}(\mathbb{R})$.

If $0 \le 1 - t < \lambda^{-2}$, then $\boldsymbol{\beta}_{\lambda,t,\rho}$ is given by

$$\boldsymbol{\beta}_{\lambda,t,\rho}(\xi) = \begin{cases} [(\nu(\frac{-2|\xi| + (2-t)\rho\pi}{t\rho\pi}))^2 - (\nu(\frac{-2\lambda^2|\xi| + (2-t)\rho\pi}{t\rho\pi}))^2]^{1/2} & \text{if } \lambda^{-2}(1-t)\rho\pi \le |\xi| \le \rho\pi, \\ 0 & \text{otherwise.} \end{cases} \tag{8.3}$$

Note that $\widetilde{\nu}(\xi) := (\nu(\frac{-2|\xi| + (2-t)\rho\pi}{t\rho\pi}))^2 - (\nu(\frac{-2\lambda^2|\xi| + (2-t)\rho\pi}{t\rho\pi}))^2 > 0$ for all ξ such that $|\xi| \in (\lambda^{-2}(1-t)\rho\pi, \rho\pi)$. Hence, $\boldsymbol{\beta}_{\lambda,t,\rho}(\xi) = \sqrt{\widetilde{\nu}(\xi)}$ is infinitely differentiable for all $|\xi| \in (\lambda^{-2}(1-t)\rho\pi, \rho\pi)$. For all other ξ such that $|\xi| \notin (\lambda^{-2}(1-t)\rho\pi, \rho\pi)$, all the derivatives of $\widetilde{\nu}(\xi)$ vanish. Then, using the Taylor expansion for $\boldsymbol{\beta}_{\lambda,t,\rho} = \sqrt{\widetilde{\nu}}$, we see that all the derivatives of $\boldsymbol{\beta}_{\lambda,t,\rho}$ vanish for all $|\xi| \notin (\lambda^{-2}(1-t)\rho\pi, \rho\pi)$. Hence, $\boldsymbol{\beta}_{\lambda,t,\rho} \in C_c^{\infty}(\mathbb{R})$.

Therefore, $\boldsymbol{\alpha}_{\lambda,t,\rho}, \boldsymbol{\beta}_{\lambda,t,\rho} \in C_c^{\infty}(\mathbb{R})$. By the definition of $\boldsymbol{\beta}_{\lambda,t,\rho}$, we have $|\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)|^2 + |\boldsymbol{\beta}_{\lambda,t,\rho}(\xi)|^2 = |\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2}\xi)|^2$ for all $\xi \in \mathbb{R}$.

Similar to the cases of $\beta_{\lambda,t,\rho}$, if $1 - t \ge \lambda^{-2}$, then we have

$$\begin{aligned} \boldsymbol{\mu}_{\lambda,t,\rho}(\xi) &= \begin{cases} \boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^2\xi) & \text{if } |\xi| \le \lambda^{-4}\rho\pi, \\ 0 & \text{if } \lambda^{-4}\rho\pi < |\xi| \le \pi, \end{cases} \\ \boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi) &= \begin{cases} \nu(\frac{2\lambda^4|\xi| - (2-t)\rho\pi}{t\rho\pi}) & \text{if } \lambda^{-4}(1-t)\rho\pi \le |\xi| \le \lambda^{-2}\rho\pi, \\ \mathbf{g}_{\lambda,t,\rho}(\xi) & \text{if } \xi \in [-\pi,\pi)\backslash \operatorname{supp} \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\cdot). \end{cases} \end{aligned} \tag{8.4}$$

In this case, obviously, $\boldsymbol{\mu}_{\lambda,t,\rho} \in C^{\infty}(\mathbb{T})$. Note that $[\frac{d^n}{d\xi^n}\nu(\frac{-2\lambda^4|\xi| + (2-t)\rho\pi}{t\rho\pi})]|_{\xi=\pm\lambda^{-2}\rho\pi} = \delta(n)$. By our choice of $\mathbf{g}_{\lambda,t,\rho}$, we see that $\boldsymbol{\upsilon}_{\lambda,t,\rho} \in C^{\infty}(\mathbb{T})$.

If $0 \le 1 - t < \lambda^{-2}$, then we have

$$\boldsymbol{\mu}_{\lambda,t,\rho}(\xi) = \begin{cases} \frac{\nu(\frac{-2\lambda^4|\xi|+(2-t)\rho\pi}{t\rho\pi})}{\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi})} & \text{if } |\xi| \le \lambda^{-4}\rho\pi, \\ 0 & \text{if } \lambda^{-4}\rho\pi < |\xi| \le \pi, \end{cases} \tag{8.5}$$

and

$$\boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi) = \begin{cases} \frac{[(\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi}))^2-(\nu(\frac{-2\lambda^4|\xi|+(2-t)\rho\pi}{t\rho\pi}))^2]^{1/2}}{\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi})} & \text{if } \lambda^{-4}(1-t)\rho\pi \le |\xi| \le \lambda^{-2}\rho\pi, \\ \mathbf{g}_{\lambda,t,\rho}(\xi) & \text{if } \xi \in [-\pi,\pi)\backslash \operatorname{supp} \boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\cdot). \end{cases} \tag{8.6}$$

Note that the function $\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi})$ is strictly positive for all ξ such that $|\xi| \le \lambda^{-4}\rho\pi$. Hence, $\frac{1}{\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi})}$ is infinitely differentiable for all ξ such that $|\xi| \le \lambda^{-4}\rho\pi$. Since ν is C^∞, the product of $\nu(\frac{-2\lambda^4|\xi|+(2-t)\rho\pi}{t\rho\pi})$ and $\frac{1}{\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi})}$ is infinitely differentiable for all ξ such that $|\xi| \le \lambda^{-4}\rho\pi$. For ξ such that $\lambda^{-4}\rho\pi \le |\xi| \le \pi$, all the derivatives of $\boldsymbol{\mu}_{\lambda,t,\rho}(\xi)$ vanish. Consequently, $\boldsymbol{\mu}_{\lambda,t,\rho} \in C^\infty(\mathbb{T})$. Observe that $[(\nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi}))^2 - (\nu(\frac{-2\lambda^4|\xi|+(2-t)\rho\pi}{t\rho\pi}))^2]^{1/2} = \nu(\frac{-2\lambda^2|\xi|+(2-t)\rho\pi}{t\rho\pi}))$ for $\lambda^{-2}(1-t)\rho\pi \le |\xi| \le \lambda^{-2}\rho\pi$. By similar arguments, we conclude that $\boldsymbol{\upsilon}_{\lambda,t,\rho} \in C^\infty(\mathbb{T})$.

Therefore, $\boldsymbol{\mu}_{\lambda,t,\rho}, \boldsymbol{\upsilon}_{\lambda,t,\rho} \in C^\infty(\mathbb{T})$. By their constructions, it is easy to check that $\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^2\xi) = \boldsymbol{\mu}_{\lambda,t,\rho}(\xi)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)$ and $\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^2\xi) = \boldsymbol{\upsilon}_{\lambda,t,\rho}(\xi)\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)$ for $\xi \in \mathbb{R}$. □

Proof of Proposition 2. Since $\boldsymbol{\gamma}_\varepsilon$ is a function in $C_c^\infty(\mathbb{R})$ and the function $f(\xi) := \xi_2/\xi_1$ or ξ_1/ξ_2 is infinitely differentiable for all $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ such that both $\xi_1 \ne 0$ and $\xi_2 \ne 0$, we see that by Taylor expansion $\boldsymbol{\Gamma}^j$ is infinitely differentiable for $\xi \in \mathbb{R}^2$ such that both $\xi_1 \ne 0$ and $\xi_2 \ne 0$. For a fixed $\xi_1 \ne 0$, we have

$$\boldsymbol{\Gamma}^j(\xi) = \sum_{\ell=-\ell_{\lambda^j}+1}^{\ell_{\lambda^j}-1} \left|\boldsymbol{\gamma}_\varepsilon\left(\lambda^j\xi_2/\xi_1 + \ell\right)\right|^2 + \left|\boldsymbol{\gamma}^{\mp}_{\lambda^j,\varepsilon,\varepsilon_0}\left(\lambda^j\xi_2/\xi_1 \pm \ell_{\lambda^j}\right)\right|^2$$

for $|\xi_2|$ small enough in view of the supports of $\boldsymbol{\gamma}_\varepsilon$ and $\boldsymbol{\gamma}_{\lambda^j,\varepsilon,\varepsilon_0}$, which implies that $\boldsymbol{\Gamma}^j(\xi)$ is infinitely differentiable at $(\xi_1, 0)$ with $\xi_1 \ne 0$. Similarly, we have $\boldsymbol{\Gamma}^j(\xi)$ is infinitely differentiable at $(0, \xi_2)$ with $\xi_2 \ne 0$. Hence, we have $\boldsymbol{\Gamma}^j \in C^\infty(\mathbb{R}^2\backslash\{0\})$. By its definition as in (4.9), $\boldsymbol{\Gamma}^j(\mathsf{E}\cdot) = \boldsymbol{\Gamma}^j$ and $\boldsymbol{\Gamma}^j(t\xi) = \boldsymbol{\Gamma}^j(\xi)$ for $t \ne 0$ and $\xi \ne 0$.

By the property of $\boldsymbol{\gamma}_\varepsilon, \boldsymbol{\gamma}_{\lambda,\varepsilon,\varepsilon_0}$ as in (4.7), it is easily seen that $1 \le \boldsymbol{\Gamma}^j(\xi) \le 2$ for $\xi \ne 0$. Now to see that (4.11) holds, we notice that the seamline element on the horizontal cone with respect to $\ell = -\ell_{\lambda^j}$ has part of the piece overlapping with the other cone. By the support of $\boldsymbol{\gamma}^+_{\lambda,\varepsilon,\varepsilon_0}$, for this seamline element, we have $\xi_2/\xi_1 \le 1 + \frac{2\varepsilon_0}{\lambda^{2j}}$. Hence, those elements on the vertical cone with support satisfying $|\xi_1/\xi_2| \le \frac{\lambda^{2j}}{\lambda^{2j}+2\varepsilon_0}$ is not affected by that seamline elements. By symmetry, same result holds for seamline elements on vertical cone affecting the horizontal cone. Therefore, (4.11) holds.

The proof for $\boldsymbol{\Gamma}_j \in C^\infty(\mathbb{R}^2\backslash\{0\})$ is similar to that for $\boldsymbol{\Gamma}^j$. By its definition in (4.10), $\boldsymbol{\Gamma}_j(\mathsf{E}\cdot) = \boldsymbol{\Gamma}_j$ and $\boldsymbol{\Gamma}_j(t\cdot) = \boldsymbol{\Gamma}_j$ for $t \ne 0$.

By the property of γ_ε as in (4.8), it is easily seen that $0 < \boldsymbol{\Gamma}_j \le 2$. Now to see that (4.12) holds, we notice that the seamline element on the horizontal cone with respect to $\ell = -\ell_{\lambda^j}$ has part of the piece overlapping with the other cone. By the support of $\boldsymbol{\gamma}_\varepsilon$, for this seamline element, we have

$$\lambda^j\xi_2/\xi_1 - \ell_{\lambda^j} \le \frac{1}{2} + \varepsilon,$$

which implies $\xi_2/\xi_1 \le \frac{1/2+\varepsilon+\ell_{\lambda^j}}{\lambda^j}$. Hence, it only affects elements in other cone with support satisfying $\xi_1/\xi_2 > \frac{\lambda^j}{1/2+\varepsilon+\ell_{\lambda^j}}$. By symmetry, the same result holds for seamline elements on vertical cone affecting the horizontal cone. Therefore, (4.12) holds. □

8.3. Proofs of results in Section 4.2

Proof of Proposition 3. Since the generators satisfy $\widehat{\varphi}, \boldsymbol{\zeta}^{j,\ell} \in C_c^\infty(\mathbb{R}^2)$ and for any bounded open set $E \subseteq \mathbb{R}^2$, $\boldsymbol{\Theta}^{J_0}(\xi)$ is the summation of finitely many terms from $\widehat{\varphi}, \boldsymbol{\zeta}^{j,\ell}$ for all $\xi \in E$, the function $\boldsymbol{\Theta}^{J_0}$ is thus also a function in $C^\infty(\mathbb{R}^2)$. By its definition, it is obvious that $\boldsymbol{\Theta}^{J_0}(\mathsf{E}\cdot) = \boldsymbol{\Theta}^{J_0}$.

For simplicity of presentation, we denote $\chi_{\{|\xi_2/\xi_1|\le 1\}} := \chi_{\{\xi\in\mathbb{R}^2:|\xi_2/\xi_1|\le 1\}}$ and similar notation applies for others. For $\xi \in \{\xi \in \mathbb{R}^2 : \max\{|\xi_1|, |\xi_2|\} < \lambda^{2J_0}\rho\pi\}$, by (4.7), we have

$$
\begin{aligned}
\boldsymbol{\Theta}^{J_0}(\xi) &\ge \left|\widehat{\varphi}(\mathsf{N}_\lambda^{J_0}\xi)\right|^2 + \sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}} \left(\left|\boldsymbol{\zeta}^{J_0,\ell}(S_\ell \mathsf{B}_\lambda^{J_0}\xi)\right|^2 + \left|\boldsymbol{\zeta}^{J_0,\ell}(S_\ell \mathsf{B}_\lambda^{J_0}\mathsf{E}\xi)\right|^2\right) \\
&\ge \left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2J_0}\xi_1)\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2J_0}\xi_2)\right|^2 + \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2J_0}\xi_1)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) \\
&\quad + \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2J_0}\xi_2)\right|^2\chi_{\{|\xi_2/\xi_1|>1\}}(\xi) > 0,
\end{aligned}
$$

and for $\xi \in \{\xi \in \mathbb{R}^2 : \max\{|\xi_1|, |\xi_2|\} > \lambda^{2J_0-2}\rho\pi\}$, we have

$$
\begin{aligned}
\boldsymbol{\Theta}^{J_0}(\xi) &= \sum_{j=J_0}^{\infty}\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2 + \left|\boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}_\lambda^{j}\mathsf{E}\xi)\right|^2\right) \\
&\ge \sum_{j=J_0}^{\infty}\left[\left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) + \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_2)\right|^2\chi_{\{|\xi_2/\xi_1|>1\}}(\xi)\right] > 0.
\end{aligned}
$$

Consequently, $\boldsymbol{\Theta}^{J_0} > 0$.

We next show that $\boldsymbol{\Theta}^{J_0} \le 2$. Again, by $\boldsymbol{\gamma}_\varepsilon$ as in (4.7), we have

$$
\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left|\boldsymbol{\eta}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) = \left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi), \quad \xi \ne 0
$$

and similarly,

$$
\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left|\boldsymbol{\zeta}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) = \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi), \quad \xi \ne 0.
$$

Hence, for $\xi \ne 0$,

$$
\begin{aligned}
&\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\boldsymbol{\eta}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2 + \left|\boldsymbol{\zeta}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2\right)\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) = \left(\left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\right|^2 + \left|\boldsymbol{\beta}_{\lambda,t,\rho}(\lambda^{-2j}\xi_1)\right|^2\right)\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) \\
&= \left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2(j+1)}\xi_1)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) = \sum_{\ell=-\ell_{\lambda^{j+1}}}^{\ell_{\lambda^{j+1}}} \left|\boldsymbol{\eta}^{j+1,\ell}(S_\ell \mathsf{B}_\lambda^{j+1}\xi)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\lim_{J\to\infty}\left(\sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}} \left|\boldsymbol{\eta}^{J_0,\ell}(S_\ell \mathsf{B}_\lambda^{J_0}\xi)\right|^2 + \sum_{j=J_0}^{J-1}\sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left|\boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}_\lambda^{j}\xi)\right|^2\right)\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) \\
&= \lim_{J\to\infty}\left|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2J}\xi)\right|^2\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) = \chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi), \quad \xi \ne 0.
\end{aligned}
$$

Now, we define

$$\widetilde{\mathbf{\Theta}}^{J_0}(\xi) := \sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}} \left(\left|\boldsymbol{\eta}^{J_0,\ell}\left(S_\ell \mathsf{B}_\lambda^{J_0}\xi\right)\right|^2 + \left|\boldsymbol{\eta}^{J_0,\ell}\left(S_\ell \mathsf{B}_\lambda^{J_0}\mathsf{E}\xi\right)\right|^2\right)$$
$$+ \sum_{j=J_0}^{\infty} \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\boldsymbol{\zeta}^{j,\ell}\left(S_\ell \mathsf{B}_\lambda^{j}\xi\right)\right|^2 + \left|\boldsymbol{\zeta}^{j,\ell}\left(S_\ell \mathsf{B}_\lambda^{j}\mathsf{E}\xi\right)\right|^2\right), \quad \xi \neq 0.$$

Then, for $\xi \neq 0$,

$$\widetilde{\mathbf{\Theta}}^{J_0}(\xi) = \sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}} \left(\left|\boldsymbol{\eta}^{J_0,\ell}\left(S_\ell \mathsf{B}_\lambda^{J_0}\xi\right)\right|^2 + \left|\boldsymbol{\eta}^{J_0,\ell}\left(S_\ell \mathsf{B}_\lambda^{J_0}\mathsf{E}\xi\right)\right|^2\right)\left(\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) + \chi_{\{|\xi_2/\xi_1|>1\}}(\xi)\right)$$
$$+ \sum_{j=J_0}^{\infty} \sum_{\ell=-\ell_{\lambda^j}}^{\ell_{\lambda^j}} \left(\left|\boldsymbol{\zeta}^{j,\ell}\left(S_\ell \mathsf{B}_\lambda^{j}\xi\right)\right|^2 + \left|\boldsymbol{\zeta}^{j,\ell}\left(S_\ell \mathsf{B}_\lambda^{j}\mathsf{E}\xi\right)\right|^2\right)\left(\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi) + \chi_{\{|\xi_2/\xi_1|>1\}}(\xi)\right)$$
$$= 1 + \left(\left|\boldsymbol{\eta}^{J_0,\pm\ell_{\lambda^{J_0}}}\left(S_{\pm\ell_{\lambda^{J_0}}} \mathsf{B}_\lambda^{J_0}\xi\right)\right|^2 + \sum_{j=J_0}^{\infty}\left|\boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}}\left(S_{\pm\ell_{\lambda^j}} \mathsf{B}_\lambda^{j}\xi\right)\right|^2\right)\chi_{\{|\xi_2/\xi_1|>1\}}(\xi)$$
$$+ \left(\left|\boldsymbol{\eta}^{J_0,\pm\ell_{\lambda^{J_0}}}\left(S_{\pm\ell_{\lambda^{J_0}}} \mathsf{B}_\lambda^{J_0}\mathsf{E}\xi\right)\right|^2 + \sum_{j=J_0}^{\infty}\left|\boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}}\left(S_{\pm\ell_{\lambda^j}} \mathsf{B}_\lambda^{j}\mathsf{E}\xi\right)\right|^2\right)\chi_{\{|\xi_2/\xi_1|\le 1\}}(\xi)$$
$$= 1 + I(\xi) + I(\mathsf{E}\xi) - \sqrt{I(\xi)I(\mathsf{E}\xi)},$$

where

$$I(\xi) := \left(\left|\boldsymbol{\eta}^{J_0,\pm\ell_{\lambda^{J_0}}}\left(S_{\pm\ell_{\lambda^{J_0}}} \mathsf{B}_\lambda^{J_0}\xi\right)\right|^2 + \sum_{j=J_0}^{\infty}\left|\boldsymbol{\zeta}^{j,\pm\ell_{\lambda^j}}\left(S_{\pm\ell_{\lambda^j}} \mathsf{B}_\lambda^{j}\xi\right)\right|^2\right)\chi_{\{|\xi_2/\xi_1|\ge 1\}}(\xi), \quad \xi \neq 0.$$

By the construction of $\boldsymbol{\alpha}_{\lambda,t,\rho}$ and $\boldsymbol{\beta}_{\lambda,t,\rho}$, we have $I \le 1$. Therefore, $1 \le \widetilde{\mathbf{\Theta}}^{J_0} \le 2$. Observe that $\mathbf{\Theta}^{J_0}(\xi) \le \widetilde{\mathbf{\Theta}}^{J_0}(\xi) \le 2$ for $\xi \neq 0$ and $\mathbf{\Theta}^{J_0}(0) = 1$, we conclude that item (i) holds.

We next show that item (ii) holds. We have

$$\mathbf{\Theta}^{J_0}(\xi) = \widetilde{\mathbf{\Theta}}^{J_0}(\xi) + \left|\widehat{\varphi}\left(\mathsf{N}_\lambda^{J_0}\xi\right)\right|^2 - \sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}}\left|\boldsymbol{\eta}^{J_0,\ell}\left(S_\ell \mathsf{B}_\lambda^{J_0}\xi\right)\right|^2.$$

Note that $\operatorname{supp}\widehat{\varphi}(\mathsf{N}_\lambda^{J_0}\cdot)$ is inside the support of $\sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}}|\boldsymbol{\eta}^{J_0,\ell}(S_\ell \mathsf{B}_\lambda^{J_0}\cdot)|^2$. Hence, for ξ outside the support of $\sum_{\ell=-\ell_{\lambda^{J_0}}}^{\ell_{\lambda^{J_0}}}|\boldsymbol{\eta}^{J_0,\ell}(S_\ell \mathsf{B}_\lambda^{J_0}\cdot)|^2$, we have $\mathbf{\Theta}^{J_0} = \widetilde{\mathbf{\Theta}}^{J_0}$. By that $\widetilde{\mathbf{\Theta}}^{J_0} = 1 + I + I(\mathsf{E}\cdot) - \sqrt{I \cdot I(\mathsf{E}\cdot)}$, we hence only need to check the overlapping coming from I and $I(\mathsf{E}\cdot)$. In fact, at scale j, the seamline element on the horizontal cone with respect to $\ell = -\ell_{\lambda^j}$ has part of the piece overlapping with the other cone. By the support of $\boldsymbol{\gamma}^+_{\lambda^j,\varepsilon,\varepsilon_0}$, for this seamline element, we have its support satisfying $\xi_2/\xi_1 \le 1 + \frac{2\varepsilon_0}{\lambda^{2j}}$. Moreover, by the support of $\boldsymbol{\beta}_{\lambda,t,\rho}$, this seamline element can only affect other elements in the vertical cone with respect to scales $j_0 = j-1, j, j+1$. Now, the support of the vertical cone element corresponding to scale j_0 and $\ell = -\ell_{\lambda^{j_0}} + s$ with s being a nonnegative integer satisfying

$$\lambda^{j_0}\xi_1/\xi_2 - \ell_{\lambda^{j_0}} + s \le \frac{1}{2} + \varepsilon,$$

which implies $\xi_1/\xi_2 \le \frac{1/2+\varepsilon+\ell_{\lambda^{j_0}}-s}{\lambda^{j_0}}$. Consequently, the seamline element on the horizontal cone affecting the elements in the vertical cone at scale j_0 means $1+\frac{2\varepsilon_0}{\lambda^{2j}} \ge \frac{\lambda^{j_0}}{1/2+\varepsilon+\ell_{\lambda^{j_0}}-s}$, which implies

$$s \le \frac{1}{2}+\varepsilon+\ell_{\lambda^{j_0}} - \frac{\lambda^{2j+j_0}}{\lambda^{2j}+2\varepsilon_0} \le \frac{1}{2}+\varepsilon+\left(\lambda^{j_0}+\frac{1}{2}-\varepsilon\right) - \frac{\lambda^{2j+j_0}}{\lambda^{2j}+2\varepsilon_0}$$
$$\le 1+\frac{2\varepsilon_0\lambda^{j_0}}{\lambda^{2j}+2\varepsilon_0} \le 1+\frac{2\varepsilon_0\lambda^{j+1}}{\lambda^{2j}+2\varepsilon_0} \le 1+\frac{2\varepsilon_0}{\lambda^{j-1}} \le 1+\frac{2\varepsilon_0}{\lambda^{J_0-1}} < 2$$

since $\varepsilon_0 < \frac{\lambda^{J_0-1}}{2}$. Hence, by that s is a nonnegative integer, we deduce that s is either 0 or 1. By symmetry, same result holds for seamline elements on vertical cone affecting the horizontal cone. Therefore, we have $\mathbf{\Theta}^{J_0}(\xi) = \mathbf{\Theta}^{J_0}(\mathsf{E}\xi) = 1$ for ξ in the support of those $\boldsymbol{\zeta}^{j,\ell}(S_\ell \mathsf{B}_\lambda^j \cdot)$ with $|\ell| < \ell_{\lambda^j} - 1$ and $j \ge J_0 + 1$. That is, item (ii) holds. □

Acknowledgments

We would like to thank the reviewers for their valuable comments and helpful suggestions that improved the presentation of the paper.

References

[1] J.-P. Antoine, R. Murenzi, P. Vanderheynst, Directional wavelets revisited: Cauchy wavelets and symmetry detection in pattern, Appl. Comput. Harmon. Anal. 6 (3) (1999) 314–345.
[2] E.J. Candès, D.L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, Comm. Pure Appl. Math. 57 (2) (2004) 219–266.
[3] E.J. Candès, L. Demannet, D. Donoho, L. Ying, Fast discrete curvelet transforms, Multiscale Model. Simul. 5 (3) (2006) 861–899.
[4] C.K. Chui, W. He, J. Stöckler, Compactly supported tight and sibling frames with maximum vanishing moments, Appl. Comput. Harmon. Anal. 13 (3) (2002) 224–262.
[5] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 61, SIAM, Philadelphia, PA, 1992.
[6] I. Daubechies, B. Han, A. Ron, Z. Shen, Framelets: MRA-based constructions of wavelet frames, Appl. Comput. Harmon. Anal. 14 (1) (2003) 1–46.
[7] K. Guo, D. Labate, W. Lim, G. Weiss, E. Wilson, Wavelets with composite dilations, Electron. Res. Announc. Amer. Math. Soc. 10 (2004) 78–87.
[8] K. Guo, G. Kutyniok, D. Labate, Sparse multidimensional representations using anisotropic dilation and shear operators, in: Wavelets and Splines, Athens, GA, 2005, Nashboro Press, Nashville, TN, 2006, pp. 189–201.
[9] K. Guo, D. Labate, W. Lim, G. Weiss, E. Wilson, Wavelets with composite dilations and their MRA properties, Appl. Comput. Harmon. Anal. 20 (2006) 231–249.
[10] K. Guo, D. Labate, Optimal sparse multidimensional representation using shearlets, SIAM J. Math. Anal. 9 (2007) 298–318.
[11] K. Guo, D. Labate, Characterization and analysis of edges using the continuous shearlet transform, SIAM J. Imaging Sci. 2 (2009) 959–986.
[12] K. Guo, D. Labate, Analysis and detection of surface discontinuities using the 3D continuous shearlet transform, Appl. Comput. Harmon. Anal. 30 (2011) 231–242.
[13] K. Guo, D. Labate, The construction of smooth Parseval frames of shearlets, Math. Model. Nat. Phenom. 8 (2013) 82–105.
[14] B. Han, On dual wavelet tight frames, Appl. Comput. Harmon. Anal. 4 (1997) 380–413.
[15] B. Han, Pairs of frequency-based nonhomogeneous dual wavelet frames in the distribution space, Appl. Comput. Harmon. Anal. 29 (2010) 330–353.
[16] B. Han, Nonhomogeneous wavelet systems in high dimensions, Appl. Comput. Harmon. Anal. 32 (2012) 169–196.
[17] B. Han, Properties of discrete framelet transforms, Math. Model. Nat. Phenom. 8 (2013) 18–47.
[18] B. Han, G. Kutyniok, Z. Shen, Adaptive multiresolution analysis structures and shearlet systems, SIAM J. Numer. Anal. 49 (2011) 1921–1946.
[19] B. Han, Z. Zhao, Tensor product complex tight framelets with increasing directionality, SIAM J. Imaging Sci. 7 (2014) 997–1034.
[20] R. Houska, The nonexistence of shearlet scaling functions, Appl. Comput. Harmon. Anal. 32 (1) (2012) 28–44.
[21] E.J. King, G. Kutyniok, X. Zhuang, Analysis of data separation and recovery problems using clustered sparsity, J. Math. Imaging Vision 48 (2014) 205–234.
[22] P. Kittipoom, G. Kutyniok, W.-Q. Lim, Construction of compactly supported shearlet frames, Constr. Approx. 35 (1) (2012) 21–72.

[23] G. Kutyniok, D. Labate, Resolution of the wavefront set using continuous shearlets, Trans. Amer. Math. Soc. 361 (2009) 2719–2754.

[24] G. Kutyniok, D. Labate, Shearlets: Multiscale Analysis for Multivariate Data, Birkhäuser, 2012.

[25] G. Kutyniok, J. Lemvig, W.-Q. Lim, Optimally sparse approximations of 3D functions by compactly supported shearlet frames, SIAM J. Math. Anal. 44 (2012) 2962–3017.

[26] G. Kutyniok, W.-Q. Lim, Compactly supported shearlets are optimally sparse, J. Approx. Theory 163 (11) (2011) 1564–1589.

[27] G. Kutyniok, T. Sauer, Adaptive directional subdivision schemes and shearlet multiresolution analysis, SIAM J. Math. Anal. 41 (2009) 1436–1471.

[28] G. Kutyniok, M. Sharam, X. Zhuang, ShearLab: a rational design of a digital parabolic scaling algorithm, SIAM J. Imaging Sci. 5 (4) (2012) 1291–1332.

[29] W.-Q. Lim, The discrete shearlet transform: a new directional transform and compactly supported shearlet frames, IEEE Trans. Image Process. 19 (5) (2010) 1166–1180.

[30] W.-Q. Lim, Nonseparable shearlet transform, IEEE Trans. Image Process. 22 (2013) 2056–2065.

[31] A. Ron, Z. Shen, Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator, J. Funct. Anal. 148 (2) (1997) 408–447.

[32] I.W. Selesnick, R.G. Baraniuk, N.G. Kingsbury, The dual-tree complex wavelet transform, IEEE Signal Process. Mag. 22 (6) (2005) 123–151.