

## Explicit schemes in seismic migration and isotropic multiscale representations

S. Jain, M. Papadakis, E. Dussaud

**ABSTRACT.** We construct examples of non-separable Isotropic Multiresolution Analyses (IMRA) for  $L^2(\mathbb{R}^d)$ . We develop a wave equation based poststack depth migration scheme using the frames arising from IMRA. If we discretise the signal at only one resolution level, then the method reduces to a so-called explicit scheme (see for example [8, 10]). The multiscale structure of IMRA, offers the possibility of reducing the cost of computation.

### 1. Introduction

Migration is a seismic imaging technique used by the oil industry to image the interior of the Earth for the purpose of oil prospecting. A detailed description of seismic migration is beyond the scope of this article. In Section 2, we give a very brief overview of what is called wave equation migration and we arrive at the so-called Phase-shift propagator operator which is used to obtain the image of the interior by downward propagating acoustic waves recorded on the surface. Our focus in this article is the efficient discretisation of this operator in terms of pairs of dual frames arising from a First Generation Isotropic Multiresolution Analysis (IMRA) described in Section 3. Seismic migration is a delicate inverse problem. To solve it we work on a 3 or 4 dimensional time-frequency space,  $\mathbb{R} \times \mathbb{R}^2$  or  $\mathbb{R} \times \mathbb{R}^3$  respectively. The underlying spaces are  $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}^d)$  with  $d = 2, 3$ . One of the main tools is the Fourier transform. We actually use two such transforms; the Fourier transform on  $L^2(\mathbb{R})$  and the Fourier transform on  $L^2(\mathbb{R}^d)$ . We refer to  $\mathbb{R}$  as the time-domain and to  $\mathbb{R}^d$  as the spatial domain. We refer to the dual group of  $\mathbb{R}^2$  or  $\mathbb{R}^3$  as the wavenumber domain to distinguish it from the dual group of  $\mathbb{R}$ , which we refer to as the frequency domain. Wavenumbers, i.e. elements of the dual

---

*Key words and phrases.* isotropic multiresolution analysis, wavelets, frames, wave-equation migration.

This research was supported in part by the following grants: Total E&P, USA Inc., NSF-DMS 0406748 and by a Research Innovation Grant by the Texas Learning and Computation Center of the University of Houston.

M. Papadakis and S. Jain are with the University of Houston, Department of Mathematics. E. Dussaud is with Total E&P, USA Inc.

Mathematics subject classification (2000): 42C40, 65T60, 35Q80, 34A55.

group of  $\mathbb{R}^d$ , are denoted by  $\boldsymbol{\xi}$ 's versus elements of the dual group of  $\mathbb{R}$ , the time domain, denoted by  $\omega$ 's.

The phase shift operator is a multiplicative operator in the wavenumber domain. It is therefore, a convolution in the space domain. If the velocity of sound waves is not assumed to be constant beneath the surface then this becomes a spatially varying convolution operator. The recursive application of this convolution operator is referred to as an explicit scheme. Explicit schemes for wave equation migration are attractive because they handle lateral variations in velocity better than implicit schemes such as Fourier finite differences [13]. However, the computational cost of explicit schemes is always a matter of concern due to the enormity of the data sets. In Section 4 we show that the IMRA-based wave equation migration (henceforth referred to as migration) reduces to a standard explicit scheme when applied to a signal at the zero resolution level. The multiscale structure of IMRA, offers the possibility of reducing the computational cost. This can be compared to the work of Margrave and his collaborators [11], who describe a sub-sampling scheme using the frequency domain.

We continue with some additional notation that we use throughout the article. The shift operator  $T_{\mathbf{k}} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$  is defined by

$$T_{\mathbf{k}}f(\mathbf{x}) = f(\mathbf{x} - \mathbf{k}) \quad \forall \mathbf{x} \in \mathbb{R}^d; \mathbf{k} \in \mathbb{Z}^d.$$

We call a  $d \times d$  matrix,  $A$ , with integer entries, **radially expansive** if  $A = a\mathcal{O}$ , for some fixed  $a > 1$  and a  $d \times d$  unitary matrix  $\mathcal{O}$ . We define the dilation  $D_A : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ , with respect to a radially expansive matrix  $A$  by,

$$D_Af(\mathbf{x}) = |\det(A)|^{1/2}f(A\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

For notational convenience we drop the subscript,  $A$ , and use  $D$  to denote the dilation operator with the understanding that it is always defined with respect to a radially expansive matrix  $A$ . It is easy to check that the translation and dilation operators defined above are unitary. Let  $\mathcal{F}$  denote the Fourier Transform on  $L^2(\mathbb{R}^d)$  defined by

$$\mathcal{F}f(\boldsymbol{\xi}) = \widehat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} f(\mathbf{x})e^{-2\pi i\langle \mathbf{x}, \boldsymbol{\xi} \rangle} d\mathbf{x} \quad \boldsymbol{\xi} \in \mathbb{R}^d$$

for all  $f \in (L^1 \cap L^2)(\mathbb{R}^d)$  and extended to  $L^2(\mathbb{R}^d)$  by Plancherel's Theorem. For convenience we adopt the notation,

$$e_{\mathbf{k}}(\boldsymbol{\xi}) = e^{-2\pi i\langle \mathbf{k}, \boldsymbol{\xi} \rangle} \quad \boldsymbol{\xi} \in \mathbb{R}^d; \mathbf{k} \in \mathbb{Z}^d \quad \text{and } \mathbb{T}^d = [-1/2, 1/2]^d.$$

Next we define the concepts of frames and Bessel families on separable Hilbert space  $H$ .

**DEFINITION 1.1.** *Let  $(H, \langle \cdot, \cdot \rangle_H)$  be a separable Hilbert space. A subset  $\{f_n\}_{n \in \Lambda}$  of  $H$ , where  $\Lambda$  is a countable indexing set, is called a **frame** if and only if there exist two constants  $0 < A, B < +\infty$ , such that*

$$(1) \quad A\|g\|_H^2 \leq \sum_{n \in \Lambda} |\langle g, f_n \rangle_H|^2 \leq B\|g\|_H^2, \quad \text{for all } g \in H.$$

The set,  $\{f_n\}_{n \in \Lambda}$ , is called a **Bessel family** if only the upper bound,  $B$ , in (1) exists and  $A = 0$ . The minimum constant  $B$  for which (1) holds is called the **upper Bessel bound** or the **upper frame bound**. The optimal constant  $A$  for which (1) holds is called the **lower frame bound**.

If  $A = B = 1$  then we refer to the frame as **Parseval**. In this case we have the following reconstruction formula:

$$g = \sum_{n \in \Lambda} \langle g, f_n \rangle_H f_n \quad \text{for all } g \in H.$$

Let  $K$  be a separable Hilbert space with an orthonormal basis  $\{e_n\}_{n \in \Lambda}$ . The operator,  $S : K \rightarrow H$ , defined via  $S(e_n) = f_n$ , is called the synthesis operator for the family  $\{f_n\}_{n \in \Lambda}$ . If  $\{f_n\}_{n \in \Lambda}$  is Bessel then  $S$  can be extended to a bounded linear operator. The adjoint of this operator is given by,  $S^*(f) = \sum_{n \in \Lambda} \langle f, f_n \rangle e_n$  for all  $f \in H$ . We refer to  $S^*$  as the analysis operator for  $\{f_n\}_{n \in \Lambda}$ . A restatement of the above reconstruction formula is:  $SS^* = I_H$  if and only if  $\{f_n\}_{n \in \Lambda}$  is a Parseval frame.

Let  $\{f_n\}_{n \in \Lambda}$  and  $\{\tilde{f}_n\}_{n \in \Lambda}$  be frames of  $H$ . If

$$g = \sum_{n \in \Lambda} \langle g, \tilde{f}_n \rangle_H f_n,$$

holds for all  $g \in H$ , then  $\{f_n\}_{n \in \Lambda}$  and  $\{\tilde{f}_n\}_{n \in \Lambda}$  are dual to each other. In operator theoretic language, if  $S$  and  $\tilde{S}$  are the synthesis operators for  $\{f_n\}_{n \in \Lambda}$  and  $\{\tilde{f}_n\}_{n \in \Lambda}$  respectively, then  $\{f_n\}_{n \in \Lambda}$  and  $\{\tilde{f}_n\}_{n \in \Lambda}$  are dual to each other if and only if  $S(\tilde{S})^* = I_H$ . For more details on frames, the reader is referred to [3].

## 2. Explicit schemes for wave equation migration

In this section we give a brief overview of the explicit wave equation migration. In a typical prospecting experiment, sound waves generated by small explosions travel from a source on the surface into the ground and are reflected back by the complex structures beneath the surface which we refer to as reflectors. These reflectors represent the surfaces where the stratigraphy changes. The reflected waves are recorded on receivers called geophones (or hydrophones in the case of off-shore exploration). If we assume that the source and the receiver are located at the same point then the data (recorded pressure wave on the receivers) is said to have zero offset. For zero-offset data the incident and reflected waves travel along the same path. A simplifying assumption is to disregard the incident wave and assume that the reflectors beneath the surface explode at time  $t = 0$ , to produce the waves that travel to the surface at half of the actual velocity and are recorded on the receivers. With this assumption the propagation distance (reflector to surface) is half the actual propagation distance (surface to reflector and back to the surface). Since, both the velocity and propagation distance are half of their actual values, the pressure wave created by the exploding reflectors arrives at the same time as the original reflected wave.

Migration refers to the process of obtaining the image of the reflectors using the pressure wave recorded on the surface. The pressure wave at  $x, y$  horizontal coordinates, depth  $z$  and time  $t$  is denoted by  $f(x, y, z, t)$ . We begin by taking the Fourier transform of  $f$  with respect to the time variable. This new function with variables  $x, y, z$  and  $\omega$  is called the ‘wavefield’ for frequency,  $\omega$ . If the velocity varies only with depth  $z$ , we solve the following differential equation proposed by Claerbout

([5]), who derives it by applying principles of classical optics;

$$(2) \quad \frac{\partial \hat{f}}{\partial z}(k_x, k_y, z, \omega) = \left( \sqrt{\frac{\omega^2}{c(z)^2} - k_x^2 - k_y^2} \right) \hat{f}(k_x, k_y, z, \omega),$$

where  $c(z)$  denotes the velocity,  $\hat{f}$  denotes the wavefield in the wavenumber domain (with respect to  $x$  and  $y$  variables), and  $k_x$  and  $k_y$  denote the wavenumbers. We refer to (2) as the one-way wave equation because it represents the upward going waves when the reflectors ‘explode’ in the subsurface. The solution to (2) is given by,

$$(3) \quad \hat{f}(z; k_x, k_y, \omega) = \exp \left( \int_{z_0}^z \sqrt{\frac{\omega^2}{c(z)^2} - k_x^2 - k_y^2} dz \right) \hat{f}(z_0; k_x, k_y, \omega)$$

For a small depth step  $\Delta z$ ,  $c$  is practically constant. Hence, for practical purposes we allow

$$(4) \quad \hat{f}(z + \Delta z; k_x, k_y, \omega) = e^{i\Delta z \sqrt{\frac{\omega^2}{c(z)^2} - k_x^2 - k_y^2}} \hat{f}(z, k_x, k_y; \omega),$$

In conclusion, the wavefield at  $z + \Delta z$  is calculated from the wavefield at  $z$  via a multiplication in the wavenumber domain. The RHS of the previous equation defines a multiplicative operator referred to as the phase shift operator. This is a convolution operator in the spatial domain which we denote by  $\mathcal{P}$ . We recursively obtain the wavefield for all  $z$  using  $\mathcal{P}$ . Integrating numerically over all frequencies  $\omega$ , we obtain the pressure wave  $f(x, y, z, 0)$  for all  $z$  at time  $t = 0$  which depicts the ‘exploding’ reflectors. This procedure is termed *post stack migration*, because, in order to apply the ‘exploding reflectors’ concept, the data has to be stacked into zero offset sections. The interested reader can find the details of the stacking process and the derivation of the one-way wave equation in [5]. If the velocity varies laterally, the operator  $\mathcal{P}$  is replaced by a ‘spatially varying convolution’ operator. We clarify this kind of convolution below. But before doing this we want to mention that in the geophysical literature, the migration process described above, is discretised by sampling the wave-field on a regular grid and by approximating  $e^{\pm i\Delta z \sqrt{\frac{\omega^2}{c^2} - k_x^2 - k_y^2}}$  (the symbol of the operator,  $\mathcal{P}$  in the wavenumber domain) with a trigonometric polynomial. This approximation of  $\mathcal{P}$  is carried out for every point of the grid on which the data are collected. The collection of all these approximations of  $\mathcal{P}$  is referred to as a propagator matrix. Propagator matrices are applied recursively on the data to calculate the propagating wavefield for all  $z$ .

Our approach is formal compared to various ad-hoc discretisations of the phase-shift operator. It also leads to a sparser propagator matrix in certain cases.

Let us now return to  $\mathcal{P}$ . For a fixed  $\omega$ , we denote the wavefield at  $(\mathbf{x}, z, \omega)$  by  $f_z(\mathbf{x})$ , where  $\mathbf{x} = (x, y)$ . The operator  $\mathcal{P}$  acts via a multiplication in the wavenumber domain,

$$\hat{f}_{z+\Delta z}(\boldsymbol{\xi}) = \widehat{(\mathcal{P}f_z)}(\boldsymbol{\xi}) = e^{2\pi i\Delta z \sqrt{\frac{\omega^2}{c^2} - \|\boldsymbol{\xi}\|^2}} \hat{f}_z(\boldsymbol{\xi}),$$

where  $\boldsymbol{\xi} = (k_x, k_y)$ . Thus,

$$\mathcal{P}f_z(\mathbf{x}) = \int_{\mathbb{R}^2} e^{2\pi i\langle \mathbf{x}, \boldsymbol{\xi} \rangle} e^{2\pi i\Delta z \sqrt{\frac{\omega^2}{c^2} - \|\boldsymbol{\xi}\|^2}} \hat{f}_z(\boldsymbol{\xi}) d\boldsymbol{\xi}.$$

For all this to work, we need to assume that  $f_z$  is appropriately band-limited and is in  $L^2(\mathbb{R}^2)$ . Now, let  $\phi$  be a continuous function with compact support. Then,

$$(5) \quad \langle \mathcal{P}f_z, \phi \rangle = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{2\pi i \langle \mathbf{x}, \boldsymbol{\xi} \rangle} e^{2\pi i \Delta z \sqrt{\frac{\omega^2}{c^2} - \|\boldsymbol{\xi}\|^2}} \widehat{f}_z(\boldsymbol{\xi}) \overline{\phi(\mathbf{x})} d\boldsymbol{\xi} d\mathbf{x}$$

The previous equation allows us to consider situations that are realistic in seismic imaging, where the velocity has lateral variation. If, the support of  $\phi$  is sufficiently small, then we can assume that  $c$  is practically constant in the support of  $\phi$ . Then, we can define  $\mathcal{P}$  via (5), on a space of band-limited functions, and allow  $c$  in (5) to be a function of  $\mathbf{x} \in \mathbb{R}^2$ . The band-limited assumption, even when  $c$  varies, implies that  $\mathcal{P}$  defined via (5) is a bounded linear operator. This gives the propagator operator we want to discretise and formalizes what we called a spatially varying convolution operator. We will describe this discretisation in Section 4.

### 3. First Generation Isotropic Multiresolution Analysis

We begin this section by defining an Isotropic Multiresolution Analysis (IMRA) of  $L^2(\mathbb{R}^d)$  (see [1]). We also construct examples of IMRAs we use for the discretisation of the propagator operator in Section 4.

**DEFINITION 3.1.** *An Isotropic Multiresolution Analysis (IMRA) of  $L^2(\mathbb{R}^d)$  with respect to a radially expansive matrix  $A$  that preserves  $\mathbb{Z}^d$  is a sequence  $\{V_j\}_{j \in \mathbb{Z}}$  of closed subspaces of  $L^2(\mathbb{R}^d)$  satisfying the following conditions:*

- $\forall j \in \mathbb{Z}, V_j \subset V_{j+1}$ ,
- $(D_A)^j V_0 = V_j$ ,
- $\cup_{j \in \mathbb{Z}} V_j$  is dense in  $L^2(\mathbb{R}^d)$ ,
- $\cap_{j \in \mathbb{Z}} V_j = \{0\}$ ,
- $V_0$  is invariant under all translations  $T_{\mathbf{k}}$  such that  $\mathbf{k} \in \mathbb{Z}^d$ ,
- If  $P_0$  denotes the orthogonal projection onto  $V_0$ , then

$$(6) \quad \forall \mathcal{O} \in SO(d), \mathcal{O}P_0\mathcal{O}^t = P_0.$$

Each  $V_j$  is referred to as a **resolution space**. We will refer to  $\phi$  as a **scaling function** associated to the IMRA,  $\{V_j\}_{j \in \mathbb{Z}}$ , if  $\{T_{\mathbf{k}}\phi\}_{\mathbf{k} \in \mathbb{Z}^d}$  is a Bessel family and  $V_0 = \overline{\text{span}}\{T_{\mathbf{k}}\phi\}_{\mathbf{k} \in \mathbb{Z}^d}$ . Note that the definition of IMRA does not require the existence of such a function. In fact, we might require several (possibly infinite) scaling functions so that their translates generate  $V_0$ . But in this article, we will restrict ourselves to the case in which  $V_0$  is generated by a single scaling function. We do not require the translates of the scaling function to form an orthonormal basis or even a frame.

**DEFINITION 3.2.** *A finite set of functions,  $\{\psi_i \in L^2(\mathbb{R}^d)\}_{i=1, \dots, n}$ , is called a **frame multiwavelet** if the set  $\{D^j T_{\mathbf{k}} \psi_i : j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$  is a frame for  $L^2(\mathbb{R}^d)$ .*

If  $\psi_i \in V_1$  for all  $i = 1, \dots, n$  then we say that the multiwavelet is associated to the IMRA  $\{V_j\}_{j \in \mathbb{Z}}$ . We can find  $\mathbb{Z}^d$ -periodic functions  $m_i$ , referred to as **high pass filters**, so that

$$\widehat{\psi}_i(A^* \cdot) = m_i \widehat{\phi}, \quad \text{for all } i = 1, \dots, n.$$

We define the **detail spaces** by  $W_j := \overline{\text{span}}\{D^j T_{\mathbf{k}} \psi_i : \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$ . Notice that although, the relation  $W_j \subset V_{j+1}$  holds, the orthogonality of the  $W_j$  and  $V_j$ , which is true for classical MRAs, may not hold in this case. This together with

the fact that we do not require the translates of the scaling function or wavelets to be frames, allows us a lot of flexibility in designing filters. For instance, we are allowed to have smooth filters versus the discontinuous ones which are necessary if we want the translates of the scaling function to form a frame (see [3]). Although, we do not have frames for each of the resolution or detail spaces, we shall prove that under some conditions on the low pass and high pass filters, the set,

$$X_{\phi\psi} := \{D^j T_{\mathbf{k}} \psi_i : j \in \mathbb{N} \cup \{0\}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\} \cup \{T_{\mathbf{k}} \phi : \mathbf{k} \in \mathbb{Z}^d\},$$

is a frame for  $L^2(\mathbb{R}^d)$ . More precisely, we will characterize the families,  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$ , that form a pair of dual frames, in terms of certain conditions on the associated low and high pass filters. This is the content of the next theorem.

**DEFINITION 3.3.** *A function,  $\phi \in L^2(\mathbb{R}^d)$ , is called **refinable** if there exists a  $\mathbb{Z}^d$ -periodic function  $m_0$  such that*

$$\widehat{\phi}(A^* \cdot) = m_0 \widehat{\phi},$$

where  $A$  is a radially expansive matrix. We refer to  $m_0$  as a **low pass filter** associated with  $\phi$ .

The spectrum of  $\phi$ , denoted by  $\sigma(\phi)$ , is defined via

$$\sigma(\phi) := \left\{ \boldsymbol{\xi} \in \mathbb{T}^d : \sum_{\mathbf{k} \in \mathbb{Z}^d} \left| \widehat{\phi}(\boldsymbol{\xi} + \mathbf{k}) \right|^2 > 0 \right\}.$$

The function,  $\boldsymbol{\xi} \rightarrow \sum_{\mathbf{k} \in \mathbb{Z}^d} \left| \widehat{\phi}(\boldsymbol{\xi} + \mathbf{k}) \right|^2$  is referred to as the auto-correlation function for  $\phi$ .

**THEOREM 3.4.** *Let  $\phi^a$  and  $\phi^s$  be refinable functions in  $L^2(\mathbb{R}^d)$  such that  $\widehat{\phi}^a$  and  $\widehat{\phi}^s$  are continuous at the origin and*

$$(7) \quad \lim_{|\boldsymbol{\xi}| \rightarrow 0} \widehat{\phi}^a(\boldsymbol{\xi}) = \widehat{\phi}^s(\boldsymbol{\xi}) = 1.$$

Let  $m_0^a$  and  $m_0^s \in L^\infty(\mathbb{T}^d)$ , be the associated low pass filters. Furthermore, let  $m_i^a, m_i^s$  for  $i = 1, \dots, n$ , be  $\mathbb{Z}^d$ -periodic measurable functions and define  $n$  pairs of wavelets  $\psi_i^a, \psi_i^s$   $i = 1, \dots, n$ , by

$$(8) \quad \widehat{\psi}_i^a(A^* \cdot) = m_i^a \widehat{\phi}^a.$$

$$(9) \quad \widehat{\psi}_i^s(A^* \cdot) = m_i^s \widehat{\phi}^s.$$

Assume that  $m_i^a, m_i^s \in L^\infty(\mathbb{T}^d)$  for all  $i = 0, \dots, n$ . Then  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  form a pair of dual frames for  $L^2(\mathbb{R}^d)$  if and only if

- (1)  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  are Bessel families,
- (2) For all  $\mathbf{q} \in (A^{*-1}\mathbb{Z}^d)/\mathbb{Z}^d$  and for a.e.  $\boldsymbol{\xi}, \boldsymbol{\xi} + \mathbf{q} \in \sigma(\phi^a) \cap \sigma(\phi^s)$ ,

$$(10) \quad \sum_{i=0}^m m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi} + \mathbf{q})} = \delta_{\mathbf{q},0}.$$

Before we prove this theorem, we state and prove a lemma required in the proof. The superscripts  $a$  and  $s$  stand for analysis and synthesis respectively. We refer to (8) and (9), for  $i = 0, \dots, n$ , as the two-scale relations, with the convention,  $\psi_0^a = \phi^a$  and  $\psi_0^s = \phi^s$ .

LEMMA 3.5. *Let  $G$  be a finite group with the group operation denoted by  $\circ$  and let  $\chi$  be a character on  $G$ . If  $\chi$  is not the identity character on  $G$ , i.e. there exists an  $h$  in  $G$  such that  $\chi(h) \neq 1$ , then*

$$\sum_{g \in G} \chi(g) = 0.$$

**Proof:** Let  $Y = \sum_{g \in G} \chi(g)$  and observe

$$\chi(h) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(h \circ g) = Y.$$

Hence, we have  $Y(\chi(h) - 1) = 0$ , which gives us the required result since  $\chi(h) \neq 1$ .  $\square$

**Proof:**[ of Theorem 3.4] The proof has two parts. First, we establish that  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  is a pair of dual frames for  $L^2(\mathbb{R}^d)$ , if and only if the following condition holds,

$$\sum_{j=0}^{j_0(\mathbf{l})} \sum_{i=1}^n \widehat{\psi}^s_i(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}^a_i(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})} = \delta_{\mathbf{l},0}.$$

for a.e.  $\boldsymbol{\xi} \in \mathbb{R}^d$  and  $\mathbf{l} \in \mathbb{Z}^d$  (The function  $j_0$  is defined in (13) below). In the second part we show that this condition is equivalent to (10).

We begin by defining a unitary operator  $U : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{T}^d, l^2(\mathbb{Z}^d))$  via

$$U(f) = \left\{ \widehat{f}(\cdot + \mathbf{k}) : \mathbf{k} \in \mathbb{Z}^d \right\}.$$

Next we define the set  $J$  by

$$J := \{(j, \mathbf{r}, i) : j \in \mathbb{N} \cup \{0\}, \mathbf{r} \in \mathbb{Z}^d / (A^{*j}\mathbb{Z}^d), i = 1, \dots, n\} \cup \{(0, 0, 0)\},$$

and a map  $S^a : L^2(\mathbb{T}^d, l^2(J)) \rightarrow L^2(\mathbb{T}^d, l^2(\mathbb{Z}^d))$ , via

$$S^a(e_{\mathbf{l}}, \delta_{(j, \mathbf{r}, i)}) = U(T_{\mathbf{l}} D^j T_{\mathbf{r}} \psi_i^a) \quad \text{for all } \mathbf{l} \in \mathbb{Z}^d, (j, \mathbf{r}, i) \in J,$$

where we adopt the convention  $\psi_0^a = \phi^a$ .

Note that,

$$(S^a e_{\mathbf{l}} \delta_{(j, \mathbf{r}, i)})(\boldsymbol{\xi}) = \left\{ e_{\mathbf{l}}(\boldsymbol{\xi}) |\det(A)|^{-j/2} e^{-2\pi i \langle \mathbf{r}, A^{*-j}(\boldsymbol{\xi} + \mathbf{k}) \rangle} \widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \right\}_{\mathbf{k} \in \mathbb{Z}^d}$$

for a.e.  $\boldsymbol{\xi} \in \mathbb{T}^d$ . Since  $X_{\phi\psi}^a$  is a Bessel family,  $S^a$  can be extended to a bounded linear transformation on  $L^2(\mathbb{T}^d, l^2(J))$ . Now  $S^a$  can be represented by a  $\mathbb{Z}^d \times J$  matrix so that the  $(\mathbf{k}, (j, \mathbf{r}, i))$ -th entry of the matrix is an operator on  $L^2(\mathbb{T}^d)$ , denoted by  $S_{\mathbf{k}, (j, \mathbf{r}, i)}^a$ . This operator acts on the modulations as follows,

$$(S_{\mathbf{k}, (j, \mathbf{r}, i)}^a e_{\mathbf{l}})(\boldsymbol{\xi}) = e_{\mathbf{l}}(\boldsymbol{\xi}) |\det(A)|^{-j/2} e^{-2\pi i \langle \mathbf{r}, (A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \rangle} \widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})),$$

Thus, we see that each  $S_{\mathbf{k}, (j, \mathbf{r}, i)}^a$  commutes with all the modulation operators  $M_{\mathbf{l}}$ . These modulation operators are defined on  $L^2(\mathbb{T}^d)$  by  $M_{\mathbf{l}}\omega(\boldsymbol{\xi}) := e_{\mathbf{l}}(\boldsymbol{\xi})\omega(\boldsymbol{\xi})$ . Hence, each  $S_{\mathbf{k}, (j, \mathbf{r}, i)}^a$  is a multiplicative operator (e.g. [6, Corollary 2.12.7]) in the following sense:

$$(S_{\mathbf{k}, (j, \mathbf{r}, i)}^a \omega)(\boldsymbol{\xi}) = \omega(\boldsymbol{\xi}) |\det(A)|^{-j/2} e^{-2\pi i \langle \mathbf{r}, (A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \rangle} \widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})),$$

for all  $\omega \in L^2(\mathbb{T}^d)$ . We denote the symbol for this operator again by  $S_{\mathbf{k},(j,\mathbf{r},i)}^a(\cdot)$ , so that,

$$S_{\mathbf{k},(j,\mathbf{r},i)}^a(\boldsymbol{\xi}) = |\det(A)|^{-j/2} e^{-2\pi i \langle \mathbf{r}, (A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \rangle} \widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})).$$

Note that  $S_{\mathbf{k},(j,\mathbf{r},i)}^a(\cdot)$  is an essentially bounded  $\mathbb{Z}^d$ -periodic function. Let  $S^a(\cdot)$  denote the  $\mathbb{Z}^d \times J$  matrix of the symbols  $S_{\mathbf{k},(j,\mathbf{r},i)}^a(\cdot)$ , then

$$(S^a \omega)(\boldsymbol{\xi}) = S^a(\boldsymbol{\xi}) \omega(\boldsymbol{\xi}) \quad \text{for all } \omega \in L^2(\mathbb{T}^d, l^2(J)).$$

Now, the adjoint of  $S^a$  can be represented by the  $J \times \mathbb{Z}^d$  matrix of operators on  $L^2(\mathbb{T}^d)$  such that the  $((j, \mathbf{r}, i), \mathbf{k})$ -th entry is  $S_{\mathbf{k},(j,\mathbf{r},i)}^{a*}$ . Since  $S_{\mathbf{k},(j,\mathbf{r},i)}^{a*}$  is the adjoint of a multiplicative operator, it itself, is a multiplicative operator with symbol  $\overline{S_{\mathbf{k},(j,\mathbf{r},i)}^a(\cdot)}$ . Therefore,

$$(S^{a*} \omega)(\boldsymbol{\xi}) = S^a(\boldsymbol{\xi})^* \omega(\boldsymbol{\xi}) \quad \text{for all } \omega \in L^2(\mathbb{T}^d, l^2(\mathbb{Z}^d)).$$

Similarly, we define  $S^s : L^2(\mathbb{T}^d, l^2(J)) \rightarrow L^2(\mathbb{T}^d, l^2(\mathbb{Z}^d))$ , via

$$S^s(e_{\mathbf{l}}, \delta_{(j,\mathbf{r},i)}) = U(T_{\mathbf{l}} D^j T_{\mathbf{r}} \psi_i^s) \quad \text{for all } \mathbf{l} \in \mathbb{Z}^d, \quad (j, \mathbf{r}, i) \in J,$$

where we adopt the convention  $\psi_0^s = \phi^s$ . Arguing as before for  $S^a$ , we obtain the following expression for  $S^s$ :

$$(S^s \omega)(\boldsymbol{\xi}) = S^s(\boldsymbol{\xi}) \omega(\boldsymbol{\xi}) \quad \text{for all } \omega \in L^2(\mathbb{T}^d, l^2(J)),$$

where  $S^s(\cdot)$  denotes the  $\mathbb{Z}^d \times J$  matrix of the symbols,

$$S_{\mathbf{k},(j,\mathbf{r},i)}^s(\boldsymbol{\xi}) = |\det(A)|^{-j/2} e^{-2\pi i \langle \mathbf{r}, (A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \rangle} \widehat{\psi}_i^s(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})),$$

The families,  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$ , is a pair of dual frames, if and only if,  $S^s S^{a*} = I_{L^2(\mathbb{T}^d, l^2(\mathbb{Z}^d))}$ , i.e.

$$(11) \quad S^s(\boldsymbol{\xi}) S^a(\boldsymbol{\xi})^* = I_{l^2(\mathbb{Z}^d)} \quad \text{for a.e. } \boldsymbol{\xi} \in \mathbb{T}^d.$$

The  $(\mathbf{k}, \mathbf{l})$ -th entry of  $S^s(\boldsymbol{\xi}) S^a(\boldsymbol{\xi})^*$  is given by

$$(12) \quad \begin{aligned} & (S^s(\boldsymbol{\xi}) S^a(\boldsymbol{\xi})^*)_{(\mathbf{k}, \mathbf{l})} \\ &= \sum_{j=0}^{\infty} \sum_{i=1}^n |\det(A)|^{-j} \sum_{\mathbf{r} \in \mathbb{Z}^d / (A^{*j} \mathbb{Z}^d)} e^{2\pi i \langle \mathbf{r}, A^{*-j}(\mathbf{k} - \mathbf{l}) \rangle} \widehat{\psi}_i^s(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} \\ & \quad + \widehat{\phi^s}(\boldsymbol{\xi} + \mathbf{k}) \overline{\widehat{\phi^a}(\boldsymbol{\xi} + \mathbf{l})}. \end{aligned}$$

Define  $j_0 : \mathbb{Z}^d \rightarrow \mathbb{Z}$  via

$$(13) \quad j_0(\mathbf{l}) = \sup\{j : A^{*-j}(\mathbf{l}) \in \mathbb{Z}^d\} \quad \text{for all } \mathbf{l} \in \mathbb{Z}^d,$$

with the convention that  $j_0(0) = +\infty$ .

Now, if  $j > j_0(\mathbf{k} - \mathbf{l})$ , then  $A^{*-j}(\mathbf{k} - \mathbf{l}) \notin \mathbb{Z}^d$  and therefore  $e_{A^{*-j}(\mathbf{k} - \mathbf{l})}$  is not the identity character on  $\mathbb{Z}^d / (A^{*j} \mathbb{Z}^d)$ . On the other hand, if  $j \leq j_0(\mathbf{k} - \mathbf{l})$  then  $A^{*-j}(\mathbf{k} - \mathbf{l}) \in \mathbb{Z}^d$  and  $e_{A^{*-j}(\mathbf{k} - \mathbf{l})}$  is the identity character on  $\mathbb{Z}^d / (A^{*j} \mathbb{Z}^d)$ . Thus using Lemma 3.5, we conclude

$$|\det(A)|^{-j} \sum_{\mathbf{r} \in \mathbb{Z}^d / (A^{*j} \mathbb{Z}^d)} e^{2\pi i \langle \mathbf{r}, A^{*-j}(\mathbf{k} - \mathbf{l}) \rangle} = \begin{cases} 0 & \text{for } j > j_0(\mathbf{k} - \mathbf{l}) \\ 1 & \text{for } j \leq j_0(\mathbf{k} - \mathbf{l}) \end{cases}.$$



Hence, for  $\mathbf{k} \neq \mathbf{l}$ , (12) reduces to,

$$(S^s(\boldsymbol{\xi})S^a(\boldsymbol{\xi}^*))_{(\mathbf{k},\mathbf{l})} = \sum_{j=0}^{j_0(\mathbf{k}-\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} \\ + \widehat{\phi}^s(\boldsymbol{\xi} + \mathbf{k}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})}.$$

Thus using (11),  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  is a pair of dual frames if and only if the following condition holds for a.e.  $\boldsymbol{\xi} \in \mathbb{T}^d$  and  $\mathbf{k}, \mathbf{l} \in \mathbb{Z}^d$ :

$$(14) \quad \sum_{j=0}^{j_0(\mathbf{k}-\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}(\boldsymbol{\xi} + \mathbf{k})) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(\boldsymbol{\xi} + \mathbf{k}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})} = \delta_{\mathbf{k},\mathbf{l}}.$$

Changing variables, in the previous equation gives that it is equivalent to:

$$(15) \quad \sum_{j=0}^{j_0(\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})} = \delta_{\mathbf{l},0}.$$

for a.e.  $\boldsymbol{\xi} \in \mathbb{R}^d$  and  $\mathbf{l} \in \mathbb{Z}^d$ . This finishes the first part of the proof.

Now, we proceed to the second part of the proof. In this part we show that (15) holds if and only if (10) holds. We will first assume (10) holds and prove that (15) holds for all  $\mathbf{l} \neq 0$ . Using the two-scale relations ((8) and (9)) and (10) we obtain,

$$\begin{aligned} & \sum_{j=0}^{j_0(\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})} = \\ &= \sum_{j=1}^{j_0(\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} \\ & \quad + \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}(\boldsymbol{\xi} + \mathbf{l}))} \left( \sum_{i=0}^n m_i^s(A^{*-1}\boldsymbol{\xi}) \overline{m_i^a(A^{*-1}(\boldsymbol{\xi} + \mathbf{l}))} \right), \\ &= \sum_{j=1}^{j_0(\mathbf{l})} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}(\boldsymbol{\xi} + \mathbf{l}))}, \\ & \quad \vdots \\ &= \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j_0(\mathbf{l})}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j_0(\mathbf{l})}(\boldsymbol{\xi} + \mathbf{l}))} + \\ & \quad \widehat{\phi}^s(A^{*-j_0(\mathbf{l})}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-j_0(\mathbf{l})}(\boldsymbol{\xi} + \mathbf{l}))} \left( \sum_{i=0}^n m_i^s(A^{*-j_0(\mathbf{l})}\boldsymbol{\xi}) \overline{m_i^a(A^{*-j_0(\mathbf{l})}(\boldsymbol{\xi} + \mathbf{l}))} \right), \\ &= \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j_0(\mathbf{l})}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j_0(\mathbf{l})}(\boldsymbol{\xi} + \mathbf{l}))} + \widehat{\phi}^s(A^{*-j_0(\mathbf{l})}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-j_0(\mathbf{l})}(\boldsymbol{\xi} + \mathbf{l}))}, \\ &= \widehat{\phi}^s(A^{*-(j_0(\mathbf{l})+1)}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-(j_0(\mathbf{l})+1)}(\boldsymbol{\xi} + \mathbf{l}))} \\ & \quad \left( \sum_{i=0}^n m_i^s(A^{*-(j_0(\mathbf{l})+1)}\boldsymbol{\xi}) \overline{m_i^a(A^{*-(j_0(\mathbf{l})+1)}(\boldsymbol{\xi} + \mathbf{l}))} \right) \\ &= 0. \end{aligned}$$

Next, using the hypothesis (10), we show that (15) holds for  $\mathbf{l} = 0$ ,

$$\begin{aligned}
& \sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} = \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} \\
&\quad + \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}\boldsymbol{\xi})} \left( \sum_{i=0}^n m_i^s(A^{*-1}\boldsymbol{\xi}) \overline{m_i^a(A^{*-1}\boldsymbol{\xi})} \right), \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}\boldsymbol{\xi})}, \\
&\quad \vdots \\
&= \sum_{j=N}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} \\
&\quad + \widehat{\phi}^s(A^{*-N}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-N}\boldsymbol{\xi})} \left( \sum_{i=0}^n m_i^s(A^{*-N}\boldsymbol{\xi}) \overline{m_i^a(A^{*-N}\boldsymbol{\xi})} \right), \\
&= \underbrace{\sum_{j=N}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})}}_{\rightarrow 0 \text{ as } N \rightarrow \infty \text{ because the series converges for a.e. } \boldsymbol{\xi} \in \mathbb{R}^d} \\
&\quad + \underbrace{\widehat{\phi}^s(A^{*-N}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-N}\boldsymbol{\xi})}}_{\rightarrow 1 \text{ as } N \rightarrow \infty \text{ by (7) and continuity of } \widehat{\phi}^s, \widehat{\phi}^a \text{ at the origin.}}, \\
&\rightarrow 1 \text{ as } N \rightarrow \infty \quad \text{for almost every } \boldsymbol{\xi} \in L^2(\mathbb{R}^d).
\end{aligned}$$

Thus, we have shown that (10) implies (15).

Next we establish that the converse implication is valid. Assume first, that (10) holds for  $\mathbf{q} = 0$ . Under the assumption, the calculations we carried out to prove that Eq. (10) implies Eq. (15) for  $\mathbf{l} \neq 0$  are still valid for almost every  $\boldsymbol{\xi}$  and all  $\mathbf{l} \in \mathbb{Z}^d$ . Now, pick  $\mathbf{q} \in (A^{*-1}\mathbb{Z}^d)/\mathbb{Z}^d$ . We have  $A^*\mathbf{q} \in \mathbb{Z}^d$ . If  $\mathbf{p}$  is an arbitrary integer grid point, i.e.  $\mathbf{p} \in \mathbb{Z}^d$  set  $\mathbf{l} := A^*(\mathbf{q} + \mathbf{p})$ . Obviously  $\mathbf{l} \in \mathbb{Z}^d$ , since  $A$  is expansive and, thus,  $A^*$  leaves the lattice  $\mathbb{Z}^d$  invariant. Observe  $j_0(\mathbf{l}) = 0$ , because  $\mathbf{q} + \mathbf{p}$  does not belong to the integer grid. Since, (15) is valid for all  $\mathbf{l} \neq 0$  we deduce

$$\begin{aligned}
0 &= \sum_{i=1}^n \widehat{\psi}_i^s(\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(\boldsymbol{\xi} + \mathbf{l})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{l})} \\
&= \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}(\boldsymbol{\xi} + \mathbf{l}))} \left( \sum_{i=0}^n m_i^s(A^{*-1}\boldsymbol{\xi}) \overline{m_i^a(A^{*-1}(\boldsymbol{\xi} + \mathbf{l}))} \right) \\
&= \widehat{\phi}^s(A^{*-1}\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^{*-1}\boldsymbol{\xi} + \mathbf{q} + \mathbf{p})} \left( \sum_{i=0}^n m_i^s(A^{*-1}\boldsymbol{\xi}) \overline{m_i^a(A^{*-1}\boldsymbol{\xi} + \mathbf{q} + \mathbf{p})} \right) \quad \text{a.e. in } \mathbb{R}^d.
\end{aligned}$$

Take  $\boldsymbol{\xi}$  and  $\boldsymbol{\xi} + \mathbf{q}$  in  $\sigma(\phi^a) \cap \sigma(\phi^s)$  such that  $A^*\boldsymbol{\xi}$  belongs to the set of points in  $\mathbb{R}^d$  for which Eq. (15) holds. The fact that Eq. (15) is equivalent to Eq. (14) implies that, if Eq. (15) holds for a  $\boldsymbol{\xi} \in \mathbb{R}^d$  then, Eq. (15) holds for all integer translates

of this point. Next, apply the previous equations for  $A^*(\boldsymbol{\xi} + \boldsymbol{\lambda})$  instead of  $\boldsymbol{\xi}$ , where  $\boldsymbol{\lambda} \in \mathbb{Z}^d$ . Then,

$$0 = \widehat{\phi}^s(\boldsymbol{\xi} + \boldsymbol{\lambda}) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{q} + \mathbf{p})} \left( \sum_{i=0}^n m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi} + \mathbf{q})} \right)$$

due to the  $\mathbb{Z}^d$ -periodicity of  $m_i^a$  and  $m_i^s$ . Since  $\mathbf{p}, \boldsymbol{\lambda}$  are arbitrary integers and  $\boldsymbol{\xi}, \boldsymbol{\xi} + \mathbf{q} \in \sigma(\phi^a) \cap \sigma(\phi^s)$  there exist some  $\boldsymbol{\lambda}_0$  and  $\mathbf{p}_0$  for which

$$\widehat{\phi}^s(\boldsymbol{\xi} + \boldsymbol{\lambda}_0) \overline{\widehat{\phi}^a(\boldsymbol{\xi} + \mathbf{q} + \mathbf{p}_0)} \neq 0.$$

Thus,

$$\sum_{i=0}^n m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi} + \mathbf{q})} = 0.$$

In order to complete the proof of the theorem we need to establish that (10) holds for  $\mathbf{q} = 0$ . To verify this set  $l = 0$  in (15):

$$\sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} = 1.$$

Observe that this implies,

$$\sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j+1}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j+1}\boldsymbol{\xi})} + \widehat{\phi}^s(A^*\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^*\boldsymbol{\xi})} = 1.$$

Using the two scale relations, (8) and (9), we infer,

$$\begin{aligned} & \sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j+1}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j+1}\boldsymbol{\xi})} + \widehat{\phi}^s(A^*\boldsymbol{\xi}) \overline{\widehat{\phi}^a(A^*\boldsymbol{\xi})} = \\ & = \sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} \left( \sum_{i=0}^n m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi})} \right), \end{aligned}$$

hence,

$$\begin{aligned} & \sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} = \\ & = \sum_{j=0}^{\infty} \sum_{i=1}^n \widehat{\psi}_i^s(A^{*-j}\boldsymbol{\xi}) \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} + \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} \left( \sum_{i=0}^n m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi})} \right), \end{aligned}$$

Therefore,  $\sum_{i=0}^n m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi})} = 1$  for a.e.  $\boldsymbol{\xi} \in \sigma(\phi^a) \cap \sigma(\phi^s)$ .  $\square$

**REMARK 3.6.** *The conditions on the filters in (10) are similar to those required for Proposition 5.2 in [7] (also [4]), from which the so-called Mixed Extension Principle follows. The statement of the previous theorem does not require the so-called Mixed Fundamental Function that is present in the said result in [7]. At first glance, this fact seems contradictory but it is not. Under the assumptions of Theorem 3.4, the Mixed Fundamental Function is equal to one almost everywhere. This follows from the proof of the so-called Mixed Oblique Extension Principle (Corollary 5.3 in [7]) if one takes into consideration the following fact. The condition*

that the sets  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  form a pair of dual frames of  $L^2(\mathbb{R}^d)$  is stronger than the condition that the affine families  $\{D^j T_{\mathbf{k}} \psi_i^a : j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$  and  $\{D^j T_{\mathbf{k}} \psi_i^s : j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$  form a pair of dual frames of  $L^2(\mathbb{R}^d)$ .

We now describe a discrete wavelet decomposition and reconstruction algorithm, first presented in [2] in the context of exact reconstruction isotropic filter banks. We use the synthesis low pass filters to construct isotropic refinable functions in order to produce IMRA examples. We show that the low and high pass filters of these isotropic filter banks satisfy the condition (10), which is an exact reconstruction filter bank equation. Hence, by Theorem 3.4 the associated families  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  form a pair of dual frames for  $L^2(\mathbb{R}^d)$ .

The wavelet decomposition will be defined in terms of the analysis filters, denoted with the superscript  $a$ , and the reconstruction will be defined by the synthesis filters, denoted with the superscript  $s$ . This is consistent with the notation in Theorem 3.4. Let  $A = a\mathcal{O}$ , be the radially expansive matrix used to define the dilation operator, where  $a > 1$  and  $\mathcal{O}$ , a unitary  $d \times d$  matrix. Now, let  $\frac{1}{2a^2} < b_3 < b_2 < b_1 < b_0 < \frac{1}{2a}$ . We denote the balls of radii  $ab_1$  and  $ab_2$  by  $\mathbb{B}_1$  and  $\mathbb{B}_2$  respectively.

We begin with the analysis low-pass filter, a smooth,  $\mathbb{Z}^d$ -periodic function  $m_0^a$  satisfying the following three properties:

- $m_0^a = 1$  inside the ball of radius  $b_2$
- $m_0^a = 0$  on  $\mathbb{T}^d \setminus B(0, b_1)$
- $m_0^a|_{\mathbb{T}^d}$  is radial,

We define  $\phi^a$  by  $\widehat{\phi^a} = m_0^a(A^{*-1})\chi_{\mathbb{T}^d}$ .

Now, let  $h^a$  be a  $\mathbb{Z}^d$ -periodic function defined by,

$$(16) \quad h^a(\boldsymbol{\xi}) = \frac{1 - m_0^a(\boldsymbol{\xi})}{|\det(A)|^{1/2}}.$$

Using  $h^a$  we define the high pass filters as follows:

$$(17) \quad m_i^a(\boldsymbol{\xi}) = e_{\mathbf{q}_{i-1}}(\boldsymbol{\xi})h^a(\boldsymbol{\xi}), \quad i = 1, \dots, |\det(A)|.$$

where  $\{\mathbf{q}_l : l = 0, 1, 2, \dots, |\det(A)| - 1\}$  are the representatives of the quotient group  $\mathbb{Z}^d / (A^* \mathbb{Z}^d)$ .

**REMARK 3.7.** We note that the set  $\{e_{\mathbf{k}} h^a : \mathbf{k} \in \mathbb{Z}^d\}$  is equal to  $\cup_{i=1}^{|\det(A)|} \{e_{2\mathbf{k}} m_i^a : \mathbf{k} \in \mathbb{Z}^d\}$ . The latter union would yield  $|\det(A)|$  high pass channels with decimation. In our implementation, all these channels are combined into an undecimated one using the filter  $h^a$ , which is the  $\mathbb{Z}^d$  periodic function defined in (16).

Accordingly, in our implementation the low-pass filtering is followed by decimation determined by  $A$  while the high pass filtered signal stays undecimated.

We use upper case characters to denote each filter in the spatial domain. Using this notation, the analysis filters produce analysis operators  $\mathcal{M}_a$  and  $\mathcal{H}_a$  given by

$$(18) \quad \mathcal{M}_a \sigma := |\det(A)|^{1/2} \mathcal{D}(M_a * \sigma) \quad \text{and} \quad \mathcal{H}_a \sigma := |\det(A)|^{1/2} H_a * \sigma, \quad \sigma \in \ell^2(\mathbb{Z}^d),$$

where  $\mathcal{D}$  denotes downsampling operator given by,

$$(\mathcal{D}\sigma)^\wedge(\boldsymbol{\xi}) = \frac{1}{|\det(A)|} \sum_{l=1}^{|\det(A)|} \widehat{\sigma}(A^{*-1}\boldsymbol{\xi} + \boldsymbol{\gamma}_l),$$

and  $\{\boldsymbol{\gamma}_l : l = 1, \dots, |\det(A)|\}$  are the representatives of the quotient group

$(A^{*-1}\mathbb{Z}^d)/\mathbb{Z}^d$ .

The range space of the operator  $\sigma \mapsto M_a * \sigma$  is the subspace of all signals whose Fourier transforms vanish outside the ball  $\{\xi : |\xi| \leq b_1\}$ . Likewise, the range of  $H_a$  contains all signals whose Fourier transforms vanish inside the ball  $\{\xi : |\xi| < b_2\}$ . Both signals  $M_a * \sigma$  and  $H_a * \sigma$  are oversampled, because the frequency support of the functions in the range of  $M_a$  or  $H_a$  is not the entire torus  $\mathbb{T}^d$ . In analogy with one-dimensional oversampling, we now use synthesis filters for the reconstruction with a bandwidth slightly greater than the bandwidth of the analysis filters.

We define the synthesis low pass filter as a smooth,  $\mathbb{Z}^d$ -periodic function  $m_0^s$  satisfying the following three properties:

- $m_0^s = 1$  inside the ball of radius  $b_1$
- $m_0^s = 0$  on  $\mathbb{T}^d \setminus B(0, b_0)$
- $m_0^s|_{\mathbb{T}^d}$  is radial,

and the synthesis high pass filter as a smooth,  $\mathbb{Z}^d$ -periodic function  $h^s$  satisfying the following three properties:

- $h^s = 0$  inside the ball of radius  $b_3$
- $h^s = \frac{1}{|\det(A)|^{1/2}}$  on  $\mathbb{T}^d \setminus B(0, b_2)$
- $h^s|_{\mathbb{T}^d}$  is radial.

We define  $|\det(A)|$  synthesis high pass filters by  $m_i^s = e_{\mathbf{q}_{i-1}} h^s$  for each  $i = 1, \dots, |\det(A)|$ , corresponding to the  $|\det(A)|$  analysis high pass channels defined above. As in the case of analysis high pass filters, the  $|\det(A)|$  high pass channels are combined into a single high pass channel with filter  $h^s$ . We define the high pass synthesis operator by  $\mathcal{H}_s \sigma = |\det(A)|^{1/2} H_s * \sigma$  and the operator  $\mathcal{M}_s \sigma := |\det(A)|^{1/2} M_s * (\mathcal{U}\sigma)$  is the low pass synthesis operator, where  $\sigma \in \ell^2(\mathbb{Z}^d)$  and  $\mathcal{U}$  is the upsampling operator,  $(\mathcal{U}\sigma)^\wedge(\xi) = \widehat{\sigma}(A^*\xi)$ .

**PROPOSITION 3.8** (Exact reconstruction formula). *Let  $\mathcal{M}_s$ ,  $\mathcal{M}_a$ ,  $\mathcal{H}_s$  and  $\mathcal{H}_a$  be operators on  $\ell^2(\mathbb{Z}^d)$  as defined above, then the following identity holds*

$$I_{\ell^2(\mathbb{Z}^d)} = \mathcal{M}_s \mathcal{M}_a + \mathcal{H}_s \mathcal{H}_a$$

**Proof:** Let  $\sigma \in \ell^2(\mathbb{Z}^d)$  then  $(\mathcal{H}_s \mathcal{H}_a \sigma)^\wedge = |\det(A)| h_s h_a \widehat{\sigma} = |\det(A)|^{1/2} h_a \widehat{\sigma}$ . From the definition of  $\mathcal{D}$  and  $\mathcal{U}$ , we infer  $(\mathcal{U} \mathcal{M}_a \sigma)^\wedge$  is a  $A^{*-1}\mathbb{Z}^d$ -periodic function. Since  $m_s$  and  $m_a$  vanish outside the  $d$ -torus  $A^{*-1}\mathbb{T}^d$  and the ball  $A^{*-1}\mathbb{B}_1$  respectively, we obtain  $(\mathcal{M}_s \mathcal{M}_a \sigma)^\wedge = m_s m_a \widehat{\sigma} = m_a \widehat{\sigma}$ . The exactness of reconstruction is verified once we observe  $m_a + |\det(A)|^{1/2} h_a = 1$ .  $\square$

**REMARK 3.9.** *The exact reconstruction formula holds even if we use  $\frac{1}{|\det(A)|^{1/2}} I_{\ell^2(\mathbb{Z}^d)}$  instead of  $\mathcal{H}_s$ . Hence, for the implementation, we do not apply the synthesis high pass filter. Nevertheless, we mention it here because if we use  $\frac{1}{|\det(A)|^{1/2}} I_{\ell^2(\mathbb{Z}^d)}$  instead of  $\mathcal{H}_s$  as the high pass synthesis operator then the family,  $X_{\phi\psi}^s$ , is not Bessel.*

**REMARK 3.10.** *In contrast to the usual tensor-product constructions both filters (low pass and high pass) are isotropic. The analysis and synthesis scaling functions are  $C^\infty$ -smooth in the wavenumber domain by definition. This, in turn, implies that the wavelets are also  $C^\infty$ -smooth in the wavenumber domain. Thus, the radial scaling functions and their associated wavelets have rapid decay in the spatial domain, a property which is very useful for handling the lateral variations of the velocity,  $c$ .*

We now define  $V_0 := \overline{\text{span}\{T_{\mathbf{k}}\phi^s\}_{\mathbf{k} \in \mathbb{Z}^d}}$  and  $V_j := D^j V_0$ , where  $\phi^s$  is defined by  $\widehat{\phi^s} = m_0^a(A^{*-1})\chi_{\mathbb{T}^d}$ . It is easy to check that  $\{V_j\}_{j \in \mathbb{Z}}$  forms an IMRA.

Next, we show that the analysis and synthesis filters defined above satisfy condition (10) for  $\mathbf{q} = 0$ :

$$\sum_{i=0}^{|\det(A)|} m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi})} = m_0^a(\boldsymbol{\xi}) + |\det(A)| \frac{1 - m_0^a(\boldsymbol{\xi})}{|\det(A)|} = 1 \text{ for all } \boldsymbol{\xi} \in \mathbb{T}^d.$$

An alternative way to verify this equation is to use Proposition 3.8. Now, for  $\mathbf{q} \in (A^{*-1}\mathbb{Z}^d)/\mathbb{Z}^d$  such that  $\mathbf{q} \neq 0$ , we have

$$m_0^s(\boldsymbol{\xi}) \overline{m_0^a(\boldsymbol{\xi} + A^{*-1}\mathbf{q})} = 0 \text{ for all } \boldsymbol{\xi} \in \mathbb{T}^d, \mathbf{q} \neq 0.$$

This is because the supports of  $m_0^s$  and  $\overline{m_0^a(\cdot + \mathbf{q})}$  are disjoint by definition. Hence,

$$\begin{aligned} \sum_{i=0}^{|\det(A)|} m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi} + \mathbf{q})} &= \sum_{i=1}^{|\det(A)|} m_i^s(\boldsymbol{\xi}) \overline{m_i^a(\boldsymbol{\xi} + \mathbf{q})} \\ &= \frac{1}{|\det(A)|^{1/2}} \overline{h_a(\boldsymbol{\xi} + \mathbf{q})} \underbrace{\left( \sum_{i=1}^{|\det(A)|} e_{\mathbf{q}_{i-1}}(\boldsymbol{\xi}) \overline{e_{\mathbf{q}_{i-1}}(\boldsymbol{\xi} + \mathbf{q})} \right)}_{=: s_{\mathbf{q}}} \end{aligned}$$

where  $\{\mathbf{q}_l : l = 0, 1, 2, \dots, |\det(A)| - 1\}$  are the representatives of the quotient group  $\mathbb{Z}^d/(A^*\mathbb{Z}^d)$ . Now, we need to show that  $s_{\mathbf{q}}$  is zero for all  $\mathbf{q} \neq 0$ .

$$s_{\mathbf{q}} = \sum_{i=1}^{|\det(A)|} e_{\mathbf{q}_{i-1}}(\boldsymbol{\xi}) \overline{e_{\mathbf{q}_{i-1}}(\boldsymbol{\xi} + \mathbf{q})} = \sum_{i=1}^{|\det(A)|} e_{\mathbf{q}_{i-1}}(-\mathbf{q}) = \sum_{i=0}^{|\det(A)|-1} e^{2\pi i \langle \mathbf{q}_i, \mathbf{q} \rangle}$$

Since,  $\mathbf{q} \neq 0$  and we know that  $\mathbf{q} \in (A^{*-1}\mathbb{Z}^d)/\mathbb{Z}^d$ , we have  $\mathbf{q} \notin \mathbb{Z}^d$ . Therefore,  $e_{\mathbf{q}}$  is not the identity character on the quotient group  $\mathbb{Z}^d/(A^*\mathbb{Z}^d)$ . Hence, by Lemma 3.5,  $s_{\mathbf{q}} = 0$ . Thus, the low and high pass filters obey the conditions in (10). All the filters are bounded and the scaling functions,  $\phi^a$  and  $\phi^s$  are continuous and equal to one at the origin by definition. Hence,  $X_{\phi^a}^a$  and  $X_{\phi^s}^s$  form a pair of dual frames for  $L^2(\mathbb{R}^d)$  by Theorem 3.4, if we can show that they are both Bessel.

We prove that  $X_{\phi^a}^a$  is a Bessel family. The proof for  $X_{\phi^s}^s$  follows from the same arguments.

Recall that  $X_{\phi^a}^a = \{D^j T_{\mathbf{k}} \psi_i^a : j \in \mathbb{N} \cup \{0\}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\} \cup \{T_{\mathbf{k}} \phi^a : \mathbf{k} \in \mathbb{Z}^d\}$ . It is easy to see that  $\{T_{\mathbf{k}} \phi^a : \mathbf{k} \in \mathbb{Z}^d\}$  is a Bessel family because the associated auto-correlation function is bounded (see Theorem 7.2.3 in [3]). Hence, it is enough to show that  $\{D^j T_{\mathbf{k}} \psi_i^a : j \in \mathbb{N} \cup \{0\}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$  is a Bessel sequence. For a fixed  $i$  and a fixed  $j$ , we have,

$$\begin{aligned} &\sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, D^j T_{\mathbf{k}} \psi_i^a \rangle|^2 = \\ &= \sum_{\mathbf{k} \in \mathbb{Z}^d} \frac{1}{|\det(A)|^j} \left| \int_{\mathbb{R}^d} \widehat{f}(\boldsymbol{\xi}) e_{\mathbf{k}(A^{*-j}\boldsymbol{\xi})} \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi})} d\boldsymbol{\xi} \right|^2 \\ &= \sum_{\mathbf{k} \in \mathbb{Z}^d} \frac{1}{|\det(A)|^j} \left| \sum_{\mathbf{l} \in \mathbb{Z}^d} \int_{A^{*j}\mathbb{T}^d} \widehat{f}(\boldsymbol{\xi} + A^{*j}\mathbf{l}) e_{\mathbf{k}(A^{*-j}\boldsymbol{\xi})} \overline{\widehat{\psi}_i^a(A^{*-j}\boldsymbol{\xi} + \mathbf{l})} d\boldsymbol{\xi} \right|^2 \end{aligned}$$

The sum over  $\mathbf{l} \in \mathbb{Z}^d$  in the last equality reduces to a sum over  $\mathbf{l} \in F$ , where  $F$  is a finite subset of  $\mathbb{Z}^d$  which remains *fixed for all  $j$* , because  $\widehat{\psi}_i^a$  has compact support. Now, applying Plancherel's theorem to the  $A^j \mathbb{Z}^d$ -periodic function  $\sum_{\mathbf{l} \in \mathbb{Z}^d} \widehat{f}(\cdot + \mathbf{l}) \widehat{\psi}_i^a(A^{*-j} \cdot + \mathbf{l})$ , we have

$$\begin{aligned}
\sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \sum_{\mathbf{k} \in \mathbb{Z}^d} |\langle f, D^j T_{\mathbf{k}} \psi_i^a \rangle|^2 &= \sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \int_{A^{*j} \mathbb{T}^d} \left| \sum_{\mathbf{l} \in F} \widehat{f}(\boldsymbol{\xi} + A^{*j} \mathbf{l}) \overline{\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\xi} + \mathbf{l})} \right|^2 d\boldsymbol{\xi} \\
&\leq \sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \int_{A^{*j} \mathbb{T}^d} \left[ \sum_{\mathbf{l} \in F} |\widehat{f}(\boldsymbol{\xi} + A^{*j} \mathbf{l})|^2 \right] \left[ \sum_{\mathbf{l}' \in F} |\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\xi} + \mathbf{l}')|^2 \right] d\boldsymbol{\xi} \\
&= \sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \sum_{\mathbf{l} \in F} \int_{A^{*j}(\mathbb{T}^d + \mathbf{l})} |\widehat{f}(\boldsymbol{\eta})|^2 \left[ \sum_{\mathbf{l}' \in F} |\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\eta} - \mathbf{l} + \mathbf{l}')|^2 \right] d\boldsymbol{\eta} \\
&\leq \sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \sum_{\mathbf{l} \in F} \int_{A^{*j}(\mathbb{T}^d + \mathbf{l})} |\widehat{f}(\boldsymbol{\eta})|^2 \left[ \sum_{\mathbf{q} \in F-F} |\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\eta} + \mathbf{q})|^2 \right] d\boldsymbol{\eta} \\
&\leq \int_{\mathbb{R}^d} |\widehat{f}(\boldsymbol{\eta})|^2 \sum_{i=1}^{|\det(A)|} \sum_{j=0}^{\infty} \left[ \sum_{\mathbf{q} \in F-F} |\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\eta} + \mathbf{q})|^2 \right] d\boldsymbol{\eta}
\end{aligned}$$

Using the definition of the  $\psi_i^a$ 's one can easily establish

$$\sum_{j=0}^{\infty} |\widehat{\psi}_i^a(A^{*-j} \boldsymbol{\eta} + \mathbf{q})|^2 \leq \frac{2}{|\det(A)|}$$

for every  $\mathbf{q}$  which implies that  $\{D^j T_{\mathbf{k}} \psi_i^a : j \in \mathbb{N} \cup \{0\}, \mathbf{k} \in \mathbb{Z}^d, i = 1, \dots, n\}$  is a Bessel sequence.

#### 4. Application of the First Generation IMRA to the solution of the one-way acoustic wave equation

In this section we describe the discretisation of the one-way wave equation in terms of the dual pairs of frames,  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$ , obtained in Section 3.

For a fixed depth,  $z$ , and a frequency,  $\omega$ , the samples of the wavefield,  $f$  on  $\mathbb{Z}^2$ , are assumed to be equal to the coefficients  $\langle f, T_{\mathbf{k}} \phi^a \rangle$  where  $\phi^a$  is the two dimensional IMRA analysis scaling function. This is based on the assumption that the wavefield is band-limited in the ball,  $\mathbb{B}_2$ . Indeed, for all such  $f$ , we have:

$$\begin{aligned}
f(\mathbf{k}) &= \int_{\mathbb{R}^2} \widehat{f}(\boldsymbol{\xi}) e^{2\pi i \langle \mathbf{k}, \boldsymbol{\xi} \rangle} d\boldsymbol{\xi} = \int_{\mathbb{R}^2} \widehat{f}(\boldsymbol{\xi}) \widehat{\phi}^a(\boldsymbol{\xi}) e^{2\pi i \langle \mathbf{k}, \boldsymbol{\xi} \rangle} d\boldsymbol{\xi} \\
(19) \quad &= \langle \widehat{f}, \widehat{\phi}^a e_{\mathbf{k}} \rangle = \langle f, T_{\mathbf{k}} \phi^a \rangle \quad \text{for all } \mathbf{k} \in \mathbb{Z}^2.
\end{aligned}$$

The second equality is true because  $\widehat{\phi}^a$  is equal to 1 on the support of  $\widehat{f}$  and the last equality follows from Plancherel's theorem. Now, if the signal (function) has a higher band limit, then we sample the signal on a finer grid and the samples are then the coefficients  $\langle f, D^j T_{\mathbf{k}} \phi^a \rangle$  for an appropriate scale  $j > 0$ .

Recall that the families  $X_{\phi\psi}^a$  and  $X_{\phi\psi}^s$  form a pair of dual frames for  $L^2(\mathbb{R}^2)$ . This implies that  $D^{j_0} X_{\phi\psi}^a$  and  $D^{j_0} X_{\phi\psi}^s$  form a pair of dual frames for  $L^2(\mathbb{R}^2)$  because

$D^{j_0}$  is a unitary operator for all  $j_0 \in \mathbb{Z}$ . Hence, every  $f \in L^2(\mathbb{R}^2)$  can be written as:

$$(20) \quad f = \sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, D^{j_0} T_{\mathbf{k}} \phi^a \rangle D^{j_0} T_{\mathbf{k}} \phi^s + \sum_{j=j_0}^{\infty} \sum_{i=1}^{|\det(A)|} \sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, D^j T_{\mathbf{k}} \psi_i^a \rangle D^j T_{\mathbf{k}} \psi_i^s.$$

In the special case that  $f$  is band limited to the disk of radius  $a^{j_0+1}b_2$  then  $\langle f, D^j T_{\mathbf{k}} \psi_i^a \rangle = 0$  for all  $j \geq j_0, i = 1, \dots, |\det(A)|, \mathbf{k} \in \mathbb{Z}^2$ , and  $\langle f, D^{j_0} T_{\mathbf{k}} \phi^a \rangle = f(A^{-j_0} \mathbf{k})$  for all  $\mathbf{k} \in \mathbb{Z}^2$ . Hence, the representation of  $f$  in (20) reduces to,

$$(21) \quad f = \sum_{\mathbf{k} \in \mathbb{Z}^2} f(A^{-j_0} \mathbf{k}) D^{j_0} T_{\mathbf{k}} \phi^s.$$

We refer to this expression as the representation at the  $j_0$  resolution level. In practise, the resolution level is not quantified. All data sets are delivered on a sampling grid. Thus, we always adopt the convention that the wavefield,  $f$ , lies in the zero-resolution space, i.e.  $f$  can be expressed by means of (21) with  $j_0 = 0$ .

**4.1. Discretisation of the propagator.** Let  $\mathcal{P}$  denote the propagator operator and  $k_0$  the ratio  $\frac{\omega^2}{c_0^2}$ , where  $c_0$  is the velocity which is assumed to be constant for the moment. The discretisation of  $\mathcal{P}$  turns out to be a Toeplitz matrix in this case. In the wavenumber domain,  $\mathcal{P}$  acts via multiplication with the symbol of the operator,

$$\widehat{\mathcal{P}(f)}(\boldsymbol{\xi}) = e^{2\pi i \Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} \widehat{f}(\boldsymbol{\xi}) \quad \boldsymbol{\xi} \in \mathbb{R}^2.$$

Next, we plug in the frame decomposition (20) to obtain a discretisation of  $\mathcal{P}$ . First,

$$(22) \quad \mathcal{P}(f) = \sum_{\mathbf{l} \in \mathbb{Z}^2} \langle \mathcal{P}(f), T_{\mathbf{l}} \phi^a \rangle T_{\mathbf{l}} \phi^s + \sum_{j=0}^{\infty} \sum_{i=1}^{|\det(A)|} \sum_{\mathbf{l} \in \mathbb{Z}^2} \langle \mathcal{P}(f), D^j T_{\mathbf{l}} \psi_i^a \rangle D^j T_{\mathbf{l}} \psi_i^s.$$

Before we proceed, note that the support of a function in the wavenumber domain is invariant under the action of the propagator,  $\mathcal{P}$  because  $\mathcal{P}$  is multiplicative in this domain. Hence, for a function,  $f$  bandlimited to  $\mathbb{B}_2$ ,  $\mathcal{P}(f)$  is also bandlimited to  $\mathbb{B}_2$ . Now, from (19) and (21), such a function,  $f$ , can be expressed in terms of the integer translates of  $\phi^s$ :

$$(23) \quad f = \sum_{\mathbf{k} \in \mathbb{Z}^2} f(\mathbf{k}) T_{\mathbf{k}} \phi^s.$$

Since, the same is also true for  $\mathcal{P}(f)$ , we have,

$$\mathcal{P}(f) = \sum_{\mathbf{l} \in \mathbb{Z}^2} \mathcal{P}(f)(\mathbf{l}) T_{\mathbf{l}} \phi^s,$$

where,

$$\mathcal{P}(f)(\mathbf{l}) = \langle \mathcal{P}(f), T_{\mathbf{l}} \phi^a \rangle = \left\langle \mathcal{P} \left( \sum_{\mathbf{k} \in \mathbb{Z}^2} f(\mathbf{k}) T_{\mathbf{k}} \phi^s \right), T_{\mathbf{l}} \phi^a \right\rangle = \sum_{\mathbf{k} \in \mathbb{Z}^2} f(\mathbf{k}) \langle \mathcal{P}(T_{\mathbf{k}} \phi^s), T_{\mathbf{l}} \phi^a \rangle$$

Thus, the vector of samples of the wavefield at  $z + \Delta z$  is given by multiplying the vector of samples of the wavefield at  $z$  with a matrix which we denote by  $\mathcal{P}^{(0)}$ . We refer to this matrix as the propagator matrix for  $V_0$ , given by,

$$\mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(0)} = \langle \mathcal{P} T_{\mathbf{k}} \phi^s, T_{\mathbf{l}} \phi^a \rangle = \left\langle \widehat{\mathcal{P}} e_{\mathbf{k}} \widehat{\phi}^s, e_{\mathbf{l}} \widehat{\phi}^a \right\rangle = \int_{\mathbb{R}^2} \widehat{\phi}^a(\boldsymbol{\xi}) e^{i \Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} e^{-2\pi i (\mathbf{k} - \mathbf{l}, \boldsymbol{\xi})} d\boldsymbol{\xi}.$$



The last equality follows from the fact  $\widehat{\phi}^s = 1$  on the support of  $\widehat{\phi}^a$ . Now, since  $\widehat{\phi}^a$  is supported in  $\mathbb{B}_1 \subset \mathbb{T}^2$ , the  $\mathbf{l}, \mathbf{k}$ -entry of  $\mathcal{P}^{(0)}$  is the  $(\mathbf{k} - \mathbf{l})$ -th Fourier coefficient of  $\widehat{\phi}^a e^{i\Delta z \sqrt{k_0^2 - \|\boldsymbol{\xi}\|^2}}$ . Observe that  $\mathcal{P}^{(0)}$  is a Toeplitz matrix. Hence, the operator is applied on the discretised signal by convolution with the first row of this matrix.

For a fixed value of the ratio  $k_0 = \frac{\omega}{c_0}$ , we refer to these Fourier coefficients as the **propagator filter** corresponding to  $k_0$ . Note that for the constant velocity model, we required only one propagator filter for each  $\omega$ . Now, if the velocity varies laterally, for each  $\omega$ , we need more than one propagator filter to construct the propagator matrix  $\mathcal{P}^{(0)}$ . In that case, we construct a ‘table’ driven migration scheme, meaning that we calculate the propagator filters for a pre-determined set of values of the ratio  $k_0 = \frac{\omega}{c_0}$  and store them in a ‘table’. The values of  $k_0$  range from 0 to  $\frac{\omega_{max}}{c_{min}}$ , where  $\omega_{max}$  is the maximum temporal frequency of the data and  $c_{min}$  is the minimum velocity for the given velocity model.

Recall, for the variable velocity model, we define the propagator weakly via (5):

$$\langle \mathcal{P}f_z, \phi \rangle = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{2\pi i \langle \mathbf{x}, \boldsymbol{\xi} \rangle} e^{2\pi i \Delta z \sqrt{\frac{\omega^2}{c(\mathbf{x})^2} - \|\boldsymbol{\xi}\|^2}} \widehat{f}_z(\boldsymbol{\xi}) \overline{\phi(\mathbf{x})} d\boldsymbol{\xi} d\mathbf{x}.$$

This equation is valid for  $\phi$  compactly supported in the spatial domain and continuous. The equation is still valid if  $\int_{\mathbb{R}^2} |\phi(\mathbf{x})|^2 (1 + |\mathbf{x}|)^{2+\delta} d\mathbf{x} < +\infty$  for  $\delta > 0$ . Hence,

$$\mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(0)} = \langle \mathcal{P}(T_{\mathbf{k}}\phi^s), T_{\mathbf{l}}\phi^a \rangle = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{2\pi i \langle \mathbf{x}, \boldsymbol{\xi} \rangle} e^{2\pi i \Delta z \sqrt{\frac{\omega^2}{c(\mathbf{x})^2} - \|\boldsymbol{\xi}\|^2}} e_{\mathbf{k}}(\boldsymbol{\xi}) \widehat{\phi}^s(\boldsymbol{\xi}) \overline{T_{\mathbf{l}}\phi^a(\mathbf{x})} d\boldsymbol{\xi} d\mathbf{x}$$

Applying the Fubini-Tonelli Theorem, we can change the order of integration. We replace  $c(\mathbf{x})$  by  $c_{\mathbf{l}}$ , the  $\mathbf{l}$ -th sample of the velocity model because  $T_{\mathbf{l}}\phi^a$  is well localized in the space domain. This yields

$$(24) \quad \mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(0)} = \int_{\mathbb{R}^2} e^{2\pi i \Delta z \sqrt{\frac{\omega^2}{c_{\mathbf{l}}^2} - \|\boldsymbol{\xi}\|^2}} e_{\mathbf{k}}(\boldsymbol{\xi}) \widehat{\phi}^s(\boldsymbol{\xi}) \overline{\widehat{\phi}^a(\boldsymbol{\xi})} d\boldsymbol{\xi}.$$

Hence, at a given depth step, the  $\mathbf{l}$ -th row of the matrix  $\mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(0)}$  is the propagator filter corresponding to the ratio  $\frac{\omega}{c_{\mathbf{l}}}$ . Hence, the  $\mathbf{l}$ -th row of the matrix can be obtained from the pre-calculated table by picking up the propagator filter corresponding to  $k_0$  which is closest to the ratio  $\frac{\omega}{c_{\mathbf{l}}}$ .

Notice that  $\widehat{\phi}^a$  acts like a smoothing function to obtain a trigonometric polynomial approximation of the phase shift operator. The matrix,  $\mathcal{P}^{(0)}$ , is practically a spatially varying convolution. This is because for the constant velocity case, we saw it was a Toeplitz matrix (i.e. each row is obtained by right shifting the previous row) and hence, a convolution. Now that the filter changes with every point on the grid (i.e.  $\mathbf{l}$ -th row depends on  $c_{\mathbf{l}}$ ), this induces what is referred to as a spatially varying convolution. As pointed out in the introduction, the discretisation of  $\mathcal{P}$  we obtained here, at the zero resolution level, is what geophysicists call a standard explicit scheme. In the geophysics literature, this kind of analysis is not carried out. Instead, the entries of the propagator matrix are calculated via (24), where instead of  $\widehat{\phi}^a \widehat{\phi}^s = \widehat{\phi}^a$ , an arbitrary smoothing filter is used.

However, the multilevel structure of the IMRA can be utilized to reduce the computational cost as we argue in the following subsections.

**4.2. Decomposing the signal and the propagator matrix using the fast isotropic wavelet algorithm.** We can decompose a wavefield belonging to  $V_0$  into a low pass component and a high pass component using the two dimensional version of the fast wavelet algorithm described in Section 3. Now, the representation of the function becomes

$$f = \underbrace{\sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, D^{-1}T_{\mathbf{k}}\phi^a \rangle D^{-1}T_{\mathbf{k}}\phi^s}_{\text{low pass component}} + \underbrace{\sum_{i=1}^{|\det(A)|} \sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, D^{-1}T_{\mathbf{k}}\psi_i^a \rangle D^{-1}T_{\mathbf{k}}\psi_i^s}_{\text{high pass component}}.$$

Now, from the definition of  $\psi_i^a$  and  $\psi_i^s$  we infer,  $D^{-1}T_{\mathbf{k}'}\psi_i^a = T_{\mathbf{k}}\psi_1^a$  and  $D^{-1}T_{\mathbf{k}'}\psi_i^s = T_{\mathbf{k}}\psi_1^s$  where  $\mathbf{k} = A\mathbf{k}' + \mathbf{p}_i$  and  $\mathbf{p}_i$  are the representatives of the quotient group  $\mathbb{Z}^2/(AZ^2)$ . For instance, if  $A$  is the dyadic dilation matrix,  $2I_{\mathbb{R}^2}$ , then  $\mathbf{p}_1 = (0, 0)$ ,  $\mathbf{p}_2 = (1, 0)$ ,  $\mathbf{p}_3 = (0, 1)$  and,  $\mathbf{p}_4 = (1, 1)$ . We therefore have the following representation of the function  $f$ :

$$f = \underbrace{\sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, T_{A\mathbf{k}}D^{-1}\phi^a \rangle T_{A\mathbf{k}}D^{-1}\phi^s}_{\text{low pass component}} + \underbrace{\sum_{\mathbf{k} \in \mathbb{Z}^2} \langle f, T_{\mathbf{k}}D^{-1}\psi_1^a \rangle T_{\mathbf{k}}D^{-1}\psi_1^s}_{\text{high pass component}}.$$

For this representation, the propagator matrix  $\mathcal{P}^{(0)}$  is decomposed into the following:

$$(25) \quad \begin{pmatrix} \mathcal{P}^{(-1)} & \mathfrak{F}^{(-1)} \\ \mathfrak{T}^{(-1)} & \mathcal{Q}^{(-1)} \end{pmatrix}$$

where  $\mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(-1)} = \langle \mathcal{P}T_{A\mathbf{k}}D^*\phi^s, T_{A\mathbf{l}}D^*\phi^a \rangle$ ,  $\mathfrak{F}_{\mathbf{l}, \mathbf{k}}^{(-1)} = \langle \mathcal{P}T_{A\mathbf{k}}D^*\phi^s, T_{\mathbf{l}}D^*\psi_1^a \rangle$ ,  $\mathfrak{T}^{(-1)} = \langle \mathcal{P}T_{\mathbf{k}}D^*\psi_1^s, T_{A\mathbf{l}}D^*\phi^a \rangle$  and  $\mathcal{Q}^{(-1)} = \langle \mathcal{P}T_{\mathbf{k}}D^*\psi_1^s, T_{\mathbf{l}}D^*\psi_1^a \rangle$ . We argue that we can ignore  $\mathfrak{F}^{(-1)}$  and  $\mathfrak{T}^{(-1)}$  to obtain the following block diagonal propagator matrix,

$$(26) \quad \begin{pmatrix} \mathcal{P}^{(-1)} & 0 \\ 0 & \mathcal{Q}^{(-1)} \end{pmatrix}$$

We now describe the process of obtaining each of the components,  $\mathcal{P}^{(-1)}$ ,  $\mathfrak{F}^{(-1)}$ ,  $\mathfrak{T}^{(-1)}$  and  $\mathcal{Q}^{(-1)}$ , for the constant velocity case. The generalisation to the variable velocity model is exactly the same as for  $\mathcal{P}^{(0)}$ . This is done according to the following equations:

$$\begin{aligned} \mathcal{P}_{\mathbf{l}, \mathbf{k}}^{(-1)} &= \langle \mathcal{P}D^*T_{\mathbf{k}}\phi^s, D^*T_{\mathbf{l}}\phi^a \rangle = |\det(A)| \left\langle \widehat{\mathcal{P}}e_{A\mathbf{k}}\widehat{\phi}^s(A^* \cdot), e_{A\mathbf{l}}\widehat{\phi}^a(A^* \cdot) \right\rangle \\ &= |\det(A)| \int_{\mathbb{R}^2} \widehat{\phi}^a(A^*\boldsymbol{\xi}) e^{i\Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} e^{-2\pi i \langle A(\mathbf{k}-\mathbf{l}), \boldsymbol{\xi} \rangle} d\boldsymbol{\xi} \end{aligned}$$

Similarly, we obtain the following expressions:

$$\begin{aligned} \mathcal{Q}_{\mathbf{l}, \mathbf{k}}^{(-1)} &= \int_{\mathbb{R}^2} \widehat{\psi}_1^a(A^*\boldsymbol{\xi}) e^{i\Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} e^{-2\pi i \langle \mathbf{k}-\mathbf{l}, \boldsymbol{\xi} \rangle} d\boldsymbol{\xi}, \\ \mathfrak{F}_{\mathbf{l}, \mathbf{k}}^{(-1)} &= |\det(A)|^{1/2} \int_{\mathbb{R}^2} \widehat{\phi}^s(A^*\boldsymbol{\xi}) \widehat{\psi}_1^a(A^*\boldsymbol{\xi}) e^{i\Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} e^{-2\pi i \langle A\mathbf{k}-\mathbf{l}, \boldsymbol{\xi} \rangle} d\boldsymbol{\xi}, \\ \mathfrak{T}_{\mathbf{l}, \mathbf{k}}^{(-1)} &= |\det(A)|^{1/2} \int_{\mathbb{R}^2} \widehat{\phi}^a(A^*\boldsymbol{\xi}) \widehat{\psi}_1^s(A^*\boldsymbol{\xi}) e^{i\Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} e^{-2\pi i \langle \mathbf{k}-A\mathbf{l}, \boldsymbol{\xi} \rangle} d\boldsymbol{\xi}. \end{aligned}$$

Now, observe that  $\widehat{\phi^a}(A^*.)$  and  $\widehat{\psi_1^a}(A^*.)$  overlap only in a small annulus with inner radius  $b_2$  and outer radius  $b_1$ . We denote this annulus by  $\mathbb{S}$ . Hence, we can ignore the contribution of the off diagonal components. We give a justification for this claim by looking more closely at  $\mathfrak{F}^{(-1)}$ .

The component  $\mathfrak{F}^{(-1)}$  acts on the sequence  $\{\langle f, T_{A\mathbf{k}} D^* \phi^a \rangle\}_{\mathbf{k} \in \mathbb{Z}^2}$ . Let  $\mathfrak{A}$  be the analysis operator for  $\{T_{A\mathbf{k}} D^* \phi^a\}_{\mathbf{k} \in \mathbb{Z}^2}$ , i.e.  $\mathfrak{A}(f) = \{\langle f, T_{A\mathbf{k}} D^* \phi^a \rangle\}_{\mathbf{k} \in \mathbb{Z}^2}$ . We claim that  $\|\mathfrak{F}^{(-1)} \mathfrak{A}(f)\|$  tends to zero, as the area of the annulus,  $\mathbb{S}$ , goes to zero, for all  $f$  such that  $\widehat{f}$  is essentially bounded. This is a reasonable assumption on  $f$  because we only have finitely many samples of the wavefield. It is therefore represented by a finite linear combination of the translates of the synthesis scaling function which is bounded in the wavenumber domain.

We write  $\mathbf{l} = A\mathbf{l}' + \mathbf{p}_i$ , and define  $|\det(A)|$  sub-matrices of  $\mathfrak{F}^{(-1)}$ , denoted by  $\mathfrak{F}_i^{(-1)}$ , via  $(\mathfrak{F}_i^{(-1)})_{\mathbf{l}', \mathbf{k}} = (\mathfrak{F}^{(-1)})_{\mathbf{l}, \mathbf{k}}$ . Recall that  $\mathbf{p}_i$  are the representatives of the quotient group  $\mathbb{Z}^2 / (A\mathbb{Z}^2)$ . Now, observe that each of  $\mathfrak{F}_i^{(-1)}$  is a Toeplitz matrix with each row made up of the Fourier coefficients of the  $A^{*-1}\mathbb{Z}^2$ -periodic function  $\widehat{\phi^s}(A^*.) \widehat{\psi_i^a}(A^*.) e^{i\Delta z \sqrt{k_0^2 - |\cdot|^2}}$ . Therefore, the action  $\mathfrak{F}_i^{(-1)}$  on  $\mathfrak{A}(f)$  is a convolution. Also note that  $\mathfrak{A}(f) = \widehat{f\phi^a}(A^*.)$ , where we identify the bandlimited function,  $\widehat{f\phi^a}(A^*.)$  with its  $A^{*-1}\mathbb{Z}^2$ -periodic extension. Therefore, applying Plancherel's theorem we conclude,

$$\begin{aligned} \|\mathfrak{F}_i^{(-1)} \mathfrak{A}(f)\|^2 &= \int_{A^{*-1}\mathbb{T}^2} \left| \widehat{\phi^s}(A^*\boldsymbol{\xi}) \widehat{\psi_i^a}(A^*\boldsymbol{\xi}) e^{i\Delta z \sqrt{k_0^2 - |\boldsymbol{\xi}|^2}} \widehat{f}(\boldsymbol{\xi}) \widehat{\phi^a}(A^*\boldsymbol{\xi}) \right|^2 d\boldsymbol{\xi} \\ &\leq \|f\|_\infty^2 \int_{A^{*-1}\mathbb{T}^2} \left| \widehat{\psi_i^a}(A^*\boldsymbol{\xi}) \widehat{\phi^a}(A^*\boldsymbol{\xi}) \right|^2 d\boldsymbol{\xi} \\ &\leq \|f\|_\infty^2 \|\widehat{\psi_i^a}(A^*.) \widehat{\phi^a}(A^*.)\|_\infty^2 |\mathbb{S}| \end{aligned}$$

where  $|\mathbb{S}|$  is the area of the annulus  $\mathbb{S}$ . This proves our claim that  $\|\mathfrak{F}^{(-1)} \mathfrak{A}(f)\|$  tends to zero as the area of  $\mathbb{S}$  tends to zero. Similar calculations can be carried out for the component,  $\mathfrak{F}^{(-1)}$ . Hence, by keeping the area of the annulus small, we can neglect the error caused by ignoring the off-diagonal components of the propagator. Our experimental results support this claim (see Figure 1).

**4.3. Increasing the sparsity of the propagator matrix.** In this subsection, we describe how the decomposition into low and high pass components can be used to increase the sparsity of the propagator matrix and thereby reduce the cost of computation. We use  $\sigma$  to denote the signal in  $V_0$ , i.e. the samples of the wavefield on  $\mathbb{Z}^2$  and refer to this sequence of samples as the original signal and the grid on which it is defined, as the original grid. After decomposing into a low pass component,  $\sigma_l$  and a high pass component,  $\sigma_h$ , using the fast wavelet algorithm, the wavefield is represented by the column vector  $(\sigma_l, \sigma_h)^T$ , where  $T$  stands for transpose. Recall that the high pass component is undecimated. Hence,  $\sigma_h$  has the same size as  $\sigma$ . Thus, as a result of our decomposition, we increase the amount of data by the size of  $\sigma_l$ . The low pass component,  $\sigma_l$  is a quarter of the size of  $\sigma$  if we use the Dyadic dilation matrix. Applying the propagator (26) on  $(\sigma_l, \sigma_h)^T$  yields  $(\mathcal{P}^{(-1)} \sigma_l, \mathcal{Q}^{(-1)} \sigma_h)^T$ . The cost of matrix-vector multiplication,  $\mathcal{Q}^{(-1)} \sigma_h$  is the same as that of the matrix-vector multiplication  $\mathcal{P}^{(0)} \sigma$ .

Nevertheless, the sparsity increases from the fact that we can discard the high pass component for certain values of the ratio,  $\omega/c$ . Wavenumbers larger than

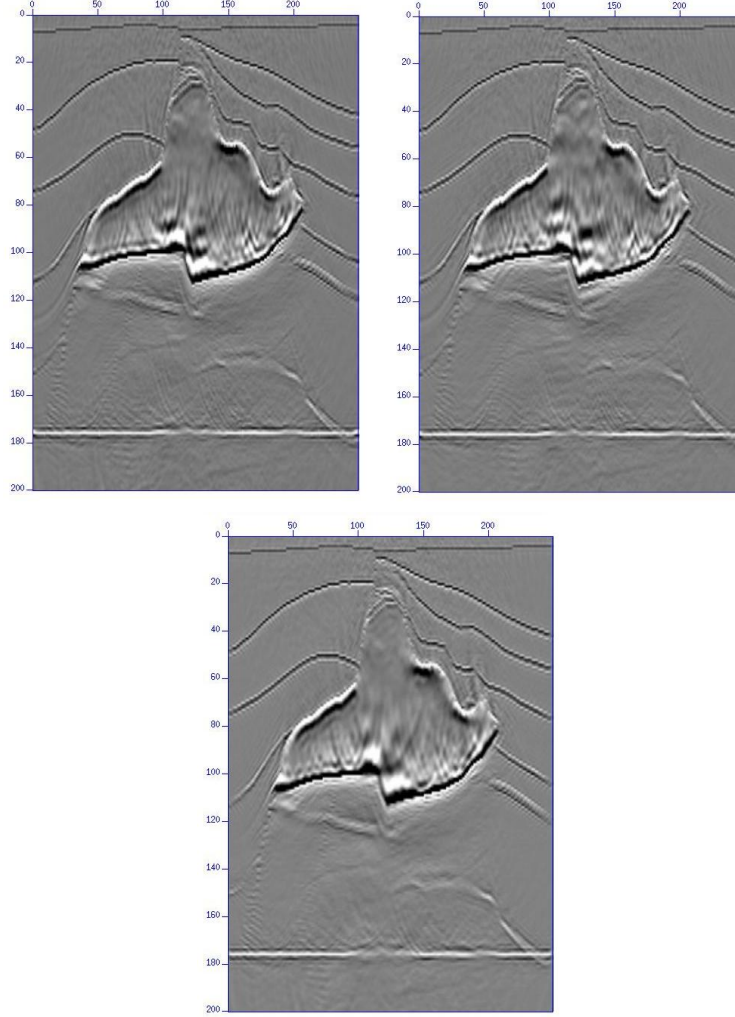


FIGURE 1. Here we show the vertical cross sections of the images obtained for the SEG Salt model using IMRA. The left image on the top was obtained by discretising the propagator at the zero resolution level. The right image on the top was obtained using one-level of decomposition with the Dyadic dilation matrix. The image on the bottom was obtained using one-level of decomposition using the Quincunx dilation matrix.

the ratio  $\omega/c$  correspond to evanescent waves which are considered non-physical [5]. Hence, we only need to propagate the waves corresponding to wavenumbers smaller than  $\omega/c$  accurately, while the evanescent waves can be damped or set to zero. Now, for ratios  $\omega/c < b_2$ , we can set the high pass component equal to zero because in this case the high pass component contains only the evanescent waves. By setting the high pass component to zero, we mean that we can set the  $l$ -th row of  $Q^{(-1)}$  equal to zero whenever  $\omega/c_l < b_2$ . In the case of dyadic dilations,

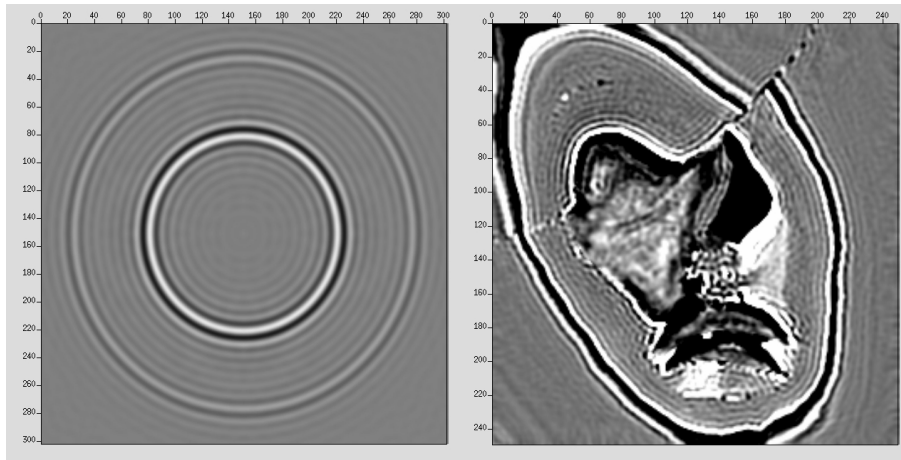


FIGURE 2. Horizontal cross-sections of the results obtained using the IMRA for the Impulse (left) and the Salt (right) models. Due to the radial nature of the filters the faults in all directions are imaged with the same accuracy and have no directional artifacts.

$\frac{1}{8} < b_2 < \frac{1}{4}$ . This reduces the computational load significantly for regions where the velocity is high and the temporal frequency is low, because in such regions we only propagate the low pass component and with the use of dyadic dilations the low pass component is one-quarter the size of the original grid. This use of the IMRA can be compared to the sub-sampling scheme proposed by Margrave et al [11]. However, unlike Margrave’s sub-sampling, the 1-level IMRA wavelet decomposition is suitable for localized variations of the velocity model. In their implementation, for a fixed frequency,  $\omega$ , the signal is sub-sampled to reduce the computation time. The IMRA decomposition is more general because we subsample not only for one fixed frequency, but also for various regions of the image where the velocity is high.

A significant gain in speed as compared to standard explicit scheme can be anticipated if we can discard the high pass component for a sufficient number points in the grid. However, Table 1 shows that with dyadic dilation, we have to propagate the high pass component for almost every grid point when the frequency is higher than 9 Hz. When this happens a decomposition ends up being more expensive. Hence, it is useful to decompose the wavefield only for very low frequencies.

To rectify this, we want the low pass filter to be supported in a ball that has bigger radius than the one we had with Dyadic dilation, so that the high pass component can be discarded for more values of  $\omega/c$ . This can be achieved by replacing the Dyadic dilation matrix with the Quincunx matrix:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

This allows the low pass band to be supported in a ball of radius  $b_0$ , where  $\frac{1}{4} < b_0 < \frac{\sqrt{2}}{4}$ . But, in this case the discretisation grid for the low pass component is half the size of the original computation grid. As we can see from Table 1, this

improves the situation because now, decomposing the wavefield is useful for higher temporal frequencies than in the case of Dyadic dilation.

Frequency in Hz	Dyadic	Quincunx
1.5 - 8.75	55	0
9 - 16.25	99.5	37.5
16.5 - 23.75	100	93.3
24 - 31.25	100	96.4
31.25 - 37.5	100	100

TABLE 1. Percentage of points in the high pass component that must be propagated for various ranges of temporal frequencies, in the case of Dyadic and Quincunx dilations, for the SEG Salt model.

## 5. Filter design and implementation

To implement either the explicit scheme or the IMRA approach, we need a trigonometric polynomial approximation of the phase shift operator. We need short filters (i.e. less terms in the trigonometric polynomial) to keep the computational cost under control. Since the extrapolation is done for small increments  $\Delta z$ , these filters must not amplify the wavefield at every depth step. Hence, we have to optimize the filter under a constraint that guarantees stability and in a way that keeps the computational cost under control. We use a weighted least squares (WLSQ) algorithm due to Thorbecke et al [15], to carry out this task. Recall that wavenumbers larger than the ratio  $\omega/c$  correspond to evanescent waves which are considered non-physical. The only requirement for the evanescent waves is that they should be damped (i.e. must be less than one in absolute value) so that they do not cause numerical instability. Hence, we use a weight that is equal to one in the propagating region ( $|\xi| < \omega/c$ ) and equal to a very small value (about  $10^{-5}$ ) in the evanescent region ( $|\xi| > \omega/c$ ). The filters calculated with this weighted least squares approach have larger errors in the evanescent region but are very accurate for  $|\xi| < \omega/c$ . Whereas, filters calculated without the weight (i.e. by truncating the Fourier series) have oscillations in the propagating region which results in numerical instability. This is illustrated in Figure 3.

For 3-D migration, we have to convolve with 2-D filters. Two-dimensional convolutions are expensive and hence, we study ways to make these convolutions computationally efficient without compromising accuracy. We can take advantage of the radial nature of the filters to optimize the 2-D convolution. Since a radial filter is even in both variables, for each point we can first sum up the four quarters of the data and perform the convolution for 1/4 of the original size of the filter i.e. just the first quadrant. Another approach is to design a 1-D filter in the radial variable and use a transformation to map it to a 2-D filter. We considered two such transformations, the McClellan transform (see [12, 9]) and the Soubaras' Laplacian synthesis (see [14]).

## 6. Acknowledgments

We authors thank Professor Bernhard G. Bodmann, Professor Donald J. Kouri and our collaborators from Total E&P, Dr. Henri Calandra, Dr. Bertrand Denel, and

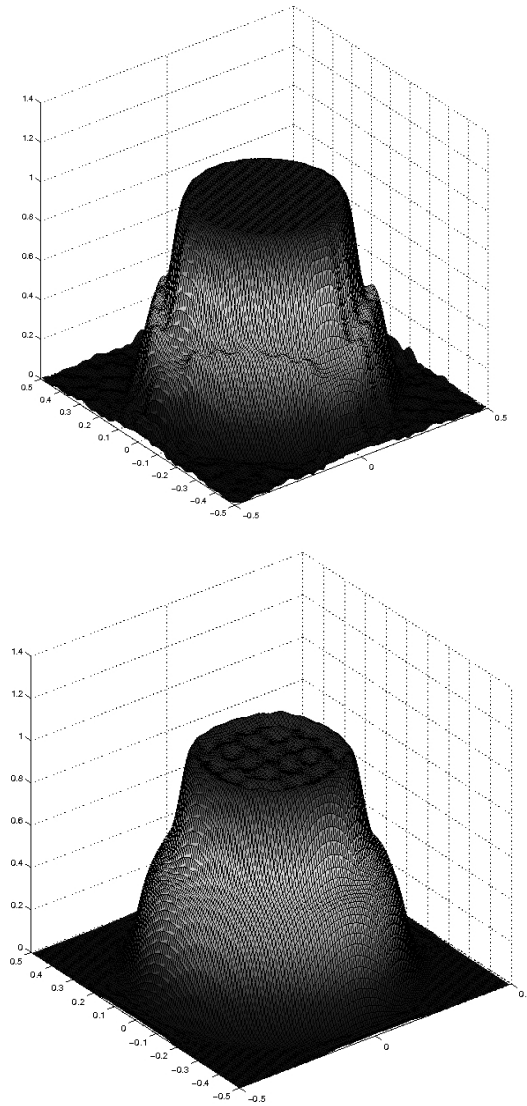


FIGURE 3. Trigonometric polynomial approximations of the propagator filter with a weighted least squares algorithm (top) and with a truncated Fourier series (bottom). Both have  $25 \times 25$  coefficients.

Dr. Paul Williamson for many helpful discussions. We also thank the anonymous referee for carefully working on the manuscript and for the useful suggestions that helped us to improve the readability of the article.

### References

- [1] S. Alexander, S. Baid, S. Jain, Manos Papadakis, and J. Romero. Isotropic Multiresolution Analysis: Geometry and analytic properties. In progress.

- [2] B.G. Bodmann and M. Papadakis et. al. Frame isotropic multiresolution analysis for micro CT scans of coronary arteries. In M. Papadakis, A. Laine, and M. Unser, editors, *Wavelets XI*, volume 5914, pages 5914I0/1–12. SPIE, 2005.
- [3] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhauser, 2002.
- [4] C. K. Chui, W. He, and J. Stockler. Compactly supported tight and sibling frames with maximum vanishing moments. *Appl. Comput. Harmon. Anal.*, 13:224–262, 2002.
- [5] Claerbout. *Imaging the Earth's Interior*. <http://sep.stanford.edu/oldreports/sep40/>.
- [6] J.B. Conway. *A Course in Operator Theory*, volume 21 of *Graduate Studies in Mathematics*. American Mathematical Society, 1999.
- [7] I. Daubechies, B. Han, A. Ron, and Z. Shen. Framelets: MRA-based constructions of wavelet frames. *Appl. Comput Harmon. Anal.*, 14(1):1–46, 2003.
- [8] G. Blacquiere et. al. 3D table-driven migration. *Geophys. Prosp.*, 37:925–958, 1989.
- [9] D. Hale. 3-D migration via McClellan transformations. *Geophysics*, 56(11):1778–1785, 1991.
- [10] O. Holberg. Towards optimum one-way wave propagation. *Geophys. Prosp.*, 36:99–114, 1988.
- [11] G.F. Margrave and et al. Improving explicit seismic depth migration with a stabilizing Wiener filter and spatial resampling. *Geophysics*, 71(3):S111–S120, 2006.
- [12] J.H. McClellan. On the design of one-dimensional and two-dimensional FIR digital filters. *PhD Thesis, Rice University*, 1973.
- [13] Ristow and Ruhl. Fourier finite-difference migration. *Geophysics*, 59(12):1882–1893, 1994.
- [14] R. Soubaras. Explicit 3-D migration using equiripple polynomial expansion and Laplacian synthesis. *Geophysics*, 61(5):1386–1393, 1996.
- [15] J.W. Thorbecke, K. Wapenaar, and G. Swinnen. Design of one-way wavefield extrapolation operators, using smooth functions in WLSQ optimization. *Geophysics*, 69(4):1037–1045, 2006.