

ADAPTIVE MULTIREOLUTION ANALYSIS STRUCTURES AND SHEARLET SYSTEMS

BIN HAN*, GITTA KUTYNIOK†, AND ZOUWEI SHEN‡

Abstract. In this paper, we first introduce the concept of an adaptive MRA (AMRA) structure which is a variant of the classical MRA structure suited to the main goal of a fast flexible decomposition strategy adapted to the data at each decomposition level. We then study this novel methodology for the general case of affine-like systems, and derive a Unitary Extension Principle (UEP) for filter design. Finally, we apply our results to the directional representation system of shearlets. This leads to a comprehensive theory for fast decomposition algorithms associated with shearlet systems which encompasses tight shearlet frames with spatially compactly supported generators within such an AMRA structure. Also shearlet-like systems associated with parabolic scaling and unimodular shear matrices are studied within this framework.

Key words. Adaptive multiresolution analysis, affine systems, fast decomposition algorithm, shearlets, tight frames

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

1. Introduction. Wavelets are nowadays indispensable as a multiscale encoding system for a wide range of more theoretically to more practically oriented tasks, since they provide optimal approximation rates for smooth 1-dimensional data exhibiting singularities. The facts that they provide a unified treatment in both the continuous as well as digital setting and that the digital setting admits a multiresolution analysis leading to a fast spatial domain decomposition were essential for their success. It can however be shown that wavelets – although perfectly suited for isotropic structures – do not perform equally well when dealing with anisotropic phenomena.

This fact has motivated the development of various types of directional representation systems for 2-dimensional data that are capable of resolving edge- or curve-like features which precisely separate smooth regions in a sparse way. Such systems include wedgelets [14], bandlets [32], contourlets [12], and curvelets [5, 6, 7]. All multiscale variants of such systems offer different advantages and disadvantages, however, neither of them provides a unified treatment of the continuous and digital settings. Curvelets, for instance, are known to yield tight frames but the digital curvelet transform is not designed within the curvelet-framework and hence, in particular, is not covered by the available theory [4]. We remind the reader that a system X is called a *tight frame* for a Hilbert space \mathcal{H} (in the literature sometimes also referred to as a *Parseval frame* or a *tight frame with bound one*), if $\|f\|^2 = \sum_{g \in X} |\langle f, g \rangle|^2$ holds for all $f \in \mathcal{H}$.

About three years ago, a novel representation system – so-called shearlets – has been proposed [17, 29], which possesses the same favorable approximation and sparsity properties as the other candidates (see [18, 27, 30]) of whom curvelets are perhaps the most advanced ones. One main point in comparison with curvelets is the fact

*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1 (bhan@math.ualberta.ca). B.H. is supported in part by NSERC Canada.

†Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany (kutyniok@math.uni-osnabrueck.de). G.K. would like to thank the Department of Mathematics at the National University of Singapore for its hospitality and support during her visit which enabled completion of a significant part of this paper.

‡Department of Mathematics, National University of Singapore, Singapore 117543 (matzuows@nus.edu.sg). Z.S. is supported in part by Grant R-146-000-113-112 at the National University of Singapore.

that angles are replaced by slopes when parameterizing directions which significantly supports the treating of the digital setting. A second main point is that shearlets fit within the general framework of affine-like systems, which provides an extensive mathematical machinery. Thirdly, it is shown in [28] that shearlets – in addition to the aforementioned favorable properties – provide a unified treatment for the continuous and digital world similar to wavelets.

Recently, several researchers [19, 31, 33] have provided approaches to introduce an MRA structure with accompanying fast spatial domain decomposition for shearlets. This would establish shearlets as the directional representation system which provides the full range of advantageous properties for 2D data which wavelets provide in 1D. However, in the previous approaches, does either the MRA structure not lead to a tight frame, infinitely many filters make an implementation difficult, or the MRA structure is not faithful to the continuum transform. Approaches to extend the present shearlet theory to higher dimensions were also already undertaken, however, for now, only with continuous parameters [9].

In wavelet frame theory, the Unitary Extension Principle (UEP) introduced in [34] has proven to be a highly useful methodology for constructing tight wavelet frames with an associated MRA structure. An MRA structure is in fact crucial for fast decomposition and reconstruction algorithms, which are key ingredients when developing efficient algorithms for wavelet frame based image restorations. This technique was initially used for high and super resolution image reconstruction in [8], with the UEP allowing to design wavelet frame systems adaptive to the problem. The ideas were then further developed to more general image deblurring in [3, 11], blind deblurring in [2], and image inpainting in [1, 16]. For more details on wavelet frames and their applications in image restorations and analysis, we refer the interested reader to the survey articles [13, 36]. One key for guaranteeing success of these applications is the fact that tight wavelet frames can be shown to provide sparse approximations for various classes of image models, in particular, through the redundancy of a frame. Systems which provide even sparser approximations, yet are associated with an MRA structure do presumably lead to more efficient algorithms for image restorations. This motivates our adventures here.

In this paper we aim at providing an MRA structure for tight shearlet frames – and in fact even for more general affine-like systems encompassing different shearlet systems as special cases – which exhibits all the favorable properties of MRA structures for wavelets. We also allow the MRA structure to be more flexible in the sense of adaptivity than ordinarily considered in the literature. We further prove sufficient conditions for such an adaptive MRA in terms of a suited Unitary Extension Principle (UEP) along with a fast spatial domain decomposition as well as approximation properties.

1.1. An Adaptive Multiresolution Analysis (AMRA). The framework of a multiresolution analysis is a well-established methodology in wavelet theory for deriving a decomposition of data into low- and high-frequency parts associated with a scaling function and wavelets which leads to a fast spatial domain decomposition. This point of view needs to be reconsidered when aiming for a decomposition of general (homogeneous) *affine-like systems*, which we coin systems being a subset of unions of *affine systems*

$$\bigcup_{M \in \Lambda} \{ \psi_{M^j; k}^\ell : j \in \mathbb{Z}, k \in \mathbb{Z}^d, \ell = 1, \dots, r \}, \quad \psi^1, \dots, \psi^r \in L_2(\mathbb{R}^d),$$

where Λ is a finite set of $d \times d$ invertible integer matrices, and where, for a function $f : \mathbb{R}^d \rightarrow \mathbb{C}$ and a $d \times d$ invertible real-valued matrix U , we used the notation

$$f_{U;k} := |\det U|^{1/2} f(U \cdot -k), \quad k \in \mathbb{Z}^d. \quad (1.1)$$

Aiming for adaptivity of the decomposition procedure to different types of data and for computational feasibility, we shall further develop the approach initiated in [20, 22, 23, 24] by considering an ‘adaptive MRA’. For this, we claim the following new paradigm for MRAs manifested in the following two requirements:

- (R1) *Nonhomogeneous systems.* If one aims for a fast decomposition algorithm, it is – as observed in [24] – sufficient to study the decomposition algorithm using a generalized filter bank in the framework of nonhomogeneous affine-like systems, i.e., considering a fixed coarsest decomposition level j , say 0, instead of analyzing the limit $j \rightarrow -\infty$. As pointed out in [24], nonhomogeneous affine-like systems do enable a canonical link between fast algorithms in the discrete domain with affine systems in the continuum domain.
- (R2) *Adaptive Filter Selection.* It is not necessary to use only one low-pass filter in a filter bank obtained via the Unitary Extension Principle which does not distinguish between low- and high-pass filters. The use of more than one low-pass filters adaptively allows us to have an adaptive MRA structure to, for instance, achieve directionality while maintaining a fast algorithm in the discrete domain.

The main advantage of (R1) is to naturally connect theoretical continuum considerations to the implementation requirements in the discrete setting, whereas (R2) allows us to change the composition of the filters adaptively and nonstationarily at each decomposition level. Therefore we coin this new paradigm for MRA an *Adaptive Multiresolution Analysis (AMRA)*.

To briefly illustrate the main idea of AMRA, let us discuss one step of the decomposition algorithm. For this, let M_ℓ , $1 \leq \ell \leq r$ be $d \times d$ invertible integer matrices, and let a_ℓ , $\ell = 1, \dots, r$, be finitely supported filters. Further, let s , $1 \leq s \leq r$ be the separator between low- and high-frequency part, and let a_ℓ , $\ell = 1, \dots, r$ be finitely supported low-pass filters. We denote the given data by v which for convenience purposes we now assume to lie in $l_2(\mathbb{Z}^d)$. Notice that from the previous decomposition step we presumably have many such low-frequency coefficients.

We then compute the next level of low-frequency coefficients by

$$v_\ell = \mathcal{T}_{a_\ell, M_\ell} v, \quad \ell = 1, \dots, s,$$

and the next level of high-frequency coefficients by

$$v_\ell = \mathcal{T}_{a_\ell, M_\ell} v, \quad \ell = s + 1, \dots, r,$$

where for any $d \times d$ invertible integer matrix M and finitely supported sequence $a : \mathbb{Z}^d \rightarrow \mathbb{C}$, the *transition operator* $\mathcal{T}_{a, M} : l_2(\mathbb{Z}^d) \rightarrow l_2(\mathbb{Z}^d)$ are defined by

$$[\mathcal{T}_{a, M} v](n) := \sum_{k \in \mathbb{Z}^d} v(k) \overline{a(k - Mn)}.$$

The next step then continues with decomposing v_ℓ , $\ell = 1, \dots, s$. The total number of decomposition steps is J , hence finite.

The reader should notice the requirements (R1) and (R2) as opposed to a ‘classical MRA-decomposition algorithm’. After projecting onto a finite fine scale J , (R1)

has only a finite number J of decomposition steps using an adaptive nonstationary filter bank. The stability of an affine or affine-like system in the continuum setting deals with the asymptotic behavior as the scale $J \rightarrow \infty$ and is a fundamental issue in wavelet analysis. Since we are dealing with tight frames, the stability issue becomes trivial in this paper. (R2) can be seen by the fact that the separator between low- and high-frequency part is somehow ‘loose’ in the sense that at each level a certain condition needs to be satisfied (see Theorem 4.1 (i) below), which does *not* distinguish between those parts. This also implies that the structure of the subspaces the data is projected onto is not as strict as for a classical MRA, but allows also non-orthogonality and non-inclusiveness. Using $s \geq 1$ low-pass filters enables, for instance, to capture the directional singularities in a data adaptively. We caution the reader that though s is called the separator between low- and high-frequency parts, the filters a_1, \dots, a_s are indeed low-pass filters themselves (that is, $\widehat{a}_\ell(0) \neq 0, \ell = 1, \dots, s$), not just symbolically called low-pass filters for superficial generality.

This decomposition algorithm is accompanied by a perfect reconstruction algorithm, which appropriately applies *subdivision operators*, which, for any $d \times d$ invertible integer matrix M and finitely supported sequence $a : \mathbb{Z}^d \rightarrow \mathbb{C}$, are defined by

$$[\mathcal{S}_{a,M}v](n) := |\det(M)| \sum_{k \in \mathbb{Z}^d} v(k)a(n - Mk).$$

Using a similar approach as in [22, 23], we shall prove that perfect reconstruction can be achieved, if and only if, in each decomposition step a generalized version of the Unitary Extension Principle [34] is satisfied by the filters a_ℓ , thereby also leading to a generalized Unitary Extension Principle. The adaptive filter banks will be constructed using the method developed in [20, 21]. Let us also mention that a simple observation made in those papers will allow us to incorporate nonstationarity and adaptivity in our algorithm by using more than one low-pass filters.

The generality of this framework will then allow us to firstly derive an AMRA for shearlets, and secondly the anticipated Shearlet Unitary Extension Principle as a corollary.

1.2. Related Work. Several research teams have previously designed MRA decomposition algorithms based on shearlets: we mention the affine system-based approach [19], the subdivision-based approach [31], and the approach based on separability [33]. However, neither of these approaches did satisfy the desired properties we posed in Subsection 1.1. Further non-MRA based approaches were undertaken, for instance, in [15]. In our opinion, these pioneer efforts demonstrate real progress in directional representation, but further progress is needed to derive an in-all-aspects satisfactory comprehensive study of a fast spatial domain shearlet transform within an appropriate MRA framework with careful attention to mathematical exactness, faithfulness to the continuum transform, and computational feasibility, ideally fulfilling all our desiderata.

A particular credit deserves the work in [31], in which the adaptivity ideas were already lurking. The main difference to this paper is the additional freedom provided by the AMRA structure. We also point out that another completely different approach is recently proposed in [B. Han, *Nonhomogeneous wavelet systems in high dimensions*, arXiv: 1002.2421, 2010], which is the high-dimensional generalization of [24], to achieve adaptivity and directionality in high dimensions with a fast algorithm.

1.3. Contribution of this Paper. The contribution of this paper is three-fold. First, we introduce the concept of an adaptive MRA (AMRA) structure suited to

the main goal of a fast flexible decomposition strategy. Secondly, we study this novel methodology for the general case of affine-like systems. And thirdly, we present a comprehensive theory for shearlet systems which encompasses tight shearlet frames with spatially compactly supported generators within such an AMRA structure along with a fast decomposition strategy.

We wish to mention that, in fact, these results are susceptible of extensive generalizations and extensions most of which are far beyond the scope of this paper and will be studied in future work. Examples are the bi-frame case as well as AMRA structures for shearlet systems in higher dimensions.

1.4. Contents. In Section 2, we introduce the notation that we employ for general affine-like systems, state the fast decomposition algorithm based on an AMRA, and prove the Unitary Extension Principle (UEP) for this situation. Section 3 is concerned with the relation to the continuum setting, i.e., with developing characterizing equations and approximation properties for the functions associated with an ARMA and a general method for constructing associated filters. This general methodology is then applied to the situation of shearlet systems in Section 4 to derive a Unitary Extension Principle for such systems.

2. An Adaptive Multiresolution Analysis for General Affine-Like Systems. As already elaborated upon in the introduction, one main idea of an AMRA is to be able to design each decomposition step adaptively, for instance, dependent on the previous decomposition. To follow this philosophy, in Subsection 2.1, we will firstly analyze one single decomposition step, which might occur at any stage of the general decomposition algorithm. Secondly, in Subsection 2.2, we will then present the large picture in the sense of the complete decomposition procedure.

Transition and subdivision operators are the key operators of decomposition and reconstruction, and our proofs will frequently require their Fourier domain versions. Hence, let us start by stating the following lemma.

LEMMA 2.1. *Let M be a $d \times d$ invertible integer matrix and $a : \mathbb{Z}^d \rightarrow \mathbb{C}$ be a finitely supported sequence. Then*

$$\widehat{\mathcal{S}_{a,M}v}(\xi) = |\det(M)|\hat{v}(M^T\xi)\hat{a}(\xi) \quad (2.1)$$

and

$$\widehat{\mathcal{T}_{a,M}v}(M^T\xi) = |\det(M)|^{-1} \sum_{\omega \in \Omega_M} \hat{v}(\xi + 2\pi\omega)\overline{\hat{a}(\xi + 2\pi\omega)}, \quad (2.2)$$

where

$$\hat{a}(\xi) := \sum_{k \in \mathbb{Z}^d} a(k)e^{-ik \cdot \xi} \quad (2.3)$$

and

$$\Omega_M := [(M^T)^{-1}\mathbb{Z}^d] \cap [0, 1)^d. \quad (2.4)$$

Proof. The two identities in (2.1) and (2.2) are well known in wavelet analysis. For the convenience of the reader, we will prove (2.1), and mention that (2.2) can be easily proved similarly.

We start by denoting $\hat{u}(\xi) := \hat{v}(\xi)\overline{\hat{a}(\xi)}$. By the definition of Fourier series in (2.3) and the property of convolution, it can be easily checked that, for each $n \in \mathbb{Z}^d$, we have $u(n) = \sum_{k \in \mathbb{Z}^d} v(k)\overline{a(k-n)}$. Therefore, by the definition of the transition operator $\mathcal{T}_{a,M}$, we have $[\mathcal{T}_{a,M}v](n) = u(Mn)$ for all $n \in \mathbb{Z}^d$. Hence, we deduce that

$$\widehat{\mathcal{T}_{a,M}v}(\mathbf{M}^T\xi) = \sum_{n \in \mathbb{Z}^d} [\mathcal{T}_{a,M}v](n)e^{-in \cdot \mathbf{M}^T\xi} = \sum_{n \in \mathbb{Z}^d} u(Mn)e^{-iMn \cdot \xi}. \quad (2.5)$$

On the other hand, we have

$$\sum_{\omega \in \Omega_M} \hat{u}(\xi + 2\pi\omega) = \sum_{k \in \mathbb{Z}^d} \sum_{\omega \in \Omega_M} u(k)e^{-ik \cdot (\xi + 2\pi\omega)} = \sum_{k \in \mathbb{Z}^d} u(k)e^{-ik \cdot \xi} \sum_{\omega \in \Omega_M} e^{-ik \cdot 2\pi\omega}.$$

Using the basic fact that $\sum_{\omega \in \Omega_M} e^{-ik \cdot 2\pi\omega} = |\det(\mathbf{M})|$ if $k \in M\mathbb{Z}^d$, and $= 0$ if $k \in \mathbb{Z}^d \setminus [M\mathbb{Z}^d]$, we conclude from the above identity that

$$\sum_{\omega \in \Omega_M} \hat{u}(\xi + 2\pi\omega) = |\det(\mathbf{M})| \sum_{k \in M\mathbb{Z}^d} u(k)e^{-ik \cdot \xi} = |\det(\mathbf{M})| \sum_{n \in \mathbb{Z}^d} u(Mn)e^{-iMn \cdot \xi}.$$

Combining the above identity with (2.5) and noting that $\hat{u}(\xi) = \hat{v}(\xi)\overline{\hat{a}(\xi)}$, we see that (2.1) holds. \square

2.1. A Unitary Extension Principle for One Decomposition Step. Let now $v \in l_2(\mathbb{Z}^d)$ be some set of data. This could be the initial data, but also data after some levels of decomposition then on a renormalized grid. We assume that we are given a sequence of arbitrary $d \times d$ matrices M_ℓ , $1 \leq \ell \leq r$, and finitely supported filters a_ℓ , $\ell = 1, \dots, r$ according to which the data shall be decomposed. Our first result is a Unitary Extension Principle (UEP) for this situation, which characterizes those matrices and filters, which allow perfect reconstruction from the decomposed data using subdivision.

THEOREM 2.2. *Let M_ℓ , $1 \leq \ell \leq r$ be $d \times d$ invertible integer matrices, and let a_ℓ , $\ell = 1, \dots, r$, be finitely supported filters. Then the following perfect reconstruction property holds:*

$$\sum_{\ell=1}^r \mathcal{S}_{a_\ell, M_\ell} \mathcal{T}_{a_\ell, M_\ell} v = v, \quad \forall v \in l_2(\mathbb{Z}^d), \quad (2.6)$$

if and only if, for any $\omega \in \Omega = \bigcup_{\ell=1}^r \Omega_{M_\ell}$, where $\Omega_{M_\ell} := [(M_\ell^T)^{-1}\mathbb{Z}^d] \cap [0, 1)^d$,

$$\sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \hat{a}_\ell(\xi)\overline{\hat{a}_\ell(\xi + 2\pi\omega)} = \delta(\omega). \quad (2.7)$$

Proof. First, notice that, by (2.1) and (2.2), for all $v \in l_2(\mathbb{Z}^d)$, (2.6) is equivalent to

$$\begin{aligned} \hat{v}(\xi) &= \sum_{\ell=1}^r [\mathcal{S}_{a_\ell, M_\ell} \widehat{\mathcal{T}_{a_\ell, M_\ell} v}](\xi) = \sum_{\ell=1}^r |\det(M_\ell)| \widehat{\mathcal{T}_{a_\ell, M_\ell} v}(\mathbf{M}_\ell^T \xi) \hat{a}_\ell(\xi) \\ &= \sum_{\ell=1}^r \sum_{\omega \in \Omega_{M_\ell}} \hat{v}(\xi + 2\pi\omega) \hat{a}_\ell(\xi) \overline{\hat{a}_\ell(\xi + 2\pi\omega)}. \end{aligned}$$

By the definition of Ω_{M_ℓ} in (2.4), we can rewrite the previous equation as

$$\hat{v}(\xi) = \sum_{\omega \in \Omega} \sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \hat{v}(\xi + 2\pi\omega) \widehat{a_\ell(\xi) \overline{a_\ell(\xi + 2\pi\omega)}} \quad \text{for all } v \in l_2(\mathbb{Z}^d). \quad (2.8)$$

Since (2.7) is equivalent to (2.8), this proves (2.7) \Rightarrow (2.6).

We now assume that (2.6) holds, hence (2.8) holds. We aim to prove that this implies (2.8). For this, we let $B_\epsilon(\xi_0)$ denote as usual an open ball around ξ_0 with radius ϵ . We now first observe that for any arbitrarily chosen but then fixed $\omega_0 \in \Omega$ and $\xi_0 \in \mathbb{R}^d$, there exists some $v \in l_2(\mathbb{Z}^d)$ and $\epsilon > 0$ such that

- (a) $\hat{v}(\xi + 2\pi\omega_0) = 1$ for all $\xi \in B_\epsilon(\xi_0)$,
- (b) $\hat{v}(\xi + 2\pi\omega) = 0$ for all $\xi \in B_\epsilon(\xi_0), \omega \in \Omega \setminus \{\omega_0\}$,
- (c) $\text{supp } \hat{v} \subseteq 2\pi\omega_0 + B_{2\epsilon}(\xi_0)$,

simply since Ω is discrete. Now we can conclude that (2.8) implies

$$\hat{v}(\xi) = \sum_{\ell \in \{1 \leq n \leq r : \omega_0 \in \Omega_{M_n}\}} \widehat{a_\ell(\xi) \overline{a_\ell(\xi + 2\pi\omega_0)}} \quad \text{for all } \xi \in B_\epsilon(\xi_0).$$

Hence again by (a) – (c), (2.8) follows for all $\xi \in B_\epsilon(\xi_0)$. Since $\omega_0 \in \Omega$ and $\xi_0 \in \mathbb{R}^d$ are arbitrarily chosen, (2.8) follows. \square

Motivated by a simple fact observed in [20, 21], we now consider a special case of Theorem 2.2 by using a common sampling lattice. The first part of the next result follows immediately from Theorem 2.2 as a special case. It significantly simplifies the condition on the filters imposed by the UEP, provided the situation allows to always use the same sampling lattice. Condition (iii) is the spatial domain expression for the UEP condition (ii), which illustrates the filter design problem in spatial domain. Certainly, this condition could also be stated in the general situation of Theorem 2.2. We however decide to omit this, since the content would be clouded by very technical details.

COROLLARY 2.3. *Let $M_\ell, 1 \leq \ell \leq r$ be $d \times d$ invertible integer matrices satisfying $M_\ell \mathbb{Z}^d = M \mathbb{Z}^d$ for all $1 \leq \ell \leq r$ for some matrix M , and let $a_\ell, \ell = 1, \dots, r$, be finitely supported filters. Then the following conditions are equivalent:*

- (i) for all $v \in l_2(\mathbb{Z}^d)$,

$$\sum_{\ell=1}^r \mathcal{S}_{a_\ell, M_\ell} \mathcal{T}_{a_\ell, M_\ell} v = v;$$

- (ii) for any $\omega \in \Omega_M = [(M^T)^{-1} \mathbb{Z}^d] \cap [0, 1)^d$,

$$\sum_{\ell=1}^r \widehat{a_\ell(\xi) \overline{a_\ell(\xi + 2\pi\omega)}} = \delta(\omega);$$

- (iii) for all $\mathbf{k}, \gamma \in \mathbb{Z}^d$,

$$\sum_{\ell=1}^r \sum_{\mathbf{n} \in \mathbb{Z}^d} a_\ell(\mathbf{k} + M\mathbf{n} + \gamma) \overline{a_\ell(M\mathbf{n} + \gamma)} = |\det(M)|^{-1} \delta(\mathbf{k}).$$

Proof. By our assumption $M_\ell \mathbb{Z}^d = M \mathbb{Z}^d$ and the definition of Ω_M in (2.4), we have $\Omega_{M_\ell} = \Omega_M$ for all $\ell = 1, \dots, r$. The equivalence of (i) and (ii) now follows directly from Theorem 2.2.

We next prove equivalence between (ii) and (iii). By using the definition of \widehat{a}_ℓ , condition (ii) is equivalent to

$$\begin{aligned}\delta(\omega) &= \sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k})} e^{i\mathbf{k} \cdot \boldsymbol{\xi}} \sum_{\mathbf{n} \in \mathbb{Z}^d} a_\ell(\mathbf{n}) e^{-i\mathbf{n} \cdot (\boldsymbol{\xi} + 2\pi\boldsymbol{\omega})} \\ &= \sum_{\ell=1}^r \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k})} a_\ell(\mathbf{n}) e^{i(\mathbf{k}-\mathbf{n}) \cdot \boldsymbol{\xi}} e^{-i\mathbf{n} \cdot 2\pi\boldsymbol{\omega}}.\end{aligned}\tag{2.9}$$

Next we denote $\Gamma_{\mathbf{M}} := (\mathbf{M}[0, 1)^d) \cap \mathbb{Z}^d$. Using the trivial relation $\mathbb{Z}^d = \Gamma_{\mathbf{M}} + \mathbf{M}\mathbb{Z}^d$ and replacing \mathbf{n} in (2.9) by $\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}$, we can rewrite (2.9) as

$$\begin{aligned}\delta(\omega) &= \sum_{\ell=1}^r \sum_{\boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k})} a_\ell(\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}) e^{i(\mathbf{k}-\mathbf{M}\mathbf{n}-\boldsymbol{\gamma}) \cdot \boldsymbol{\xi}} e^{-i(\mathbf{M}\mathbf{n}+\boldsymbol{\gamma}) \cdot 2\pi\boldsymbol{\omega}} \\ &= \sum_{\ell=1}^r \sum_{\boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k} + \mathbf{M}\mathbf{n} + \boldsymbol{\gamma})} a_\ell(\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}) e^{i\mathbf{k} \cdot \boldsymbol{\xi}} e^{-i\boldsymbol{\gamma} \cdot 2\pi\boldsymbol{\omega}}.\end{aligned}$$

Considering this equation in matrix form

$$(e^{-i\boldsymbol{\gamma} \cdot 2\pi\boldsymbol{\omega}})_{\boldsymbol{\omega} \in \Omega_{\mathbf{M}}, \boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} \left(\sum_{\ell=1}^r \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k} + \mathbf{M}\mathbf{n} + \boldsymbol{\gamma})} a_\ell(\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}) e^{i\mathbf{k} \cdot \boldsymbol{\xi}} \right)_{\boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} = (\delta(\omega))_{\boldsymbol{\omega} \in \Omega_{\mathbf{M}}},$$

we can conclude by taking the inverse that

$$\begin{aligned}&\left(\sum_{\ell=1}^r \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k} + \mathbf{M}\mathbf{n} + \boldsymbol{\gamma})} a_\ell(\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}) e^{i\mathbf{k} \cdot \boldsymbol{\xi}} \right)_{\boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} \\ &= |\det(\mathbf{M})|^{-1} (e^{i\boldsymbol{\gamma} \cdot 2\pi\boldsymbol{\omega}})_{\boldsymbol{\omega} \in \Omega_{\mathbf{M}}, \boldsymbol{\gamma} \in \Gamma_{\mathbf{M}}} (\delta(\omega))_{\boldsymbol{\omega} \in \Omega_{\mathbf{M}}} \\ &= |\det(\mathbf{M})|^{-1} (1, \dots, 1)^\top.\end{aligned}$$

Thus (ii) is equivalent to the equation

$$\sum_{\ell=1}^r \sum_{\mathbf{k}, \mathbf{n} \in \mathbb{Z}^d} \overline{a_\ell(\mathbf{k} + \mathbf{M}\mathbf{n} + \boldsymbol{\gamma})} a_\ell(\mathbf{M}\mathbf{n} + \boldsymbol{\gamma}) e^{i\mathbf{k} \cdot \boldsymbol{\xi}} = |\det(\mathbf{M})|^{-1} \quad \text{for all } \boldsymbol{\gamma} \in \mathbb{Z}^d,$$

which in turn is equivalent to (iii). \square

This corollary implies, by defining the equivalence relation \sim on the $d \times d$ invertible integer matrices by $\mathbf{M}_1 \sim \mathbf{M}_2 \iff \mathbf{M}_1 \mathbb{Z}^d = \mathbf{M}_2 \mathbb{Z}^d$, that only one representative of each involved class needs to satisfy (2.7). In other words, only the generated lattice $\mathbf{M}_1 \mathbb{Z}^d$ does matter in the condition equivalent to perfect reconstruction. Inside each class we have very much freedom to choose the dilation matrices as necessary by the application. We refer the reader to [20, 21] for the application of this observation on the construction of wavelets.

2.2. Fast Decomposition Algorithm. In the previous subsection, we derived characterizing conditions for sequences of arbitrary $d \times d$ matrices \mathbf{M}_ℓ , $1 \leq \ell \leq r$, and finitely supported filters a_ℓ , $\ell = 1, \dots, r$ which allow perfect reconstruction in each level of an AMRA. This now enables us to present a general adaptive decomposition

algorithm, where the matrices and filters in each level can be chosen according to Theorem 2.2. The adaptivity of our algorithm is essentially built on the multilevel decomposition and reconstruction in a nonstationary fashion. Though multilevel and nonstationarity will unavoidably make the presentation and notation of our algorithm more challenging, it worths our effort to present explicitly our algorithm using tree structure. This also allows the reader and us to follow our description here for a practical implementation of our algorithm in a near future.

We first require some notation to carefully keep track of the decomposition steps and positions in the generated tree structure by suitable indexes. While reading the definitions, we recommend the reader to also take a look at Figure 2.1, which illustrates the indexing of the decomposition. To reflect the multilevel and nonstationarity following the commonly used tree structure in computer science, we shall use vectors $(\beta_1, \dots, \beta_j)$ assigned to matrices, filters, and data, where each entry β_j indicates whether this object is related to the computation of scale j or a coarser scale, and if yes, whether the data at scale j to reach this object was generated by a low-pass or high-pass filter.

Let us now be more specific. Recall that $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The original data is assigned the index $\mathbf{0} := (0, \dots, 0) \in \mathbb{N}_0^J$, and we set $\mathcal{L}_0^L = \{\mathbf{0}\} = \{(0, \dots, 0)\}$. For the indexing of the low-pass filters, we define $\mathcal{I}_0^L = \{1, \dots, s_0\}$ and, for the high-pass filters, $\mathcal{I}_0^H = \{s_0 + 1, \dots, r_0\}$, where $s_0 \leq r_0$ and r_0 is a positive integer. In that sense s_0 partitions the filters into low- and high-frequency filters. If $s_0 = r_0$, then $\mathcal{I}_0^H = \emptyset$. Thus, matrices and filters in the first step of the decomposition are labeled by $(\ell, 0, \dots, 0) \in \mathbb{N}_0^J$, $\ell \in \mathcal{I}_0^L$ for the low-frequency objects, and $(\ell, 0, \dots, 0) \in \mathbb{N}_0^J$, $\ell \in \mathcal{I}_0^H$ for the high-frequency objects. For our convenience, we further introduce the notation

$$\mathcal{L}_1^L = \mathcal{L}_{1, \mathbf{0}}^L = \{(\ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}_0^L\}$$

and

$$\mathcal{L}_1^H = \mathcal{L}_{1, \mathbf{0}}^H = \{(\ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}_0^H\}$$

to denote the indices for the low- and high-frequency objects in the first decomposition layer. The associated filters $a_\beta, \beta \in \mathcal{L}_1^L \cup \mathcal{L}_1^H$ for the first decomposition step are required to satisfy condition (ii) in Theorem 2.2, i.e.,

$$\sum_{\beta \in \{\gamma \in \mathcal{L}_1^L \cup \mathcal{L}_1^H : \omega \in \Omega_{M_\gamma}\}} \widehat{a_\beta}(\xi) \overline{\widehat{a_\beta}(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \Omega_0 = \Omega_{(0, \dots, 0)},$$

where M_γ are $d \times d$ invertible integer matrices, $\Omega_{M_\gamma} := [(M_\gamma^T)^{-1}\mathbb{Z}^d] \cap [0, 1)^d$, and $\Omega_{(0, \dots, 0)} := \cup_{\gamma \in \mathcal{L}_1^L \cup \mathcal{L}_1^H} \Omega_{M_\gamma}$.

The matrices and filters which are used in the j th step will then be labeled as follows. We first assume that in the $(j-1)$ th step the sets \mathcal{L}_{j-1}^L and \mathcal{L}_{j-1}^H were already constructed. Then, for some $\beta = (\beta_1, \dots, \beta_{j-1}, 0, \dots, 0) \in \mathcal{L}_{j-1}^L$, we define

$$\mathcal{I}_\beta^L = \{1, \dots, s_\beta\} \quad \text{and} \quad \mathcal{I}_\beta^H = \{s_\beta + 1, \dots, r_\beta\} \quad \text{with} \quad 1 \leq s_\beta \leq r_\beta,$$

as single labels in the new decomposition layer for the matrices and filters. To take the whole tree structure into account, we now define the set of new low-pass related indices arising from $\beta = (\beta_1, \dots, \beta_{j-1}, 0, \dots, 0)$ by

$$\mathcal{L}_{j, \beta}^L = \{(\beta_1, \dots, \beta_{j-1}, \ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}_\beta^L\}$$

and the same for the set of all high-pass related indices,

$$\mathcal{L}_{j,\beta}^H = \{(\beta_1, \dots, \beta_{j-1}, \ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}_\beta^H\}.$$

We further set

$$\mathcal{L}_j^L = \cup_{\beta \in \mathcal{L}_{j-1}^L} \mathcal{L}_{j,\beta}^L \quad \text{and} \quad \mathcal{L}_j^H = \cup_{\beta \in \mathcal{L}_{j-1}^L} \mathcal{L}_{j,\beta}^H,$$

the complete set of low-pass and high-pass indices in step j , respectively. Now let $\alpha \in \mathcal{L}_{j-1}^L$. To ensure perfect reconstruction, according to Theorem 2.2, the associated filters a_β , $\beta \in \mathcal{L}_{j,\alpha}^L \cup \mathcal{L}_{j,\alpha}^H$ must satisfy

$$\sum_{\beta \in \{\gamma \in \mathcal{L}_{j,\alpha}^L \cup \mathcal{L}_{j,\alpha}^H : \omega \in \Omega_{M_\gamma}\}} \widehat{a}_\beta(\xi) \overline{\widehat{a}_\beta(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \Omega_\alpha,$$

where $\Omega_\alpha := \cup_{\gamma \in \mathcal{L}_{j,\alpha}^L \cup \mathcal{L}_{j,\alpha}^H} \Omega_{M_\gamma}$.

Next we describe the general multi-level decomposition algorithm explicitly using the indexing that we just introduced. For illustrative purposes, before presenting the complete algorithm, we display the first decomposition step as well as one part of the second decomposition step in Figure 2.1.

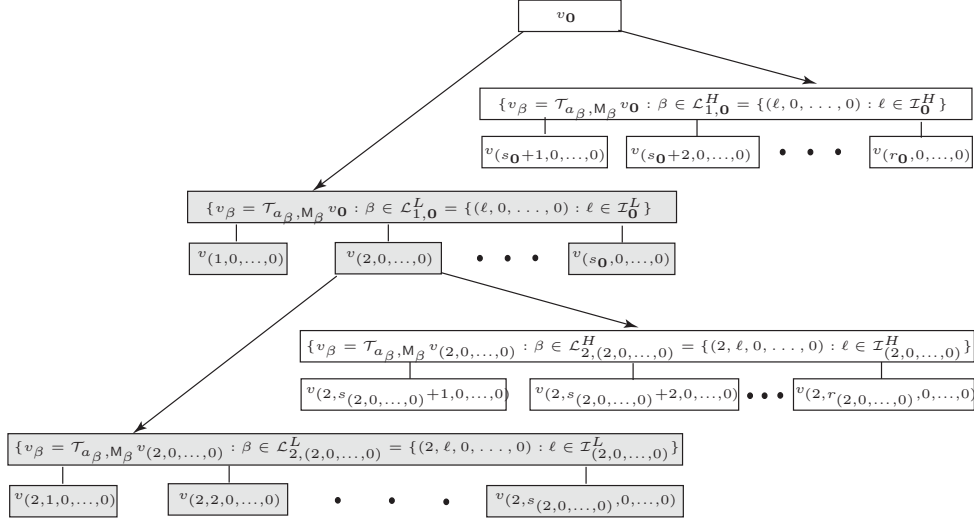


FIG. 2.1. The decomposition structure of (FAD) illustrated through the first complete step and the second step shown exemplarily through the decomposition of $v_{(2,0,\dots,0)}$. The low-frequency components, which will be processed further, are gray-shaded.

In Figure 2.2, we now describe the general multi-level decomposition algorithm explicitly. This decomposition can be implemented as the usual fast wavelet transform with a tree structure. We remind the reader of the introduced notion of transition and subdivision in Subsection 1.1.

The fact that the filters are chosen to be perfect reconstruction filters allows us to reconstruct the data by application of appropriate subdivision operators as displayed in Figure 2.3.

3. Relation with Continuum Setting. We next intend to relate the ‘digital conditions’ on the filters to ‘continuum conditions’ for associated function systems, in particular, frame systems.

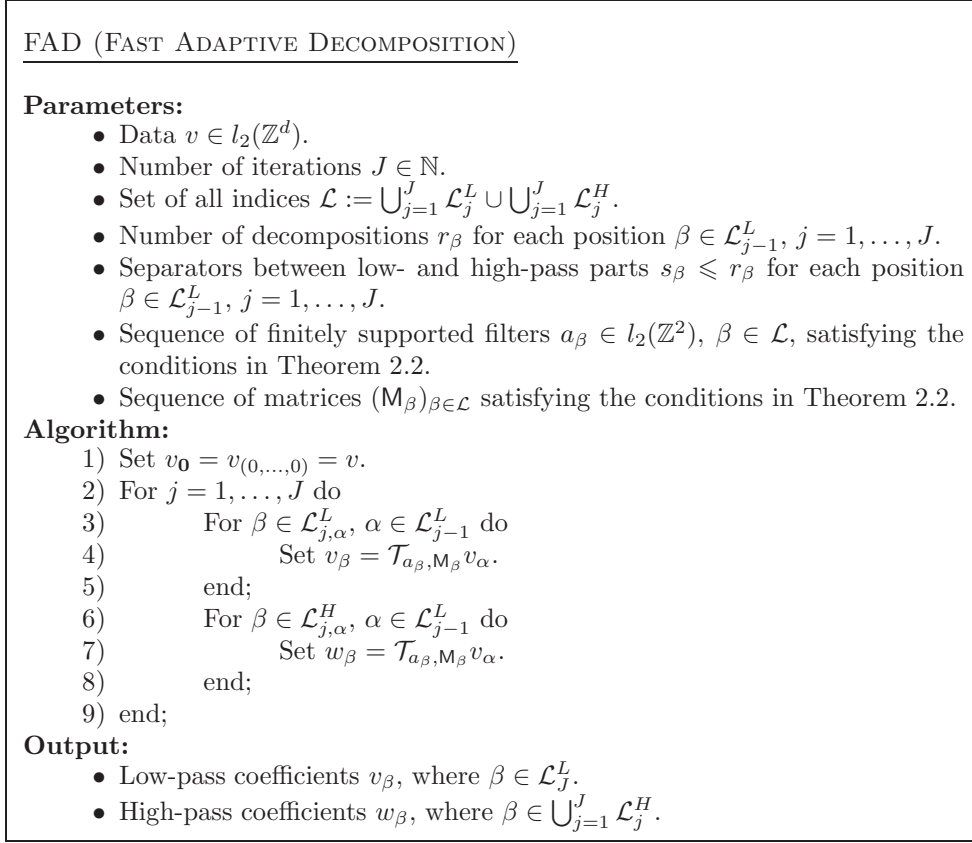


FIG. 2.2. The FAD Algorithm for a fast adaptive decomposition using affine-like systems.

3.1. Characterization Equations. As before in Theorem 2.2, we only consider one level. The conditions we will derive then need to be satisfied for each step in the iteration, and the choice of a non-stationary or stationary scheme is left to the user.

For a function $f : \mathbb{R}^d \rightarrow \mathbb{C}$ and an invertible $d \times d$ matrix U , we shall use the notation as in (1.1).

We can now formulate the condition on perfect reconstruction in terms of a condition in the function setting. Notice that here we again consider the most general situation of Theorem 2.2 as opposed to Corollary 2.3.

THEOREM 3.1. *Let M_ℓ , $0 \leq \ell \leq r$ be $d \times d$ invertible integer matrices, and let a_ℓ , $\ell = 1, \dots, r$, be finitely supported filters. Let ϕ be a nontrivial compactly supported function in $L_2(\mathbb{R}^d)$. Then the following conditions are equivalent:*

(i) For all $v \in l_2(\mathbb{Z}^d)$,

$$\sum_{\ell=1}^r \mathcal{S}_{a_\ell, M_\ell} \mathcal{T}_{a_\ell, M_\ell} v = v;$$

(ii) For each $f, g \in L_2(\mathbb{R}^d)$,

$$\sum_{k \in \mathbb{Z}^d} \langle f, \phi_{M_0; k} \rangle \langle \phi_{M_0; k}, g \rangle = \sum_{\ell=1}^r \sum_{k \in \mathbb{Z}^d} \langle f, \psi_{M_\ell^{-1} M_0; k}^\ell \rangle \langle \psi_{M_\ell^{-1} M_0; k}^\ell, g \rangle, \quad (3.1)$$

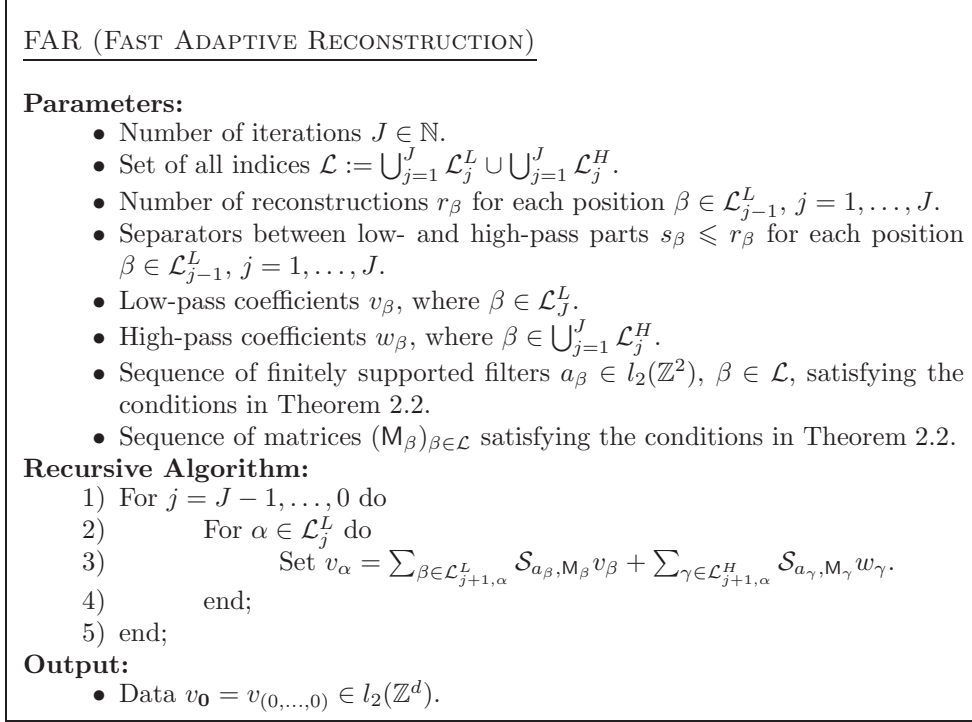


FIG. 2.3. The FAR Algorithm for a fast adaptive reconstruction using affine-like systems.

where $\phi_{M_0; \mathbf{k}}$ is defined as in (1.1) and ψ^1, \dots, ψ^r are defined by $\widehat{\psi^\ell}(M_\ell^\top \xi) := \widehat{a_\ell}(\xi) \widehat{\phi}(\xi)$, that is,

$$\psi^\ell := |\det(M_\ell)| \sum_{\mathbf{k} \in \mathbb{Z}^d} a_\ell(\mathbf{k}) \phi(M_\ell \cdot -\mathbf{k}).$$

Proof. We first recall that (i) is equivalent to (2.7) by Theorem 2.2. The strategy will now be to show that (ii) is equivalent to (2.7).

For $f, g, \eta \in L_2(\mathbb{R}^d)$ and an invertible real-valued matrix U , we have

$$\langle f, \eta_{U; \mathbf{k}} \rangle = \langle |\det(U)|^{-1/2} f(U^{-1} \cdot), \eta(\cdot - \mathbf{k}) \rangle = \langle f_{U^{-1}; 0}, \eta_{I_d; \mathbf{k}} \rangle. \quad (3.2)$$

It is now easy to see that (3.1) holds if and only if it holds for $M_0 = I_d$. So, we assume that $M_0 = I_d$ in the following proof. Using the Fourier-based approach (e.g. [24, Lemma 3]) and Parseval identity, we have the following known identity:

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \eta_{I_d; \mathbf{k}} \rangle \langle \eta_{I_d; \mathbf{k}}, g \rangle = (2\pi)^d \int_{\mathbb{R}^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi \mathbf{k})} \widehat{\eta}(\xi) \widehat{\eta}(\xi + 2\pi \mathbf{k}) d\xi \quad (3.3)$$

for all $f, g, \eta \in L_2(\mathbb{R}^d)$. In particular, (3.3) holds if $\eta = \phi$. Moreover, applying (3.2)

and (3.3), we have

$$\begin{aligned}
\sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{M_\ell^{-1}; \mathbf{k}}^\ell \rangle \langle \psi_{M_\ell^{-1}; \mathbf{k}}^\ell, g \rangle &= \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f_{M_\ell; 0}, \psi_{I_d; \mathbf{k}}^\ell \rangle \langle \psi_{I_d; \mathbf{k}}^\ell, g_{M_\ell; 0} \rangle \\
&= (2\pi)^d \int_{\mathbb{R}^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f_{M_\ell; 0}}(\xi) \overline{\widehat{g_{M_\ell; 0}}(\xi + 2\pi\mathbf{k})} \widehat{\psi^\ell}(\xi) \widehat{\psi^\ell}(\xi + 2\pi\mathbf{k}) d\xi \\
&= (2\pi)^d |\det(M_\ell)|^{-1} \int_{\mathbb{R}^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}((M_\ell^\top)^{-1}\xi) \overline{\widehat{g}((M_\ell^\top)^{-1}(\xi + 2\pi\mathbf{k}))} \widehat{\psi^\ell}(\xi) \widehat{\psi^\ell}(\xi + 2\pi\mathbf{k}) d\xi.
\end{aligned}$$

Changing from the variable ξ to $M_\ell \xi$ in the last identity and summing over ℓ , we end up with

$$\begin{aligned}
&\sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{M_\ell^{-1}; \mathbf{k}}^\ell \rangle \langle \psi_{M_\ell^{-1}; \mathbf{k}}^\ell, g \rangle \\
&= (2\pi)^d \int_{\mathbb{R}^d} \sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi(M_\ell^\top)^{-1}\mathbf{k})} \widehat{\psi^\ell}(M_\ell^\top \xi) \widehat{\psi^\ell}(M_\ell^\top \xi + 2\pi\mathbf{k}) d\xi.
\end{aligned} \tag{3.4}$$

Since $\widehat{\psi^\ell}(M_\ell^\top \xi) = \widehat{a}_\ell(\xi) \widehat{\phi}(\xi)$, we have

$$\widehat{\psi^\ell}(M_\ell^\top \xi) \widehat{\psi^\ell}(M_\ell^\top \xi + 2\pi\mathbf{k}) = \overline{\widehat{a}_\ell(\xi)} \widehat{a}_\ell(\xi + 2\pi(M_\ell^\top)^{-1}\mathbf{k}) \widehat{\phi}(\xi) \widehat{\phi}(\xi + 2\pi(M_\ell^\top)^{-1}\mathbf{k}).$$

Note that \mathbb{Z}^d is the disjoint union of $M_\ell^\top \omega + M_\ell^\top \mathbb{Z}^d$ for every $\omega \in \Omega_{M_\ell}$. Now we deduce that (3.4) becomes

$$\begin{aligned}
&\sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{M_\ell^{-1}; \mathbf{k}}^\ell \rangle \langle \psi_{M_\ell^{-1}; \mathbf{k}}^\ell, g \rangle \\
&= (2\pi)^d \int_{\mathbb{R}^d} \sum_{\ell=1}^r \sum_{\omega \in \Omega_{M_\ell}} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi\omega + 2\pi\mathbf{k})} \overline{\widehat{a}_\ell(\xi)} \widehat{a}_\ell(\xi + 2\pi\omega) \\
&\quad \cdot \widehat{\phi}(\xi) \widehat{\phi}(\xi + 2\pi\omega + 2\pi\mathbf{k}) d\xi \\
&= (2\pi)^d \int_{\mathbb{R}^d} \sum_{\omega \in \Omega} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi\omega + 2\pi\mathbf{k})} \widehat{\phi}(\xi) \widehat{\phi}(\xi + 2\pi\omega + 2\pi\mathbf{k}) \\
&\quad \cdot \sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \overline{\widehat{a}_\ell(\xi)} \widehat{a}_\ell(\xi + 2\pi\omega) d\xi,
\end{aligned}$$

where $\Omega := \cup_{\ell=1}^r \Omega_{M_\ell}$ as in Theorem 2.2. Hence, (3.1) is equivalent to

$$\begin{aligned}
&\int_{\mathbb{R}^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi\mathbf{k})} \widehat{\phi}(\xi) \widehat{\phi}(\xi + 2\pi\mathbf{k}) d\xi \\
&= \int_{\mathbb{R}^d} \sum_{\omega \in \Omega} \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi + 2\pi\omega + 2\pi\mathbf{k})} \widehat{\phi}(\xi) \widehat{\phi}(\xi + 2\pi\omega + 2\pi\mathbf{k}) \\
&\quad \cdot \sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \overline{\widehat{a}_\ell(\xi)} \widehat{a}_\ell(\xi + 2\pi\omega) d\xi.
\end{aligned} \tag{3.5}$$

Since (2.7) holds, it is obvious that the above identity in (3.5) holds and therefore, (iii) holds.

Conversely, if (ii) holds, then (3.5) holds. Now we use a similar argument as in [24, Lemma 5] to prove that this implies condition (ii). For this, we note that $\Omega + \mathbb{Z}^d := \{\omega + \mathbf{k} : \omega \in \Omega, \mathbf{k} \in \mathbb{Z}^d\}$ is a discrete set without any accumulation point. For any arbitrarily chosen but then fixed $\omega_0 \in \Omega$ and $\xi_0 \in \mathbb{R}^d$, we can take any functions $f, g \in L_2(\mathbb{R}^d)$ such that the support of \hat{f} is contained inside $B_\epsilon(\xi_0)$ and the support of \hat{g} is contained inside $B_\epsilon(\xi + 2\pi\omega_0)$. As long as ϵ is small enough, it is not difficult to see that

$$\hat{f}(\xi)\overline{\hat{g}(\xi + 2\pi\mathbf{k})} = 0 \quad \text{for all } \mathbf{k} \in [\Omega + \mathbb{Z}^d] \setminus \{\omega_0\}.$$

Now it is easy to see that (3.5) becomes

$$\begin{aligned} & \delta(\omega_0) \int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\overline{\hat{\phi}(\xi)}\hat{\phi}(\xi)d\xi \\ &= \int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi + 2\pi\omega_0)}\overline{\hat{\phi}(\xi)}\hat{\phi}(\xi + 2\pi\omega_0) \sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \overline{\hat{a}_\ell(\xi)}\hat{a}_\ell(\xi + 2\pi\omega_0)d\xi. \end{aligned}$$

Note that $\hat{\phi}(\xi) \neq 0$ for almost every $\xi \in \mathbb{R}^d$. Since the above identity holds for all functions f, g in $L_2(\mathbb{R}^d)$ as long as the support of \hat{f} is contained inside $B_\epsilon(\xi_0)$ and the support of \hat{g} is contained inside $B_\epsilon(\xi + 2\pi\omega_0)$, we now easily deduce from the above identity that $\sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{M_n}\}} \overline{\hat{a}_\ell(\xi)}\hat{a}_\ell(\xi + 2\pi\omega_0) = \delta(\omega_0)$ for almost every $\xi \in B_\epsilon(\xi_0)$. Since $\omega_0 \in \Omega$ and $\xi_0 \in \mathbb{R}^d$ can be arbitrarily chosen, we now conclude that (2.7) holds.

The theorem is proved. \square

If $r = 1$ and $M_1 = I_d$ in Theorem 3.1, then we must have $\hat{a}_1 = 1$ (that is, $a_1 = \delta$) and consequently for this particular case, $\mathcal{T}_{a_1, I_d}v = v$ and $\mathcal{S}_{a_1, I_d}v = |\det(M)|v$, that is, up to a multiplicative constant, the data is just copied.

3.2. Tight Frame Structure. For any positive integer J , we now construct a tight affine-like frame \mathcal{AS}_J in $L_2(\mathbb{R}^d)$ corresponding to the decomposition and reconstruction algorithms, (FAD) (Figure 2.2) and (FAR) (Figure 2.3), respectively. Let a be a filter on \mathbb{Z}^d and M_0 be a $d \times d$ dilation matrix, for example, $M_0 = 2I_d$ and a is a tensor product filter. We assume that

$$\hat{\phi}(\xi) := \prod_{j=1}^{\infty} \hat{a}((M_0^T)^{-j}\xi), \quad \xi \in \mathbb{R}^d$$

is a well-defined function in $L_2(\mathbb{R}^d)$ and we also assume that there exist filters b_1, \dots, b_r such that

$$\hat{a}(\xi)\overline{\hat{a}(\xi + 2\pi\omega)} + \sum_{\ell=1}^r \hat{b}_\ell(\xi)\overline{\hat{b}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \Omega_{M_0}.$$

Based on these, we then define

$$\widehat{\psi}^\ell(M_0^T\xi) := \hat{b}_\ell(\xi)\hat{\phi}(\xi), \quad \ell = 1, \dots, r.$$

Then the system

$$WS(\psi^1, \dots, \psi^r) := \{\psi_{M_0^j; \mathbf{k}}^1, \dots, \psi_{M_0^j; \mathbf{k}}^r : j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^d\}$$

forms a tight frame in $L_2(\mathbb{R}^d)$ by the Unitary Extension Principle of [34] (see also [10, 21]).

Next, we use this standard tight frame system to derive a new tight affine-like frame \mathcal{AS}_J in $L_2(\mathbb{R}^d)$ corresponding to the decomposition and reconstruction algorithms, (FAD) (Figure 2.2) and (FAR) (Figure 2.3), respectively.

First, denote

$$\psi^{(0,\dots,0)} = \phi.$$

Then, let $\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H$ with $\beta \in \mathcal{L}_{J,\beta^{(1)}}^L$, $\beta^{(1)} \in \mathcal{L}_{J-1}^L$, or $\beta \in \mathcal{L}_{j,\beta^{(1)}}^H$, $\beta^{(1)} \in \mathcal{L}_{j-1}^L$ for some $j \in \{1, \dots, J\}$. Then we recursively define affine functions by

$$\widehat{\psi^\beta}(\mathbf{M}_\beta^T \xi) := \widehat{a_\beta}(\xi) \widehat{\psi^{\beta^{(1)}}}(\xi).$$

Further, let $\beta^{(2)} \in \mathcal{L}_{j-2,\mu^{(2)}}^L$, etc. Using this sequence, we recursively define matrices by

$$\mathbf{N}_\beta := \mathbf{M}_\beta^{-1} \mathbf{M}_{\beta^{(1)}}^{-1} \cdots \mathbf{M}_{\beta^{(j-1)}}^{-1} \mathbf{M}_{\beta^{(j)}}^{-1}.$$

We now define the associated affine-like system in the following way:

DEFINITION 3.2. *Retaining the introduced notions and definitions, the affine-like system \mathcal{AS}_J is defined by*

$$\mathcal{AS}_J := \{\psi_{\mathbf{M}_0^j; \mathbf{k}}^1, \dots, \psi_{\mathbf{M}_0^j; \mathbf{k}}^r : j \geq J, \mathbf{k} \in \mathbb{Z}^d\} \cup \{\psi_{\mathbf{N}_\beta \mathbf{M}_0^j; \mathbf{k}}^\beta : \mathbf{k} \in \mathbb{Z}^d, \beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H\},$$

where we also employ the notation introduced in (1.1).

We next show that this system – although in general not forming an orthonormal basis – still always constitutes a tight frame.

THEOREM 3.3. *For any positive integer J , \mathcal{AS}_J is a tight frame for $L_2(\mathbb{R}^d)$.*

Proof. By Theorem 3.1, it is not difficult to deduce that

$$\sum_{\beta \in \mathcal{L}_J^L \cup \mathcal{L}_J^H} \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^J; \mathbf{k}}^\beta \rangle|^2 = \sum_{\beta \in \mathcal{L}_{J-1}^L} \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^J; \mathbf{k}}^\beta \rangle|^2.$$

Now from the above relation, we see that

$$\sum_{\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H} \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^J; \mathbf{k}}^\beta \rangle|^2 = \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{M}_0^J; \mathbf{k}}^{(0,\dots,0)} \rangle|^2 = \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \phi_{\mathbf{M}_0^J; \mathbf{k}} \rangle|^2.$$

By our given assumption on ϕ and ψ^1, \dots, ψ^r , it is known that $\{\phi_{\mathbf{M}_0^j; \mathbf{k}} : \mathbf{k} \in \mathbb{Z}^d\} \cup \{\psi_{\mathbf{M}_0^j; \mathbf{k}}^\ell : j \geq J, \mathbf{k} \in \mathbb{Z}^d, \ell = 1, \dots, r\}$ is a tight frame for $L_2(\mathbb{R}^d)$. Consequently,

$$\sum_{\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H} \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^J; \mathbf{k}}^\beta \rangle|^2 + \sum_{j=J}^{\infty} \sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, \psi_{\mathbf{M}_0^j; \mathbf{k}}^\ell \rangle|^2 = \|f\|_{L_2(\mathbb{R}^d)}^2.$$

Hence, \mathcal{AS}_J is a tight frame for $L_2(\mathbb{R}^d)$. \square

Next we analyze the approximation order of the affine-like system \mathcal{AS}_J . For this, for $\tau \geq 0$, we recall that $H^\tau(\mathbb{R}^d)$ consists of all functions $f \in L_2(\mathbb{R}^d)$ satisfying

$$\|f\|_{H^\tau(\mathbb{R}^d)}^2 := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \|\xi\|^{2\tau} d\xi < \infty.$$

We say that a filter $a : \mathbb{Z}^d \mapsto \mathbb{C}$ has τ sum rules if

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} a(\mathbf{n} + \mathbf{M}_0 \mathbf{k})(\mathbf{n} + \mathbf{M}_0 \mathbf{k})^\beta = \sum_{\mathbf{k} \in \mathbb{Z}^d} a(\mathbf{M}_0 \mathbf{k})(\mathbf{M}_0 \mathbf{k})^\beta$$

for all $\mathbf{n} \in \mathbb{Z}^d$ and for all $\beta = (\beta_1, \dots, \beta_d) \in (\mathbb{N} \cup \{0\})^d$ such that $0 \leq \beta_1, \dots, \beta_d < \tau$, $\beta_1 + \dots + \beta_d < \tau$.

It follows from Theorem 3.3 that for any $f \in L_2(\mathbb{R}^d)$, expanding f under the tight frame \mathcal{AS}_J , we have

$$f = \sum_{\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H} \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^j; \mathbf{k}}^\beta \rangle \psi_{\mathbf{N}_\beta \mathbf{M}_0^j; \mathbf{k}}^\beta + \sum_{j=J}^{\infty} \sum_{\ell=1}^r \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{\mathbf{M}_0^j; \mathbf{k}}^\ell \rangle \psi_{\mathbf{M}_0^j; \mathbf{k}}^\ell$$

with the series converging absolutely in $L_2(\mathbb{R}^d)$. We now consider the truncated series

$$P_J f := \sum_{\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H} \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{\mathbf{N}_\beta \mathbf{M}_0^j; \mathbf{k}}^\beta \rangle \psi_{\mathbf{N}_\beta \mathbf{M}_0^j; \mathbf{k}}^\beta.$$

By the same argument as in Theorem 3.3, it is not difficult to deduce that

$$P_J f = \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \phi_{\mathbf{M}_0^j; \mathbf{k}} \rangle \phi_{\mathbf{M}_0^j; \mathbf{k}}$$

The next theorem now follows from well-known results (see, e.g., [10, 22, 26] and references therein).

THEOREM 3.4. *Let \mathcal{AS}_J be the tight frame for $L_2(\mathbb{R}^d)$ introduced in Definition 3.2. Suppose that \mathbf{M}_0 is isotropic, that is, \mathbf{M}_0 is similar to a diagonal matrix with all entries having the same modulus. If the filter a has τ sum rules, then \mathcal{AS}_J has approximation order τ , that is, there exists a positive constant C such that*

$$\|f - P_J f\|_{L_2(\mathbb{R}^d)} \leq C |\det \mathbf{M}_0|^{-\tau J/d} |f|_{H^\tau(\mathbb{R}^d)} \quad \text{for all } f \in H^\tau(\mathbb{R}^d), J \in \mathbb{N}.$$

3.3. General Construction of an AMRA. Based on the work in [20, 21], we shall first propose a general construction recipe for an AMRA. Then we shall present two concrete examples in Subsections 3.4 and 4.3 to illustrate the general construction. Let us start by recalling the comments after Corollary 2.3, in which it was pointed out that the construction of a sequence of dilation matrices \mathbf{M}_ℓ , $1 \leq \ell \leq r$, i.e., invertible integer matrices and filters a_ℓ , $\ell = 1, \dots, r$ satisfying (2.7), depends only on the generated lattice $\mathbf{M}_\ell \mathbb{Z}^d$, $1 \leq \ell \leq r$. Hence we first present a very general construction of such a sequence of matrices and filters provided that the matrices satisfy $\mathbf{M}_\ell \mathbb{Z}^d = \mathbf{M} \mathbb{Z}^d$ for all $1 \leq \ell \leq r$ for some matrix \mathbf{M} , which was the hypothesis of Corollary 2.3.

For this, we first fix a sublattice Γ of \mathbb{Z}^d . Obviously, there exist many $d \times d$ matrices \mathbf{M} such that $\mathbf{M} \mathbb{Z}^d = \Gamma$. We now construct a set of tight frame filters for such a lattice Γ so that

$$\sum_{\ell=1}^r \widehat{a}_\ell(\xi) \overline{\widehat{a}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \Omega_{\mathbf{M}}. \quad (3.6)$$

As already mentioned before, the above equations in (3.6) only depend on Ω_M , which in turn only depend on the lattice $M\mathbb{Z}^d = \Gamma$. Here we employ a technique as used in [20, Cor. 3.4]. Since M is an integer matrix, it is a well-known fact from linear algebra that there exist two integer matrices E and F such that $|\det E| = |\det F| = 1$ and

$$M = EDF, \quad D = \text{diag}(d_1, \dots, d_m, 1, \dots, 1), \quad d_1 \geq \dots \geq d_m > 1.$$

For the dilation matrix D , we first construct a tensor product tight affine frame filter bank as follows. For each $d_n > 1$, one can easily construct a one-dimensional tight frame filter bank $u_\ell, \ell = 1, \dots, r_{d_n}$ such that

$$\sum_{\ell=1}^{r_{d_n}} \widehat{u}_\ell(\xi) \overline{\widehat{u}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \{0, \frac{1}{d_n}, \dots, \frac{d_n-1}{d_n}\}. \quad (3.7)$$

For the construction of (compactly supported) one-dimensional tight frame filters, there exist a long list of references; examples are [10, 21, 22, 34].

Now we construct a tensor product filter bank by constructing filters $U_{(\ell_1, \dots, \ell_m)}$ on the lattice \mathbb{Z}^d as follows:

$$U_{(\ell_1, \dots, \ell_m)}(\beta_1, \dots, \beta_m, \beta_{m+1}, \dots, \beta_d) := u_{\ell_1}(\beta_1) \cdots u_{\ell_m}(\beta_m), \quad \beta_1, \dots, \beta_d \in \mathbb{Z}. \quad (3.8)$$

This generates a total of $r := \prod_{j=1}^m r_{d_j}$ filters. We reorder them as U_1, \dots, U_r . By the definition of Ω_{DF} , we see that Ω_{DF} has the tensor product structure:

$$\Omega_{DF} = \{0, \frac{1}{d_1}, \dots, \frac{1}{d_1}\} \times \cdots \times \{0, \frac{1}{d_m}, \dots, \frac{1}{d_m}\} \times \{0\} \times \cdots \times \{0\}. \quad (3.9)$$

By the tensor-product structure in (3.8) and the tensor-product structure of Ω_{DF} in the (3.9), using (3.7), one can easily check that

$$\sum_{\omega \in \Omega_{DF}} \widehat{U}_\ell(\xi) \overline{\widehat{U}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \xi \in \mathbb{R}^d. \quad (3.10)$$

Now we define $\widehat{a}_\ell(\xi) := \widehat{U}_\ell(E^T \xi)$, $\xi \in \mathbb{R}^d$ and $\ell = 1, \dots, r$. By $M = EDF$ and the definition of Ω_M in (2.4), we have $\Omega_M = \{E^T \zeta \mid \zeta \in \Omega_{DF}\}$. Changing from the variable ξ to $E^T \xi$ in (3.10), we deduce that

$$\sum_{\omega \in \Omega_M} \widehat{a}_\ell(\xi) \overline{\widehat{a}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \xi \in \mathbb{R}^d. \quad (3.11)$$

In other words, using a tensor product construction, for any sublattice of \mathbb{Z}^d generated by $M\mathbb{Z}^d$, we can easily obtain a set of (compactly supported) tight affine(-like) frame filters satisfying (3.11). It is very important to notice that the above construction only depends on the lattice $M\mathbb{Z}^d$ instead of the dilation matrix M itself. For more details on construction of high-dimensional wavelet filter banks using the tensor product methods, see [20, 21].

Next, we shall present a general construction of tight affine(-like) frame filters to fulfill the need in this paper. Suppose that we are given a group of dilation matrices $M_\ell, \ell = 1, \dots, r$. To achieve directionality, as discussed above, only the dilation matrices for the low-pass filters will be important. In other words, if shear matrices are involved, i.e., direction-based matrices, they play an important role for low-pass

filter only and for high-pass filters, the choice of the dilation matrices does not matter, since no further decomposition will be performed for high-pass coefficients.

We group these dilation matrices into subgroups according to their lattices $M_\ell \mathbb{Z}^d$: if the lattices $M_\ell \mathbb{Z}^d = M_{\ell'} \mathbb{Z}^d$ are the same, then $M_\ell, M_{\ell'}$ are grouped into the same group. For each dilation matrix in a subgroup, we only use a fixed set of tight affine frame filters that are constructed for that lattice. In other words, for a given lattice, we have a set of tight affine frame filters and we have complete freedom in choosing the direction-based matrices, for instance, shear matrices, to achieve directionality as long as the resulting lattice is the same given lattice.

Suppose now that we choose N sets of such tight affine frame filters for all the groups of dilation matrices. Since these N sets of tight affine frames are completely independent, when we put them together to get one whole set of tight affine frame filters, we have to multiply the factor $\frac{1}{\sqrt{N}}$ to every involved filter in the set. Now it is straightforward to see that the total collection of all such N sets of renormalized tight affine frame filters indeed forms a collection of tight affine frame filters satisfying (2.7), where r is the total number of all the involved filters. It is also very important to notice that the renormalization of the filters does not reduce the directionality of the tight affine frame system, since the coefficients in the same band has the same ordering of magnitude as the one without renormalization.

3.4. An Example of Applications of AMRA. Based on the general construction of an AMRA discussed in the previous subsection, we now present one explicit example of an AMRA and discuss applications to image denoising, see also [25]. Even though we will construct a wavelet frame system consisting of piecewise linear spline exhibiting a relatively simple AMRA structure, some improvements in image denoising can already be observed as compared to regular MRA framelet based denoising. Since however applications are not the focus of this paper, the example given here shall only illustrate the potential of our framework and show how to utilize this general construction in particular applications.

Let now the filters $u_0, u_1,$ and u_2 be defined by

$$u_0(\xi) := \frac{1}{4}e^{i\xi}(1 + e^{-i\xi})^2, \quad u_1(\xi) := \frac{\sqrt{2}}{4}(e^{-i\xi} - e^{i\xi}), \quad u_2(\xi) := \frac{1}{4}e^{i\xi}(e^{-i\xi} - 1)^2. \quad (3.12)$$

Notice that u_0 is the refinement mask of the centralized piecewise B-spline $\phi(x) = \max(1 - |x|, 0)$. Then the following UEP condition is satisfied (see [34]):

$$\sum_{\ell=0}^2 \widehat{u}_\ell(\xi) \overline{\widehat{u}_\ell(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \{0, \frac{1}{2}\}.$$

At each level j , we now choose the dilation matrices $M_{j,\ell}, \ell = 1, \dots, r_j$ by

$$M_{j,\ell} = E_{j,\ell} M_0,$$

where $M_0 = 2I_2$ and $E_{j,\ell}, \ell = 1, \dots, r_j$ with $r_j \leq 5$ are chosen from the set of shear matrices:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \right\},$$

which we label for later purposes by E_1, \dots, E_5 . We wish to mention that $M_{j,\ell} \mathbb{Z}^2 = 2\mathbb{Z}^2$ for all $\ell = 1, \dots, r_j$ is ensured by this choice. At each level j , we now use the

TABLE 3.1

PSNR values of Numerical results using framelet based ℓ_1 norm regularizations for image denoising with $\sigma = 20$.

denoising	“bowl256”	“barbara512”	“baboon512”	“zebra512”
MRA framelet	28.77	27.39	25.81	28.19
AMRA framelet	29.88	28.05	25.91	28.57
Sequence of E_k	(E_1, E_5, E_3)	(E_2, E_4, E_3)	(E_1, E_4, E_2)	(E_1, E_3, E_4)

following tensor product filters:

$$a_{\ell_1, \ell_2}(\beta_1, \beta_2) := u_{\ell_1}(\beta_1)u_{\ell_2}(\beta_2), \quad \beta_1, \beta_2 \in \mathbb{Z}^2, \quad \ell_1, \ell_2 = 0, 1, 2,$$

where $a_{0,0}$ is the low-pass filter and the other filters a_{ℓ_1, ℓ_2} are high-pass filters. The choice of $E_{j, \ell}$ for each level j is adaptive to the content of the given images. We expect such an adaptive nonstationary framelet system with associated AMRA to provide much sparser approximations of images, while associated transforms show the same computational efficiency and implementation simplicity as for 2D tensor product framelet systems.

The benefit of using such an AMRA framelet system is demonstrated for image denoising, see also [25]. In the experiments, the noise component is synthesized by Gaussian noise with noise level $\sigma = 20$. Then the image is recovered from the noisy observation X by solving the following balanced minimization mode of [37, 13]:

$$f := W^T v; \quad v = \min_{\tilde{v}} \frac{1}{2} \|X - W^T \tilde{v}\|_2^2 + \frac{\kappa}{2} \|(I - WW^T)\tilde{v}\|_2^2 + \lambda \|\tilde{v}\|_1,$$

where W denotes the linear decomposition – the inverse being W^T due to the tight frame property – and $X = f + \epsilon$ denotes the noisy observation of the original image f corrupted by Gaussian noise ϵ with variance σ . The APG method as given in [37] is then used to solve the above minimization with $\kappa = 5$ and $\lambda = 0.11$. Throughout the experiments, a three-level decomposition ($J = 3$) of AMRA framelets is applied and at each level the choice of the matrix $E_{j, \ell}$ is automatically determined by some simple analysis on the output of the framelet coefficients. The PNSR values of the results from both regular framelets and AMRA framelets are listed in Table 3.1 and sample images are shown in Figure 3.1. It is evident from the numerical results that AMRA framelets are superior to regular MRA framelet in terms of both PSNR values and visual quality.

4. Shearlet Systems with an Associated AMRA. In this section, we will apply our general framework to the special case of shearlets. This leads to shearlet systems, especially compactly supported ones, associated with an AMRA structure and fast decomposition and reconstruction algorithms. We anticipate that our considerations will improve the applicability of shearlets to various applications.

4.1. Traditional Shearlet Systems. The continuum shearlet transform with discrete parameters [17, 29] for functions in $L_2(\mathbb{R}^d)$ uses a two-parameter dilation group, where one parameter indexes scale, and the second parameter indexes orientation. For each $c > 0$ and $t \in \mathbb{R}$, let A_c denote the *parabolic scaling matrix* and S^t denote the *shear matrix* of the form

$$A_c := \begin{pmatrix} c & 0 \\ 0 & \sqrt{c} \end{pmatrix}, \quad c > 0 \quad \text{and} \quad S^t := \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad (4.1)$$



FIG. 3.1. (a) Original image “Barbara”; (b) Noisy image with $\sigma = 20$; (c) Results from MRA framelet; (d) Results from AMRA framelet

respectively. To provide an equal treatment of the x - and y -axis, we split the frequency plane into the horizontal cone

$$\mathcal{C}_0 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \geq 1, |\xi_1/\xi_2| \geq 1\},$$

the vertical cone

$$\mathcal{C}_1 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \geq 1, |\xi_1/\xi_2| \leq 1\},$$

as well as a centered rectangle

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

(see Figure 4.1 (a)).

For cone \mathcal{C}_0 , at scale $j \geq 0$, orientation $k = -2^j, \dots, 2^j$, and spatial position $m \in \mathbb{Z}^2$, the associated *shearlets* are then defined by

$$\sigma_\eta = 2^{j\frac{3}{4}} \psi(S^k A_{4^j} \cdot -m),$$

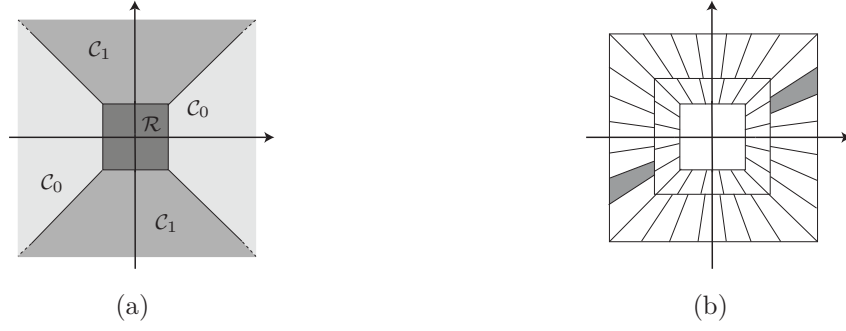


FIG. 4.1. (a) The cones \mathcal{C}_0 and \mathcal{C}_1 and the centered rectangle \mathcal{R} in frequency domain. (b) The tiling of the frequency domain induced by discrete shearlets.

where $\eta = (j, k, m, \iota)$ index scale, orientation, position, and cone. The shearlets for \mathcal{C}_1 are defined likewise by symmetry, as illustrated in Figure 4.1 (b), and we denote the resulting *shearlet system* by

$$\{\sigma_\eta : \eta \in \mathbb{N}_0 \times \{-2^j, \dots, 2^j\} \times \mathbb{Z}^2 \times \{0, 1\}\}. \quad (4.2)$$

This is an affine-like system as defined before.

Notice that we chose a scaling of 4^j . Shearlet systems can be defined similarly for a scaling of 2^j , however, in this case the odd scales have to be handled particularly carefully.

One particular interesting shearlet generator is the function $\psi \in L_2(\mathbb{R}^d)$ defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where $\psi_1 \in L^2(\mathbb{R})$ is a wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and $\text{supp } \hat{\psi}_1 \subseteq [-4, -\frac{1}{4}] \cup [\frac{1}{4}, 4]$, and $\psi_2 \in L^2(\mathbb{R})$ is a ‘bump’ function satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and $\text{supp } \hat{\psi}_2 \subseteq [-1, 1]$. Filling in the low frequency band appropriately, with this particular generator the shearlet system (4.2) can be proven (see [17, Thm. 3]) to form a tight frame for $\{f \in L_2(\mathbb{R}^d) : \text{supp } \hat{f} \subset \mathcal{C}_0 \cup \mathcal{C}_1\}$.

Concluding, the definition just discussed shows that shearlets live on anisotropic regions of width 2^{-2j} and length 2^{-j} at various orientations, which are parameterized by slope rather than angle as for second generation curvelets.

4.2. Shearlet Unitary Extension Principle. Let us now apply Theorem 2.2 to the special situation of shearlet systems. This leads to the following result, which we could coin the ‘Shearlet Unitary Extension Principle’.

THEOREM 4.1. *Let S^{t_1}, \dots, S^{t_r} be a selection of 2×2 shear matrices in (4.1). Let $a_\ell, \ell = 1, \dots, r$, be finitely supported filters on \mathbb{Z}^2 . Then the following perfect reconstruction property in the discrete domain holds:*

$$\sum_{\ell=1}^r \mathcal{S}_{a_\ell, S^{t_\ell} A_4} \mathcal{T}_{a_\ell, S^{t_\ell} A_4} v = v, \quad v \in l_2(\mathbb{Z}^2),$$

if and only if, for any $\omega \in \Omega = \bigcup_{\ell=1}^r \Omega_{S^{t_\ell} A_4}$, where $\Omega_{S^{t_\ell} A_4} := [(S^{-t_\ell})^\top (\frac{1}{4}\mathbb{Z} \times \frac{1}{2}\mathbb{Z})] \cap [0, 1)^2$,

$$\sum_{\ell \in \{1 \leq n \leq r : \omega \in \Omega_{S^{t_n} A_4}\}} \hat{a}_\ell(\xi) \overline{\hat{a}_\ell(\xi + 2\pi\omega)} = \delta(\omega),$$

where δ denotes the Dirac sequence such that $\delta(0) = 1$ and $\delta(\omega) = 0$ for $\omega \neq 0$.

4.3. Shearlet Constructions and Algorithms. We next focus on explicit shearlet constructions. For this, we recall that the dilation for shearlets is composed of a shear matrix S^{t_ℓ} with $t_\ell \in \mathbb{Z}$ and a parabolic scaling matrix A_c with $c = 4$ in (4.1). Thus, at each level of the decomposition, the matrices M_ℓ will be chosen as

$$M_\ell = S^{t_\ell} A_4,$$

for some shear matrices S^{t_ℓ} such that the selection of the shear matrix parameters $t_\ell \in \mathbb{Z}$ should be driven by the particular application at hand, and A_4 is the parabolic scaling matrix, where the choice of the value 4 (in contrast to 2) avoids technicalities caused by square roots. As already elaborated upon before, which of those matrices will be labeled low- and which one high-pass is left to the user.

Aiming towards examples for possible filters for the Shearlet Unitary Extension Principle, we first observe that there exists only two different lattices in Theorem 4.1: Consider the product $S^{t_\ell} A_4$. If t_ℓ is an even integer, then

$$S^{t_\ell} A_4 \mathbb{Z}^2 = A_4 \mathbb{Z}^2;$$

if t_ℓ is an odd integer, then

$$S^{t_\ell} A_4 \mathbb{Z}^2 = S A_4 \mathbb{Z}^2, \quad \text{where } S := S^1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Hence, we only need to design two sets of tight shearlet frame filters in advance for the lattices $A_4 \mathbb{Z}^2$ and $S A_4 \mathbb{Z}^2$, respectively, following the general construction in Subsection 3.3. Then we have complete freedom in choosing the shear matrix parameters t_ℓ to obtain a whole set of tight shearlet frame filters satisfying perfect reconstruction (see condition (i) in Theorem 4.1).

Let us now provide a concrete construction, which is also a special case of the general construction of an AMRA described in Subsection 3.3. For this, let u_0, u_1, u_2 be the filters defined in (3.12) and set

$$\begin{aligned} \hat{u}_0(\xi) &:= \frac{1}{16} e^{i3\xi} (1 + e^{-i\xi})^2 (1 + e^{-i2\xi})^2, \\ \hat{u}_1(\xi) &:= \frac{\sqrt{2}}{16} e^{i3\xi} (e^{-i\xi} - 1) (1 + e^{-i\xi})^3 (1 + e^{-i2\xi}), \\ \hat{u}_2(\xi) &:= \frac{1}{16} e^{i3\xi} (e^{-i\xi} - 1)^2 (1 + e^{-i\xi})^4, \\ \hat{u}_3(\xi) &:= \frac{1}{4} e^{i\xi} (e^{-i\xi} - 1) (1 + e^{-i\xi}), \\ \hat{u}_4(\xi) &:= \frac{\sqrt{2}}{8} e^{i\xi} (e^{-i\xi} - 1)^2, \\ \hat{u}_5(\xi) &:= \frac{1}{4} e^{-i\xi} (e^{-i\xi} - 1) (1 + e^{-i\xi}), \\ \hat{u}_6(\xi) &:= \frac{\sqrt{2}}{8} e^{-i\xi} (e^{-i\xi} - 1)^2. \end{aligned}$$

One can directly check that $\tilde{u}_0, \dots, \tilde{u}_6$ satisfy the Unitary Extension Principle with dilation 4:

$$\sum_{\ell=0}^6 \overline{\hat{u}_\ell(\xi)} \hat{u}_\ell(\xi + 2\pi\omega) = \delta(\omega), \quad \omega \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}.$$

For the lattice $A_4\mathbb{Z}^2$, we use the following tensor product tight framelet filter bank:

$$a_{\ell_1, \ell_2}(\beta_1, \beta_2) := \tilde{u}_{\ell_1}(\beta_1)u_{\ell_2}(\beta_2), \quad \beta_1, \beta_2 \in \mathbb{Z}^2, \ell_1 = 0, \dots, 6, \ell_2 = 0, 1, 2,$$

where $a_{0,0}$ is the low-pass filter and the other filters a_{ℓ_1, ℓ_2} are high-pass filters.

For the lattice $SA_4\mathbb{Z}^2$, we simply use the sheared tight framelet filter bank $\tilde{a}_{\ell_1, \ell_2}, \ell_1 = 0, \dots, 6$ and $\ell_2 = 0, 1, 2$, where

$$\widehat{\tilde{a}_{\ell_1, \ell_2}}(\xi) := \widehat{a_{\ell_1, \ell_2}}(S^T \xi), \quad \xi \in \mathbb{R}^2.$$

Depending on the application at hand, we can adaptively choose the shear matrices to achieve flexibility and directionality for various types of data. There are certainly various other choices of tight framelet filter banks which can be employed in our construction of an AMRA.

Each such a choice of filters for each level then leads to a Fast Adaptive Shearlet Decomposition associated with a Fast Adaptive Shearlet Reconstruction. Those two algorithms are described in Figures 4.2 and 4.3. Notice that here – as already in the general version of (FAD) and (FAR) – the shear matrices can be chosen differently at each level of the decomposition. We can envision that this adaption can be made flexible dependent on a quick analysis, say, thresholding of the data outputted in the previous decomposition step. The algorithm leaves all those possibilities open. The great flexibility provided here should be utilizable for each particular application.

REFERENCES

- [1] J.-F. Cai, R. H. Chan and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal. **24** (2008), 131–149.
- [2] J.-F. Cai, H. Ji, C. Liu and Z. Shen, *Blind motion deblurring from a single image using sparse approximation*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.
- [3] J.-F. Cai, S. Osher and Z. Shen, *Linearized Bregman iteration for frame based image deblurring*, SIAM J. Imaging Sci. **2** (2009), 226–252.
- [4] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms*, Multiscale Model. Simul. **5** (2006), 861–899.
- [5] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.
- [6] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197.
- [7] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: II. Discretization of frames*, Appl. Comput. Harmon. Anal. **19** (2005), 198–222.
- [8] R. H. Chan, S. D. Riemenschneider, L. Shen, and Z. Shen, *Tight frame: an efficient way for high-resolution image reconstruction*, Appl. Comput. Harmon. Anal. **17** (2004), 91–115.
- [9] S. Dahlke, G. Steidl, and G. Teschke, *The Continuous Shearlet Transform in Arbitrary Space Dimensions*, J. Fourier Anal. Appl., to appear.
- [10] I. Daubechies, B. Han, A. Ron, and Z. Shen, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harmon. Anal. **14** (2003), 1–46.
- [11] I. Daubechies, G. Teschke, and L. Vese, *Iteratively solving linear inverse problems under general convex constraints*, Inverse Problems and Imaging **1** (2007), no. 1, 29.
- [12] M. N. Do and M. Vetterli, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans. Image Process. **14** (2005), 2091–2106.
- [13] B. Dong, Z. Shen, *MRA-based wavelet frames and applications*, IAS Lecture Notes Series, Summer Program on “The Mathematics of Image Processing”, Park City Mathematics Institute, 2010.
- [14] D. L. Donoho, *Wedgelets: nearly minimax estimation of edges*, Ann. Stat. **27** (1999), 859–897.
- [15] G. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal. **25** (2008), 25–46.

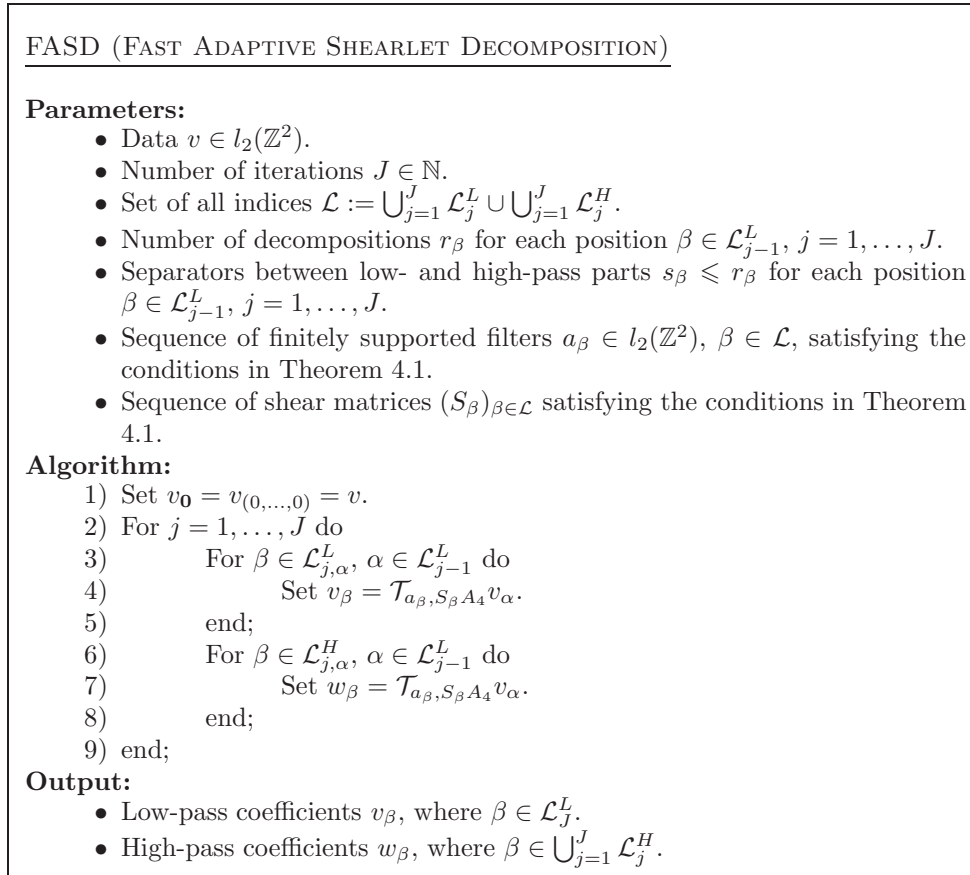


FIG. 4.2. The FASD Algorithm for a fast adaptive decomposition using shearlet systems.

- [16] M. Elad, J.L. Starck, P. Querre, and D.L. Donoho, *Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)*, Applied and Computational Harmonic Analysis **19** (2005), no. 3, 340–358.
- [17] K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators*, Wavelets and Splines (Athens, GA, 2005), Nashboro Press, Nashville, TN (2006), 189–201.
- [18] K. Guo and D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, SIAM J. Math. Anal. **39** (2007), 298–318.
- [19] K. Guo, D. Labate, W. Lim, G. Weiss, and E. Wilson, *Wavelets with Composite Dilations and their MRA Properties*, Appl. Comput. Harmon. Anal. **20** (2006), 231–249.
- [20] B. Han, *Symmetry property and construction of wavelets with a general dilation matrix*, Linear Algebra Appl. **353** (2002), 207–225.
- [21] B. Han, *Compactly supported tight wavelet frames and orthonormal wavelets of exponential decay with a general dilation matrix*, J. Comput. Appl. Math., **155** (2003), 43–67.
- [22] B. Han, *Dual multiwavelet frames with high balancing order and compact fast frame transform*, Appl. Comput. Harmon. Anal. **26** (2009), 14–42.
- [23] B. Han, *The structure of balanced multivariate biorthogonal multiwavelets and dual multiwavelets*, Math. Comp., **79** (2010), 917–951.
- [24] B. Han, *Pairs of frequency-based nonhomogeneous dual wavelet frames in the distribution space*, Appl. Comput. Harmon. Anal. **29** (2010), 330–353.
- [25] H. Ji, C. Jiu, and Z. Shen *Adaptive MRA framelet based image restoration*, in preparation.
- [26] R. Q. Jia, *Approximation properties of multivariate wavelets*, Math. Comp. **67** (1998), 647–665.
- [27] G. Kutyniok, *Sparsity Equivalence of Anisotropic Decompositions*, preprint.
- [28] G. Kutyniok, M. Shahram, and D. L. Donoho, *Development of a Digital Shearlet Transform*

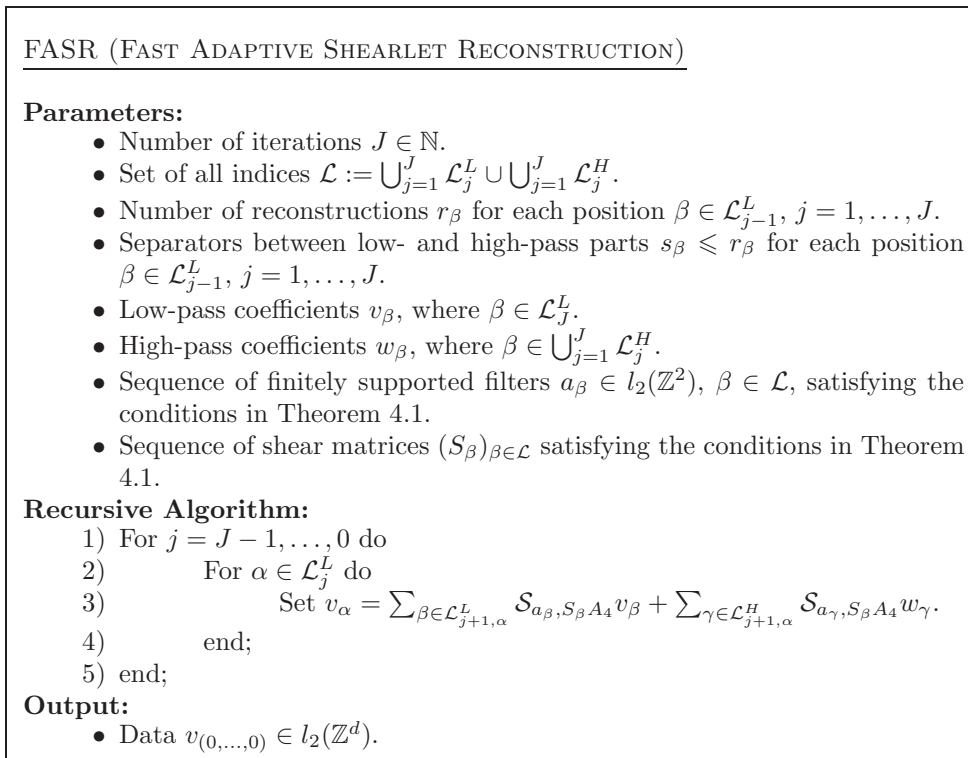


FIG. 4.3. The FASR Algorithm for a fast adaptive reconstruction using shearlet systems.

- Based on Pseudo-Polar FFT, Wavelets XIII (San Diego, CA, 2009), 74460B-1 - 74460B-13, SPIE Proc. 7446, SPIE, Bellingham, WA, 2009.*
- [29] G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.
- [30] G. Kutyniok and W.-Q. Lim, *Compactly supported shearlets are optimally sparse*, J. Approx. Theory, to appear.
- [31] G. Kutyniok and T. Sauer, *Adaptive directional subdivision schemes and Shearlet Multiresolution Analysis*, SIAM J. Math. Anal. **41** (2009), 1436–1471.
- [32] E. Le Pennec and S. Mallat, *Bandelet image approximation and compression*, Multiscale Model. Simul. **4** (2005), 992–1039.
- [33] W. Lim, *Discrete Shearlet Transform: New Multiscale Directional Image Representation* Proceedings of SAMP'09, Marseille 2009.
- [34] A. Ron and Z. Shen, *Affine systems in $L^2(\mathbb{R}^d)$: the analysis of the analysis operator*, J. of Funct. Anal. **148** (1997), 408–447.
- [35] A. Ron and Z. Shen, *Affine systems in $L^2(\mathbb{R}^d)$: dual systems*, J. Fourier Anal. Appl. **3** (1997), 617–637.
- [36] Z. Shen, *Wavelet frames and image restorations*, in: R. Bhatia (Ed.), Proceedings of the International Congress of Mathematicians 2010, Vol IV, Hyderabad, India, Hindustan Book Agency, pp. 2834–2863.
- [37] Z. Shen, K. C. Toh, and S. Yun, *An accelerated proximal gradient algorithm for frame based image restorations via the balanced approach*, preprint (2009).