Higher-dimensional wavelets and the Douglas-Rachford algorithm

David Franklin
School of Math and Phys Sci
University of Newcastle
Callaghan NSW 2308
Australia
david.franklin@newcastle.edu.au

Jeffrey A. Hogan
School of Math and Phys Sci
University of Newcastle
Callaghan NSW 2308
Australia
jeff.hogan@newcastle.edu.au

Matthew Tam
Institute for Num and Appl Math
University of Göttingen
Lotzestr. 16-18, 37083 Göttingen
Germany
m.tam@math.uni-goettingen.de

***Abstract*—We recast the problem of multiresolution-based wavelet construction in one and higher dimensions as a feasibility problem with constraints which enforce desirable properties such as compact support, smoothness and orthogonality of integer shifts. By employing the Douglas-Rachford algorithm to solve this feasibility problem, we generate one-dimensional and non-separable two-dimensional multiresolution scaling functions and wavelets.**

I. Multiresolution analysis, scaling functions, wavelets and Douglas-Rachford

The construction of a compactly supported smooth orthogonal scaling function–wavelet pair, (φ, ψ), on the line was first achieved by Daubechies in [8] with the help of the multiresolution structure introduced independently by Mallat [18] and Meyer [19]. The problem reduces to the construction of a matrix-valued function $U : \mathbb{R} \to \mathbb{R}^{2\times 2}$ satisfying certain restrictions designed to force φ and ψ to have desirable properties for signal processing. Construction of *nonseparable* multiresolution wavelets in higher dimensions has proved more elusive, although there are some constructions which rely on lifting one-dimensional constructions [2], [4], [11], [13]–[16], [22]. In this paper, we describe work done by Franklin in his PhD thesis [10] on the application of optimisation techniques to the construction of compactly supported smooth orthogonal multiresolution wavelets in one- and two-dimensions.

In what follows, the collection of $n \times n$ matrices with complex entries is denoted $\mathbb{C}^{n\times n}$ and the collection of unitary $n \times n$ matrices is denoted $\mathcal{U}(n)$. If $x \in \mathbb{R}^n$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$ with each α_i a non-negative integer, we denote $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$. We also let $Q_M^n = \{0, 1, \ldots, M-1\}^n \subset \mathbb{Z}^n$.

A. Multiresolution analysis and wavelet matrices

Definition 1: A *multiresolution analysis* (MRA) $(\{V_j\}_{j=-\infty}^{\infty}, \varphi)$ for $L^2(\mathbb{R}^n)$ is a sequence of closed subspaces $\{V_j\}_{j=-\infty}^{\infty} \subset L^2(\mathbb{R}^n)$ and a function $\varphi \in V_0$ such that

(i) $V_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$
(ii) $\cap_{j=-\infty}^{\infty} V_j = \{0\}$ and $\overline{\cup_{j=-\infty}^{\infty} V_j} = L^2(\mathbb{R})$
(iii) $f(x) \in V_j \iff f(2x) \in V_{j+1}$
(iv) $f(x) \in V_0 \iff f(x-k) \in V_0 \ (k \in \mathbb{Z}^n)$.
(v) $\{\varphi(x-k)\}_{k=\infty}^{\infty}$ is an orthonormal basis for V_0.

Given an MRA structure, it can be shown that there exists an ℓ^2 sequence $\{g_k^0\}_{k\in\mathbb{Z}^n}$ such that, for $m_0(\xi) = \sum_{k\in\mathbb{Z}^n} g_k^0 e^{-2\pi i\langle k,\xi\rangle}$ $(\xi \in \mathbb{R})$, we have

$$\hat{\varphi}(2\xi) = m_0(\xi)\hat{\varphi}(\xi). \tag{1}$$

A necessary (but not sufficient) condition for orthonormality of the collection $\{\varphi(x-k)\}_{k\in\mathbb{Z}^n}$ is the *quadrature mirror filter* (QMF) condition $\sum_{p\in V^n} |m_0(\xi + p/2)|^2 = 1$ for almost every ξ, where V^n is the collection of the 2^n vertices of the unit cube $[0,1]^n \subset \mathbb{R}^n$.

Let $W_0 = V_1 \ominus V_0$. Then there are $2^n - 1$ *wavelets* $\{\psi^\varepsilon\}_{\varepsilon=1}^{2^n-1}$ such that the collection $\{\psi^\varepsilon(x-k);\ k \in \mathbb{Z},\ 1 \le \varepsilon \le 2^n - 1\}$ is an orthonormal basis for W_0. In this case, the collection $\{2^{nj/2}\psi^\varepsilon(2^j x - k);\ k \in \mathbb{Z}^n,\ 1 \le \varepsilon \le 2^n - 1\}$ is an orthonormal *wavelet basis* for $L^2(\mathbb{R}^n)$. The challenge is finding such functions φ and ψ^ε $(1 \le \varepsilon \le 2^n - 1)$.

Again because of the MRA structure, there are ℓ^2 sequences $\{g_k^\varepsilon\}_{k\in\mathbb{Z}^n}$ such that with $m_\varepsilon(\xi) = \sum_{k\in\mathbb{Z}^n} g_k^\varepsilon e^{2\pi i\langle k,\xi\rangle}$ we have

$$\hat{\psi^\varepsilon}(2\xi) = m_\varepsilon(\xi)\hat{\varphi}(\xi). \tag{2}$$

Define the mapping $U : \mathbb{R}^n \to \mathbb{C}^{2^n\times 2^n}$ by

$$U(\xi)_{\varepsilon p} = (m_\varepsilon(\xi + \frac{p}{2}))_{0\le\varepsilon\le 2^{n-1},\ p\in V^n}. \tag{3}$$

If $\{\psi^\varepsilon(x-k);\ k \in \mathbb{Z}^n,\ 1 \le \varepsilon \le 2^n - 1\}$ is an orthonormal collection in W_0, then $U(\xi)$ must be unitary for all ξ. This condition is not sufficient to ensure orthonormality. Nevertheless, once such a function U is given, it is possible to check that the wavelets generated by it form an orthonormal basis for $L^2(\mathbb{R}^n)$ (see [6] for details).

Beyond unitarity of $U(\xi)$, we also need to impose further conditions. To ensure *completeness* of the wavelet basis, it is enough to ensure the multiresolution property $\overline{\cup_{j=-\infty}^{\infty} V_j} = L^2(\mathbb{R}^n)$. We shall also seek conditions to ensure that the scaling function and associated wavelets are themselves smooth and *compactly supported*, and that the length of the support is a quantity to be minimised wherever possible.

By the multidimensional Paley-Wiener theorem [23], compact support of the scaling function and associated wavelets is equivalent to the filters $\{m_\varepsilon\}_{\varepsilon=0}^{2^n-1}$ being trigonometric polynomials. We will seek filters whose coefficients g_k^ε are zero unless $0 \le k_i \le M-1$ for all $0 \le \varepsilon \le 2^n - 1$ and some fixed positive integer M.

The completeness and smoothness conditions can be imposed using conditions 3 and 4 in Problem 1 below.

Problem 1: Find a matrix-valued function $U(\xi)$ of the form (3) such that

1. Each function m_ε $(0 \le \varepsilon \le 2^n - 1)$ is of the form $m_\varepsilon(\xi) = \sum_{k \in Q_M^n} g_k^\varepsilon e^{2\pi i \langle k, \xi \rangle}$;
2. $U(\xi)$ is unitary for a.e. ξ;
3. $U(0) = \begin{pmatrix} 1 & 0^T \\ 0 & V \end{pmatrix}$ with $V \in \mathcal{U}(2^n - 1)$ and $0^T = (0, 0, \dots, 0) \in \mathbb{C}^{2^n - 1}$;
4. $\partial^\alpha U(0) = \begin{pmatrix} a_\alpha & 0^T \\ 0 & A_\alpha \end{pmatrix}$ $(|\alpha| \le d)$ with $A_\alpha \in \mathbb{C}^{(2^n-1)\times(2^n-1)}$ and $a_\alpha \in \mathbb{C}$.

B. Optimisation Preliminaries

Projection operators: Let $\mathcal{E}$ be a finite-dimensional Hilbert space. If $S \subseteq \mathcal{E}$, its *(metric) projector* is the point-to-set mapping given by

$$P_S(x) := \{s \in S : \|s - x\| \le d(x, S)\},$$

where $d(x, S) = \inf_{s \in S} \|x - s\|$. It is straightforward to check that $P_S(x) \neq \emptyset$ for all $x \in \mathcal{E}$ so long as S is nonempty and closed. We write $P_S(x) = p$ to mean $P_S(x) = \{p\}$.

Proposition 2 (Properties of projectors): Let $\mathcal{E}$ be a finite dimensional Hilbert space.

1) Let $C_0, C_1, \dots, C_{M-1} \subseteq \mathcal{E}$ be nonempty closed sets and define $C := C_0 \times \cdots \times C_{M-1} \subseteq \mathcal{E}^M$. Then
$$P_C = P_{C_0} \times \cdots \times P_{C_{M-1}}.$$
2) Let $L : \mathcal{E} \to \mathcal{E}$ be an isometric isometry and $C \subseteq \mathcal{E}$ be a nonempty closed set. Then
$$P_{L(C)} = L \circ P_C \circ L^{-1}.$$

In what follows, the unit sphere is denoted $\mathbb{S} := \{x \in \mathcal{E} : \|x\| = 1\}$. The singular value decomposition (SVD) of a matrix $A \in \mathbb{C}^{n\times n}$ is $A = USV^*$ where $U, V \in \mathcal{U}(n)$ and $S \in \mathbb{C}^{n \times n}$ is a diagonal matrix with the diagonal entries (the *singular values* of A) being the eigenvalues of $\sqrt{A^*A}$.

Proposition 3 (Examples of projectors): Let $\mathcal{E}, \mathcal{E}'$ be finite dimensional Hilbert spaces.

1) Let $L : \mathcal{E} \to \mathcal{E}'$ be linear and denote $C := \{x \in \mathcal{E} : Lx = 0\}$. If LL^* is invertible, then
$$P_C(x) = x - L^*(LL^*)^{-1}(Lx) \quad \forall x \in \mathcal{E}.$$
2) Let $x \in \mathcal{E}$. Then $P_{\mathbb{S}}(x) = \begin{cases} \frac{x}{\|x\|} & x \neq 0, \\ \mathbb{S} & x = 0. \end{cases}$
3) Let $X \in \mathbb{C}^{n \times n}$. Then $P_{\mathcal{U}(n)}(X) = \{UV^* : X = USV^* \text{ is an SVD}\}$.

Projection Algorithms and Feasibility Problems: Given finitely many closed sets $C_1, \dots, C_n \subseteq \mathcal{E}$ with nonempty intersection, the corresponding *feasibility problem* is

$$\text{find } x \in \bigcap_{k=1}^n C_k. \tag{4}$$

Projection algorithms are a family of iterative algorithms which can be used to solve (4) by, in each step, utilising only projectors onto the individual sets (rather than the entire intersection at once). The two most important examples of projection algorithms are the *method of cyclic projections* [5] and the *Douglas–Rachford (DR) method* [3], [17].

In this work, we employ the DR method which is the following fixed point iteration: Given closed sets $C, D \subset \mathcal{E}$ and $x_0 \in \mathcal{E}$, choose any sequence (x_k) satisfying

$$x_{k+1} \in T(x_k) \text{ where } T := \frac{I + R_C R_D}{2}, \tag{5}$$

and $R_A := 2P_A - I$ denotes the *reflector* with respect to a set A. Here we note that sequence (x_k) is only required to satisfy the inclusion since, in general, the operator $T : \mathcal{E} \to 2^{\mathcal{E}}$ is a point-to-set mapping.

When applying a method based on (5), the sequence of interest (i.e., the one that solves (4)) is not (x_k) itself, but one of its projections onto the set A. In order to be concrete, we state a general convergence result for the convex setting in Theorem 4.

Although the DR algorithm as described above only directly applies to (4) with $n = 2$, the general problem (4) can always be cast as a two set problem. To do so, we define the following two subsets of $\mathcal{E}^n$:

$$C := C_1 \times C_2 \times \cdots \times C_n, \ D := \{(x, x, \dots, x) \in \mathcal{E}^n : x \in \mathcal{E}\}.$$

Then the following equivalence holds

$$x \in \bigcap_{k=1}^n C_k \iff (x, x, \dots, x) \in C \cap D.$$

From here onwards, when speaking of applying the DR algorithm to a feasibility problem, we will mean its product space reformulation.

Theorem 4 (Behaviour of the DR algorithm [3, Theorem 3.13]): Suppose $C, D \subseteq \mathcal{E}$ are closed and convex with nonempty intersection. Let $x_0 \in \mathcal{E}$ and set $x_{k+1} = T(x_k)$ $(k \in \mathbb{N})$. Then the sequence (x_k) converges to a point $x \in \operatorname{Fix} T := \{x : Tx = x\}$ and, moreover, $P_D(x) \in C \cap D$.

Beyond the case of convex sets, there is insufficient theory to fully justify application of projection methods. Indeed, most non-convex results in the literature rely on restrictive regularity notions and yield only local convergence guarantees [7], [12], [21]. Nevertheless, projection methods have been empirically observed to still perform well in certain non-convex settings including matrix completion [1]. This experience suggests use of the DR method in the setting outlined in following section.

II. Wavelets on the line

Here the wavelet matrix $U = U(\xi)$ takes the form

$$U(\xi) = \begin{pmatrix} m_0(\xi) & m_1(\xi) \\ m_0(\xi + 1/2) & m_1(\xi + 1/2) \end{pmatrix}. \tag{6}$$

For compact support, we insist that m_0 and m_1 be trigonometric polynomials of the form $m_0(\xi) = \sum_{k=0}^{M-1} h_k e^{2\pi i k \xi}$, $m_1(\xi) = \sum_{k=0}^{M-1} g_k e^{2\pi i k \xi}$ from which we see that $U(\xi) =$

$\sum_{k=0}^{M-1} A_k e^{2\pi i k\xi}$ with each $A_k \in \mathbb{C}^{2\times 2}$. This allows for a discretisation of the problem. Let $U_j = U(j/M)$ $(j \in Q_M^1)$, i.e., $U_j = \sum_{k=0}^{M-1} A_k e^{2\pi i jk/M}$. The sampling procedure produces an *ensemble* of matrices $\mathcal{U} = (U_0, U_1, \ldots, U_{M-1}) \in (\mathbb{C}^{2\times 2})^M$. The coefficient matrices A_k may be obtained from the sample matrices U_j by a discrete Fourier transform:

$$A_k = (\mathcal{F}_M \mathcal{A})_k = \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2\pi i jk/M}, \tag{7}$$

with inverse $U_j = (\mathcal{F}_M^{-1}\mathcal{A})_j$. From this we see that properties of $U(\xi)$, which are encoded in the coefficient matrices A_k, are also encoded in the sample matrices U_j. The sampling procedure imposes some structure on the ensembles. When U is defined as in (6), then $U(\xi + 1/2) = \sigma U(\xi)$ where $\sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the "row swap" matrix. The samples U_j must reflect this symmetry. In particular, when $M \geq 4$ is even, the ensemble $\mathcal{U}$ of samples must satisfy $U_{j+M/2} = \sigma U_j$ $(j \in Q_{M/2}^1)$. The collection of enembles $\mathcal{V} \in (\mathbb{C}^{2\times 2})^M$ with this symmetry property is denoted $(\mathbb{C}^{2\times 2})_\sigma^M$.

Unitarity of each sample $U_j = U(j/M)$ $(j \in Q_M^1)$ is insufficient to ensure unitarity of $U(\xi)$ for all ξ. However, it transpires that unitarity of the $2M$ samples $U(\frac{j}{2M})$ $(j \in Q_{2M}^1))$ is sufficient. These matrices may be obtained from $\mathcal{U}$ as follows. Let $(\chi_M)_j = e^{\pi i j/M}$ $(j \in Q_M^1)$, $\mathcal{U}$ be as above and

$$\tilde{\mathcal{U}} = \left(U\left(\frac{1}{2M}\right), U\left(\frac{3}{2M}\right), \ldots, U\left(1 - \frac{1}{2M}\right)\right).$$

i.e., $(\tilde{\mathcal{U}})_\ell = U(\frac{2\ell+1}{2M})$ $(\ell \in Q_M^1)$. Then $\tilde{\mathcal{U}} = \mathcal{F}_M^{-1}\chi_M \mathcal{F}_M(\mathcal{U})$.

Finally we note that the regularity condition 4 from Problem 1 may be written in terms of the sample matrices U_j:

$$\sum_{j=0}^{M-1} j^\ell A_j = \frac{1}{M} \sum_{k=0}^{M-1} \alpha_{\ell k} U_k$$

where $\alpha_{\ell k} = \dfrac{1}{M} \sum_{j=0}^{M-1} j^\ell e^{-2\pi i kj/M}$.

Problem 1 can now be viewed as the following three-set feasibility problem (posed in the subspace $(\mathbb{C}^{2\times 2})_\sigma^M$):

Problem 2: Given an even integer $M \geq 4$, find $\mathcal{U} = (U_0, \ldots, U_{M-1}) \in C_1 \cap C_2 \cap C_3 \subseteq (\mathbb{C}^{2\times 2})_\sigma^M$ where

$$C_1 := \left\{ \mathcal{U} : U_0 = \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix}, |z| = 1, U_j \in \mathcal{U}(2), j \in Q_{M/2}^1 \right\}$$

$$C_2 := \left\{ \mathcal{U} : (\mathcal{F}_M \chi_M (\mathcal{F}_M)^{-1}(\mathcal{U}))_j \in \mathcal{U}(2), j \in Q_{M/2}^1 \right\},$$

$$C_3 := \left\{ \mathcal{U} : \sum_{j=0}^{M-1} \alpha_{\ell k} U_k \in \operatorname{diag}(\mathbb{C}^{2\times 2}),\ 1 \leq \ell \leq d \right\}.$$

Given an arbitrary starting point (i.e, an ensemble $\mathcal{U}^0 \in (\mathbb{C}^{2\times 2})_\sigma^M$), we apply the DR algorithm in the Hilbert space $\mathcal{E} = (\mathbb{C}^{2\times 2})_\sigma^M$ to find an ensemble $\mathcal{U} \in \cap_{j=1}^3 C_j$. Then (7) is used to compute the coefficient matrices A_k, whose entries contain the scaling and wavelet dilation equation coefficients h_k and g_k. The cascade algorithm [9] is used to determine the values of the scaling function φ at dyadic rationals and the Fourier transform of equation (2) is used to compute the values of the wavelet ψ. Despite the constraints C_1 and C_2 being non-convex, the DR algorithm converges for a high proportion of random starting ensembles, producing smooth compactly supported scaling functions and wavelets with orthogonal shifts. In particular, for each even integer M, the algorithm successfully reproduces the Daubechies systems of that order, provided the appropriate number of derivative constraints are applied in C_3. The algorithm has also provided hitherto unseen examples, a symmetric example of which appears in Figure 1.

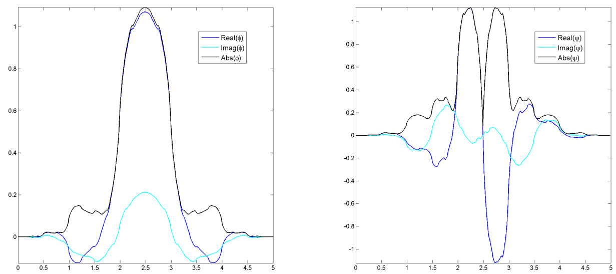

Fig. 1. Plots of the scaling function φ and associated wavelet ψ on their support $[0, 5]$, discovered by the DR method (with $M = 6$). The real component of the respective functions is denoted in blue, the imaginary component in light blue, and the magnitude in black.

III. Wavelets on the Plane

In two dimensions, the wavelet matrix $U(\xi)$ $(\xi \in \mathbb{R}^2)$ takes the form

$$U(\xi)^T = \begin{pmatrix} m_0(\xi) & m_0(\xi+q_1) & m_0(\xi+q_2) & m_0(\xi+q_3) \\ m_1(\xi) & m_1(\xi+q_1) & m_1(\xi+q_2) & m_1(\xi+q_3) \\ m_2(\xi) & m_2(\xi+q_1) & m_2(\xi+q_2) & m_2(\xi+q_3) \\ m_3(\xi) & m_3(\xi+q_1) & m_3(\xi+q_2) & m_3(\xi+q_3) \end{pmatrix}$$

where $q_1 = (\frac{1}{2}, 0)$, $q_2 = (0, \frac{1}{2})$ and $q_3 = (\frac{1}{2}, \frac{1}{2})$ and to ensure compact support, we require that each m_ε $(0 \leq \varepsilon \leq 3)$ be a trigonometric polynomial, i.e., for some even integer $M \geq 4$, we have $m_\varepsilon(\xi) = \sum_{k \in Q_M^2} g_k^\varepsilon e^{2\pi i \langle k, \xi\rangle}$. $U(\xi)$ may therefore be written as $U(\xi) = \sum_{k \in Q_M^2} A_k e^{2\pi i \langle k, \xi \rangle}$ with each $A_k \in \mathbb{C}^{4\times 4}$. We sample $U(\xi)$ at the points $\xi = j/M$ with $j \in Q_M^2$ to obtain matrices $U_j = U(j/M)$ and an ensemble $\mathcal{U} = (U_j)_{j \in Q_M^2} \in (\mathbb{C}^{4\times 4})^{M\times M}$. The coefficient matrices and sample ensemble are related through the (two-dimensional) discrete Fourier transform:

$$A_k = \frac{1}{M^2} \sum_{j \in Q_M^2} U_j e^{-2\pi i \langle j, k\rangle/M}.$$

From the definition of $U(\xi)$ above we see that U must satisfy the *consistency conditions* $U(\xi + q_1) = \sigma_1 U(\xi)$, $U(\xi + q_2) = \sigma_2 U(\xi)$ where σ_1 and σ_2 are the row swap matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

The samples U_j must also satisfy these conditions in the sense that $U_{j+Mq_1} = \sigma_1 U_j$ and $U_{j+Mq_2} = \sigma_2 U_j$ for $j \in Q^2_{M/2}$. We therefore restrict attention to the subspace $(\mathbb{C}^{4\times 4})^{M\times M}_{\sigma_1,\sigma_2}$ of $(\mathbb{C}^{4\times 4})^{M\times M}$ consisting of ensembles that satisfy the consistency conditions. By $1 \otimes U(3)$ we mean the collection of 4×4 matrices of the form $\begin{pmatrix} 1 & 0^T \\ 0 & V \end{pmatrix}$ where $V \in \mathcal{U}(3)$. By $\mathbb{C} \otimes \mathbb{C}^{3\times 3}$ we mean the collection of 4×4 matrices of the form $\begin{pmatrix} b & 0^T \\ 0 & B \end{pmatrix}$ where $b \in \mathbb{C}$ and $B \in \mathbb{C}^{3\times 3}$. Wavelet construction is then reduced to solving the following feasibility problem.

Problem 3: Let $\mathcal{E} = (\mathbb{C}^{4\times 4})^{M\times M}_{\sigma_1,\sigma_2}$. Find an ensemble $\mathcal{U} \in \cap^3_{j=1} C_j$ where

$$C_1 = \{\mathcal{U} \in \mathcal{E} : U_0 \in 1 \otimes \mathcal{U}(3),\, U_j \in \mathcal{U}(4),\, j \in Q^2_{M/2}\}$$
$$C_2 = \{\mathcal{U} \in \mathcal{E} : (\mathcal{F}_M{}^{-1} \chi_\ell \mathcal{F}_M \mathcal{U})_j \in \mathcal{U}(4)\ 1 \leq \ell \leq 3,\ j \in Q^2_{M/2}\}$$
$$C_3 = \{\mathcal{U} \in \mathcal{E} : \sum_{j \in Q^2_M} a_{\alpha,j} U_j \in \mathbb{C} \otimes \mathbb{C}^{3\times 3},\ \ |\alpha| \leq d\}.$$

Here $(\chi_\ell)_j = e^{\pi i \langle p_\ell, j\rangle/M}$, $a_{\alpha,j} = \sum_{k\in Q^2_M} k^\alpha e^{-2\pi i\langle j,k\rangle/M}$ and $\{p_\ell\}^3_{\ell=0} = V^2$.

The DR algorithm is used to solve Problem 3, seeded by a starting ensemble $\mathcal{U} \in (\mathbb{C}^{4\times 4})^{M\times M}$. The ensemble found in the intersection of the three constraint sets given must pass *Cohen's criterion* for orthogonality [6] and a test for *non-separability*. As in the one-dimensional case, the output is sufficient to determine a scaling function and associated wavelets. The output of a typical run of the algorithm is given in Figure 2, which shows a smooth non-separable orthogonal MRA scaling function φ and three associated wavelets ψ_1, ψ_2, ψ_3 all supported on $[0, 5]^2$.

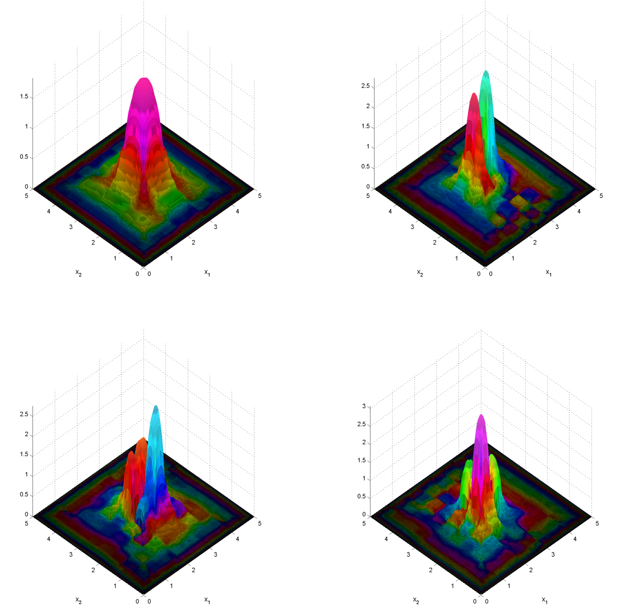

Fig. 2. Phase plots (where absolute value determines height and phase determines colour) of the scaling function φ and associated wavelets ψ_1, ψ_2 and ψ_3 for their support $[0, 5]$, discovered by the DR method ($M = 6$).

IV. Acknowledgements

JAH thanks the CARMA Centre at the University of Newcastle for its continued support. JAH is also supported by ARC Grant DP160101537. Thanks Roy. Thanks HG.

References

[1] Aragón Artacho, F. J., Borwein, J. M., & Tam, M. K. (2014). Douglas–Rachford feasibility methods for matrix completion problems. *The ANZIAM Journal*, 55(4):299–326.

[2] Ayache, A. (1999). Construction of non separable dyadic compactly supported orthonormal wavelet bases for $L^2(\mathbb{R}^2)$ of arbitrarily high regularity. *Revista Matemática Iberoamericana*, 15(1):37–58.

[3] Bauschke, H. H., Combettes, P. L. & Luke, D. R. (2004). Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192.

[4] Belogay, B. & Wang, Y. (1999). Arbitrarily smooth orthogonal nonseparable wavelets in $\mathbb{R}^2$. *SIAM Journal on Mathematical Analysis*, 30(3):678–697.

[5] Bregman, L. M. (1965). The method of successive projection for finding a common point of convex sets. *Doklady Akademii Nauk*, 162(3):688–692.

[6] Cohen, A. (1990). Ondelettes, analysées multirésolutions et filtres miroir en quadrature. *Ann. Inst. H. Poincaré, Anal non linéaire* 7:439–459.

[7] Dao, M. N. & Tam, M. K. (2018). Union averaged operators with applications to proximal algorithms for min-convex functions. *J. Optim. Theory Appl.* published online November 2018.

[8] Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* 41:909–996.

[9] Daubechies, I. (1992). *Ten Lectures on Wavelets.* SIAM.

[10] Franklin, D. J. (2018). *Projection Algorithms for Non-Separable Wavelets and Clifford Fourier Analysis*. PhD Thesis, University of Newcastle, Australia.

[11] He, W. & Lai, M.-J. (2000). Examples of bivariate nonseparable compactly supported orthonormal continuous wavelets. *IEEE Transactions on Image Processing*, 9(5):949–953.

[12] Hesse, R., Luke, D. R. & Neumann, P. (2014). Alternating projections and Douglas–Rachford for sparse affine feasibility. *IEEE Transactions on Signal Processing*, 62(18):4868–4881.

[13] Karoui, A. (2005). A note on the design of nonseparable orthonormal wavelet bases of $L^2(\mathbb{R}^3)$ *Applied Mathematics Letters.* 18(3):293–298.

[14] Kovacevic, J. & Vetterli, M. (1992). Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathbb{R}^n$. *IEEE Transactions on Information Theory*, 38(2):533-555.

[15] Lai, M. & Roach, D.W. (1999). Nonseparable symmetric wavelets with short support. in *Wavelet Applications in Signal and Image Processing VII*. International Society for Optics and Photonics, 3813:132–147.

[16] Lai, M.-J. (2002). Methods for constructing nonseparable compactly supported orthonormal wavelets in *Wavelet Analysis: Twenty Years' Developments.* 231–251. World Scientific

[17] Lions, P.-L. & Mercier, B. (1979). Splitting algorithms for the sum of two nonlinear operators, *SIAM Journal on Numerical Analysis*, 16(6):964–979.

[18] Mallat, S. (1989). Multiresolution approximation and wavelets. *Trans. Amer. Math. Soc.*, 315:69–88.

[19] Meyer, Y. (1986). Ondelettes, fonctions splines et analyses graduées. Lectures given at the University of Torino.

[20] Meyer, Y. (1989). *Wavelets and Operators.* Cambridge University Press.

[21] Phan, H. M. (2016). Linear convergence of the Douglas–Rachford method for two closed sets. *Optimization*, 65(2):369–385.

[22] San Antolin, A. & Zalik, R.A. (2013). A family of nonseparable scaling functions and compactly supported tight framelets *Journal of Mathematical Analysis and Applications.* 404(2):201–211.

[23] Stein, E.M. & Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton.