Multiwavelet Construction via the Lifting Scheme *

Geoffrey M. Davis
Microsoft Research
1 Microsoft Way, Redmond, WA 98052

Vasily Strela
Department of Mathematics & Computer Science, Drexel University
Philadelphia, PA 19104

Radka Turcajová
Department of Mathematics, Sam Houston State University
Huntsville, TX 77341

Abstract

Lifting provides a simple method for constructing biorthogonal wavelet bases. We generalize lifting to the case of multiwavelets, and in so doing provide useful intuition about the additional degrees of freedom made available in the construction of multiwavelets. We show that any compactly supported multiwavelet transform can be decomposed into a sequence of lifting steps. Finally, we compare lifting to the two-scale similarity transform construction method.

1 Introduction

The recent work of Geronimo et al [5] has generated considerable interest in multiwavelet constructions. In contrast to the scalar wavelet case, in which all basis functions are generated from translations and dilations of a single wavelet and scaling function, multiwavelet bases are constructed from translates and dilations of a *vector* of wavelets and scaling functions. Allowing multiple prototypes for the basis elements provides additional

*Portions of this material were presented at the AMS Special Session on Wavelets and Multiwavelets in January, 1997.

degrees of freedom that can be used to construct basis functions with interesting and useful properties. For example, in the scalar wavelet case, no symmetric, orthogonal, and compactly supported bases exist apart from the Haar basis and trivial variations. In contrast, the Geronimo-Hardin-Massopust (GHM) multiwavelet basis, is symmetric, orthogonal, compactly supported, continuous, and reproduces constant and linear functions.

The original construction of the GHM basis, using fractal interpolation functions, is quite complicated. The goal of this paper is to provide a simple method for constructing biorthogonal multiwavelet bases with prescribed properties. We will show how to adapt lifting, a technique proposed for constructing scalar wavelets by Sweldens in [25], for the construction of multiwavelets. Lifting provides valuable intuition about the degrees of freedom made available in the multiwavelet case, and it motivates some new design criteria for using these new degrees of freedom effectively. As an example we use the new degrees of freedom to obtain finer control of the tradeoff between length of support and vanishing moments. We also show that all compactly supported biorthogonal multiwavelet bases can be achieved by applying a finite sequence of simple lifting steps to a simple initial basis. The result parallels that of [4], but there are additional complications in the multiwavelet case that we discuss.

In the last secton, we will compare lifting to the two-scale similarity transform. This is an alternative method of multiwavelet basis construction that parallels Daubechies' spectral factorization method.

Independently developed presentations of lifting for multiwavelets can be found in [6] and [15].

2 Lifting for Scalar Wavelets

Lifting is an iterative procedure for constructing biorthogonal wavelet bases. The procedure takes a simple initial basis and, through successive modifications of the basis functions, fine-tunes such properties as the number of vanishing moments and the order of approximation. The canonical lifting step involves either modifying the wavelets while holding the scaling functions fixed or modifying the dual wavelets while holding the dual scaling functions fixed. This cycle of modifications is repeated until the desired basis properties are obtained.

2.1 Notation

Let $\varphi(x)$ and $\psi(x)$ be a scaling and a wavelet function, respectively, that satisfy the recurrence relations

$$\begin{aligned} \varphi(x) &= \sqrt{2}\sum_{k\in\mathrm{Z}} h_k \varphi(2x-k) \\ \psi(x) &= \sqrt{2}\sum_{k\in\mathrm{Z}} g_k \varphi(2x-k) \end{aligned} \quad (1)$$

and that generate a multiresolution analysis of L^2. We will denote the dual scaling function and dual wavelet functions by $\tilde{\varphi}(x)$ and $\tilde{\psi}(x)$, respectively, and their recurrence relation coefficients by $\tilde{h}_k$ and $\tilde{g}_k$. The conditions of biorthogonality can be expressed as

$$\begin{aligned} \langle\varphi(x-k),\tilde{\varphi}(x-l)\rangle &= \delta(k-l), & \langle\psi(x-k),\tilde{\varphi}(x-l)\rangle &= 0, \\ \langle\psi(x-k),\tilde{\psi}(x-l)\rangle &= \delta(k-l), & \langle\varphi(x-k),\tilde{\psi}(x-l)\rangle &= 0, \end{aligned} \quad (2)$$

for all $k, l \in \mathbb{Z}$.

For the recurrence coefficients h_k, we define the *symbol* $h(z)$ by $h(z) = \sum_{k\in\mathrm{Z}} h_k z^{-k}$. The symbols $g(z)$, $\tilde{h}(z)$, and $\tilde{g}(z)$ are defined similarly. In this paper, we focus exclusively on the contruction of compactly supported biorthogonal bases. This means that only a finite number of the coefficients h_k, $\tilde{h}_k$, g_k, and $\tilde{g}_k$ will be nonzero. The symbols therefore are polynomials. (Here and elswhere in this paper, by polynomial we mean a Laurent polynomial, i.e. a finite series.)

2.2 Characterizing Families of Biorthogonal Wavelet Bases

The theoretical motivation for lifting is a lemma due to Vetterli and Herley [30] that provides a simple parametrization of all biorthogonal families with fixed scaling function φ that satisfy perfect reconstruction conditions.

Lemma 1 (Vetterli-Herley) *Suppose that* $\{\varphi, \psi_A, \tilde{\varphi}_A, \tilde{\psi}_A\}$ *and* $\{\varphi, \psi_B, \tilde{\varphi}_B, \tilde{\psi}_B\}$ *are compactly supported families satisfying the conditions of biorthogonality (2). Let* $\{h(z),\, g_A(z),\, \tilde{h}_A(z),\, \tilde{g}_A(z)\}$ *and* $\{h(z),\, g_B(z),\, \tilde{h}_B(z),\, \tilde{g}_B(z)\}$ *be the symbols for the corresponding recurrence coefficients. Then, up to a monomial in* z*, we have*

$$\begin{aligned} g_B(z) &= g_A(z) + s(z^2)h_A(z) \\ \tilde{h}_B(z) &= \tilde{h}_A(z) - \overline{s(\bar{z}^{-2})}\tilde{g}_A(z) \\ \tilde{g}_B(z) &= \tilde{g}_A(z) \end{aligned} \quad (3)$$

where $s(z)$ *has finite degree.*

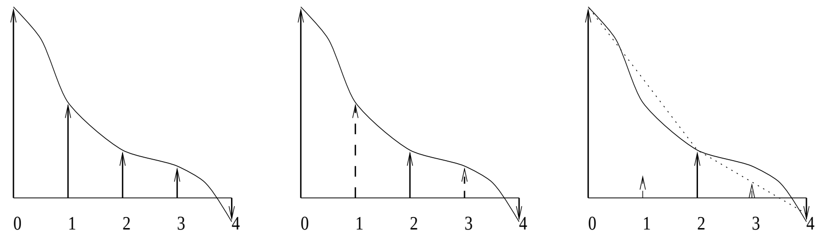

Figure 1: On the left we have a function $f(x)$. Its resolution 0 approximation $A_0 f(x)$ consists of a set of Diracs at unit intervals. In the center we have the coarser resolution -1 approximation (the solid samples) together with the resolution -1 wavelet coefficients (the dashed samples). On the right we have the resolution -1 approximation and the modified wavelet coefficients given by the prediction errors from a linear predictor. Note that because $f(x)$ is smooth and slowly varying, the prediction errors are small.

The Vetterli-Herley Lemma is a special case of Lemma 2 below, and we defer the proof until later.

The Vetterli-Herley lemma provides a complete characterization of all compactly supported wavelets that complement a fixed scaling function $\varphi(x)$ in such a way that the resulting dual scaling and wavelet functions are also compactly supported. Every finite degree polynomial s yields a new finite biorthogonal filter set. Note, however, that not all of these filters lead to biorthogonal functions: h and $\tilde{h}$ must also satisfy some additional constraints to ensure that well-behaved solutions exist to the recurrence relations. The omitted monomial corresponds to a trivial change of basis by replacing the wavelet functions by their integer translates and rescaling them.

What is interesting about the lemma is its interpretation in the spatial domain. Take an initial family $\{\varphi_A, \psi_A, \tilde{\varphi}_A, \tilde{\psi}_A\}$. For a fixed scaling function φ_A, the possible wavelets are

$$\psi_B(x) = \psi_A(x) + \sum_{k \in \mathrm{Z}} s_k \varphi_A(x-k). \quad (4)$$

Thus, we can generate a new wavelet from an old by adding linear combinations of translates of the scaling function. Sweldens[25] calls such a modification of φ_A a *lifting step*.

2.3 Contruction of scalar wavelets by lifting

One can obtain useful intuition about lifting by considering an initial basis derived from the degenerate set

$$\begin{array}{ll} \varphi(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{otherwise} \end{cases} & \psi(x) = \begin{cases} 1 & x = \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \\ \tilde{\varphi}(x) = \delta(x) & \tilde{\psi}(x) = \delta(x - \frac{1}{2}). \end{array} \tag{5}$$

The resolution j approximation of a function $f(x)$ is given by $A_j f(x) = \sum \langle f(x), \tilde{\varphi}_k^j(x) \rangle \varphi_k^j(x)$; here $\varphi_k^j(x) = 2^{j/2}\varphi(2^j x - k)$ and $\tilde{\varphi}_k^j$ is defined similarly. For this particular set of functions the approximation $A_j f$ consists of samples of f, with a sampling rate determined by j. We are sacrificing some rigor for the sake of insight here, and there are a few analytical complications that we will note only in passing. For example, $A_j f(x)$ converges only weakly to $f(x)$ and only when $f(x)$ is sufficiently smooth. Furthermore, $A_j f(x)$ is not even a function, but rather a distribution. Nevertheless, this example is quite helpful in conveying the essence of lifting.

In our example, each step of the "wavelet transform" divides the samples of the current resolution j approximation of f into samples at even points (the "low-pass" coefficients) and odd points (the "high-pass" coefficients). In typical applications of wavelets, the goal of the transform is to obtain an efficient representation of a piecewise smooth function, i.e. we would like to be able to approximate a given function accurately with a linear combination of a small number of wavelets. The transform in our example does not in general yield such an efficient representation (unless we have *a priori* reason to assume that our function samples will be mostly zero).

If the function in question is smooth and slowly varying, we can form a more efficient representation by using the even samples to predict the odd samples. For example, we can use as a predictor a degree-N polynomial fitted to a set of even coefficients surrounding the coefficient to be predicted. Rather than keeping the odd samples, we keep the difference between the odd samples and a prediction based on the even neighbors. Figure 1 illustrates this procedure.

The even and odd coefficients have the form $\langle f(x), \tilde{\varphi}(x-k) \rangle$ and $\langle f(x), \tilde{\psi}(x-k) \rangle$, respectively. Replacing the odd coefficients with the differences from their predicted values is equivalent to replacing $\tilde{\psi}(x-k)$ with $\tilde{\psi}(x-k) - \sum c_k \tilde{\varphi}(x-k)$, where c_k are the coefficients of the linear combination of even samples used to predict the odd samples. Thus, we see that this new, more efficient representation is obtained by performing a lifting step on the dual wavelet. Applying a predictor that is exact (i.e. the updated odd coefficients will all be zero) for polynomials of order up to N corresponds to creating a new dual wavelet with $N+1$ vanishing moments.

Performing a subsequent lifting step that adds vanishing moments to the primal wavelet results in a low-pass filtering of the even coefficients. Performing several lifting steps corresponds to applying multi-level prediction and yields a wider variety of wavelet bases. Daubechies and Sweldens showed in [4] that any compactly supported biorthogonal wavelet basis can be obtained (up to some translations and rescaling) by applying a finite number of lifting steps to the trivial Dirac basis described above.

3 The Motivation for Multiwavelets

Singularities play an important role in many signal processing applications. For example, discontinuities in the brightness function correspond to edges in images. If we are to obtain a compact representation for functions containing singularities, we need to pay close attention to the transform-domain behavior of the singularities.

Wavelet bases with high approximation order are better able to approximate smooth functions with a small number of terms than bases with lower order. However, high approximation order comes at the price of longer support for the dual wavelet. When a predictor of order N is used, a wavelet coefficient is negligible if the function is very nearly a polynomial of order not greater than N over the entire support of appropriately translated and dilated wavelet $\tilde{\psi}$. If it is not well-approximated by an order N polynomial, the wavelet coefficient will be large. The wider the support of $\tilde{\psi}$ is, the more large coefficients will be generated by each singularity. To obtain compact representation of a function with singularities we thus need to balance the approximation order needed to efficiently approximate smooth portions of the function with the length of support that sufficiently localizes the singularities. The right balance depends on the smoothness of the function and the prevalence of singularities. In image processing applications, the balance is typically attained with two to four orders of approximation.

Consider our Dirac basis example. Because we are predicting odd coefficients from even coefficients, each additional order after the second increases the length of the support of the filter by two. The tradeoff between filter length and order would be less severe if we were able to use both even and odd coefficients for prediction. However, in general, using for prediction even and odd coefficients simultaneously makes it difficult to invert the transform. With a minor modification to our lifting procedure, however, we can incorporate odd coefficients into the prediction process and thereby gain more control over the tradeoff between support and order of approximation.

We divide the samples we are trying to predict, the odd samples, into

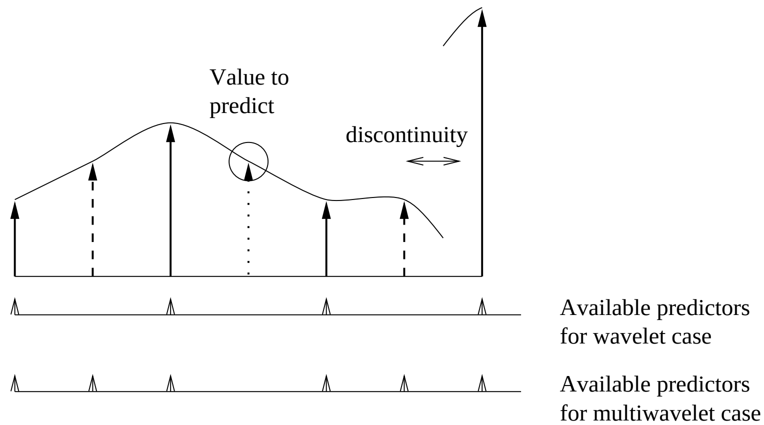

Figure 2: Predicting a function sample near an edge in the wavelet case and the multiwavelet case.

two classes: the samples with indices $4k+1$ and those with indices $4k+3$. We first predict the points with indices $4k+3$ using both the even points as well as the index $4k+1$ points. Next we predict the points with indices $4k+1$ using the even points. The result, illustrated in figure 2, is shorter prediction window for the $4k+3$ predictor and no change for the $4k+1$ predictor.

The process of partitioning the points to be predicted into two sets makes possible a better tradeoff between filter length and order of approximation for half of our coefficients. However, the result is two different dual wavelets and two different primal scaling functions. The first set of wavelets corresponds to using both the even samples and the $4k+1$-indexed samples for prediction, and the second set corresponds to using only the even-indexed samples for prediction. We have gone from a system of biorthogonal wavelets to a system of biorthogonal *multiwavelets*. In fact, as we will show, the *only* significant difference between wavelets and multiwavelets is the fact that we can predict wavelet coefficients with other wavelet coefficients and we can predict scaling function coefficients using other scaling functions coefficients.

In general, to construct a multiwavelet family with r different scaling functions, we subdivide our initial samples x_k into $2r$ *polyphase components* $\{x_{rk+n},\ k \in \mathbb{Z}\}$. Polyphase components with n even correspond to scaling function coefficients. The odd components correspond to wavelet coefficients. In performing prediction steps, we are free to use coefficients from *any* of the other polyphase components. We will formalize this idea in the sections that follow.

4 Lifting for Multiwavelets

4.1 Notation

For multiwavelet bases, instead of one scaling function $\varphi(x)$ and one wavelet $\psi(x)$ we have a vector of r scaling functions (a *multiscaling function*), $\boldsymbol{\Phi}(x) = [\varphi_0(x) \ \ldots \ \varphi_{r-1}(x)]^T$ and vector of r wavelet functions (a *multiwavelet*) $\boldsymbol{\Psi}(x) = [\psi_0(x) \ \ldots \ \psi_{r-1}(x)]^T$. These vectors satisfy recurrence relations similar to (1):

$$\begin{aligned} \boldsymbol{\Phi}(x) &= \sqrt{2}\sum_{k\in Z} \boldsymbol{H}_k \boldsymbol{\Phi}(2x-k) \\ \boldsymbol{\Psi}(x) &= \sqrt{2}\sum_{k\in Z} \boldsymbol{G}_k \boldsymbol{\Phi}(2x-k). \end{aligned} \tag{6}$$

Here $\boldsymbol{H}_k$ and $\boldsymbol{G}_k$ are $r \times r$ real matrices and we assume that only a finite number of them is nonzero. We also have r vectors of dual scaling and wavelet functions $\tilde{\boldsymbol{\Phi}}$ and $\tilde{\boldsymbol{\Psi}}$, respectively, satisfying similar recurrence relations with coefficients $\tilde{\boldsymbol{H}}_k$ and $\tilde{\boldsymbol{G}}_k$. We now have *matrix symbols* $\boldsymbol{H}(z)$, $\tilde{\boldsymbol{H}}(z)$, $\boldsymbol{G}(z)$, and $\tilde{\boldsymbol{G}}(z)$, where $\boldsymbol{H}(z) = \sum_{k\in Z} \boldsymbol{H}_k z^{-k}$, and so on.

The conditions of biorthogonality we can be also written in matrix notation in a fashion reminding the similar equations in the scalar case:

$$\begin{aligned} &\langle \boldsymbol{\Phi}(x-k), \tilde{\boldsymbol{\Phi}}(x-l)\rangle = \delta(k-l)\boldsymbol{I}, && \langle \boldsymbol{\Psi}(x-k), \tilde{\boldsymbol{\Phi}}(x-l)\rangle = \boldsymbol{0} \\ &\langle \boldsymbol{\Psi}(x-k), \tilde{\boldsymbol{\Psi}}(x-l)\rangle = \delta(k-l)\boldsymbol{I} && \langle \boldsymbol{\Phi}(x-k), \tilde{\boldsymbol{\Psi}}(x-l)\rangle = \boldsymbol{0}. \end{aligned} \tag{7}$$

Here, $\langle \, . \, , \, . \, \rangle$ does not stand for a proper inner product. Rather, it is a vector generalization of the L^2 inner product, $\langle \boldsymbol{f}(x), \boldsymbol{g}(x)\rangle = \int \boldsymbol{f}(x)(\boldsymbol{g}(x))^* \, dx$ ($*$ denotes the ordinary conjugated transpose). The value yielded thus is an $r \times r$ matrix.

4.2 Characterizing Families of Biorthogonal Multiwavelet Bases

As in the scalar case, we obtain a simple characterization of the set of multiwavelet bases that are biorthogonal to fixed multiscaling function. Our characterization generalizes the Vetterli-Herley lemma, and as in the scalar case, we will use this lemma as the starting point for basis construction using lifting.

As a preliminary, we define $\boldsymbol{T}^\dagger(z)$ to be the *paraconjugated transpose* of $\boldsymbol{T}$, $\boldsymbol{T}^\dagger(z) = (\boldsymbol{T}(\bar{z}^{-1}))^*$. If $\boldsymbol{T}(z) = \sum_{k\in\mathrm{Z}} \boldsymbol{T}_k z^{-k}$, then $\boldsymbol{T}^\dagger(z) = \sum_{k\in\mathrm{Z}} \tilde{\boldsymbol{T}}_k^* z^k$, where $*$ denotes the ordinary conjugated transpose. Note that the paraconjugated transpose coincides with the ordinary conjugated transpose on the unit circle.

Lemma 2 *Suppose we have two compactly supported multiwavelet families that satisfy the conditions of biorthogonality and that share a multiscaling function* $\boldsymbol{\Phi}$. *Denote the families* $\{\boldsymbol{\Phi}, \boldsymbol{\Psi}_A, \tilde{\boldsymbol{\Phi}}_A, \tilde{\boldsymbol{\Psi}}_A\}$ *and* $\{\boldsymbol{\Phi}, \boldsymbol{\Psi}_B, \tilde{\boldsymbol{\Phi}}_B, \tilde{\boldsymbol{\Psi}}_B\}$. *Let* $\{\boldsymbol{H}(z), \boldsymbol{G}_A(z), \tilde{\boldsymbol{H}}_A(z), \tilde{\boldsymbol{G}}_A(z)\}$ *and* $\{\boldsymbol{H}(z), \boldsymbol{G}_B(z), \tilde{\boldsymbol{H}}_B(z), \tilde{\boldsymbol{G}}_B(z)\}$ *be the symbols for the recurrence relations for these functions. Then*

$$\begin{aligned} \boldsymbol{G}_B(z) &= \boldsymbol{T}(z^2)(\boldsymbol{G}_A(z) + \boldsymbol{S}(z^2)\boldsymbol{H}_A(z))) \\ \tilde{\boldsymbol{H}}_B(z) &= \tilde{\boldsymbol{H}}_A(z) - \boldsymbol{S}^\dagger(z^2)\tilde{\boldsymbol{G}}_A(z) \\ \tilde{\boldsymbol{G}}_B(z) &= (\boldsymbol{T}^\dagger(z^2))^{-1}\tilde{\boldsymbol{G}}_A(z), \end{aligned} \tag{8}$$

where $\boldsymbol{S}(z)$ *and* $\boldsymbol{T}(z)$ *are of finite degree and the determinant of* $\boldsymbol{T}(z)$ *is a monomial.*

As in the scalar case, the lemma has a straightforward interpretation in the spatial domain. When $\boldsymbol{T}(z)$ is the identity, we simply obtain a vector analog of (4). Namely, we have $\boldsymbol{\Psi}_B(x)$ is $\boldsymbol{\Psi}_A(x)$ plus a linear combination of multiscaling functions. Because $\boldsymbol{S}(z)$ is not necessarily diagonal, the components of $\Psi_B(x)$ can all be built from different linear combinations of scaling functions, which gives us some new (but not very interesting) degrees of freedom in our construction.

An important difference between scalar wavelets and multiwavelets comes from nontrivial values of $\boldsymbol{T}(z)$. We have the restriction that $\boldsymbol{T}(z)$ is of finite degree and has a monomial determinant. We will call such matrices *unimodular*. From Cramer's rule, we see that it implies that $\boldsymbol{T}(z)$ is invertible and that $\boldsymbol{T}^{-1}(z)$ is also of finite degree. The inverse must be finite so as the resulting dual wavelet fuctions were compactly supported. The allowable modifications to $\boldsymbol{\Psi}$ thus consist of a vector lifting step that is similar to that for scalar wavelets followed by a finitely invertible intermixing of the different wavelet components of $\boldsymbol{\Psi}$. In scalar case, where there is only one scaling function, the only such possible intermixing is choosing a different translate for a prototype and recaling it. $\boldsymbol{T}(z)$ corresponds to the monomial omitted in (3).

In the example in Section 3, we split our samples into four polyphase components. Predicting the odd samples from the even corresponds to the wavelet case. Using two different predictors that use even samples to predict the odd, one for the $4k + 1$-indexed coefficients and one for the

$4k+3$-indexed coefficients, corresponds an application of Lemma 2 with $\boldsymbol{T}(z)=\boldsymbol{I}$. Our use of both even samples and $4k+1$-indexed samples to predict the $4k+3$-indexed samples corresponds to using a nontrivial $\boldsymbol{T}(z)$, as we are using wavelet coefficients to predict other wavelet coefficients. The constraint of unimodularity on $\boldsymbol{T}(z)$ is ensures that the transform will be invertible.

Before we prove Lemma 2, we introduce the polyphase representation, which will simplify the proof.

4.3 The Polyphase Representation

The polyphase representation[28] provides a more compact way of describing the mechanics of lifting. In the polyphase representation, the conditions of biorthogonality (2) and (7) are expressed in terms of a simple maxtix product. A lifting step is equivalent to performing an elementary matrix operation on each of the matrices in this product.

We define the *even and odd polyphase components* $\boldsymbol{H}_e(z)$ and $\boldsymbol{H}_o(z)$ for the recurrence coefficients $\boldsymbol{H}_k$ by

$$\boldsymbol{H}_e(z^2)=\frac{1}{2}[\boldsymbol{H}(z)+\boldsymbol{H}(-z)], \qquad \boldsymbol{H}_o(z^2)=\frac{z}{2}[\boldsymbol{H}(z)-\boldsymbol{H}(-z)].$$

The polyphase components for the other recurrence coefficients $\boldsymbol{G}_k$, $\tilde{\boldsymbol{G}}_k$, and $\tilde{\boldsymbol{H}}_k$ are defined similarly. The *polyphase matrix* $\boldsymbol{P}(z)$ is then given by

$$\boldsymbol{P}(z)=\begin{bmatrix} \boldsymbol{H}_e(z) & \boldsymbol{H}_o(z) \\ \boldsymbol{G}_e(z) & \boldsymbol{G}_o(z) \end{bmatrix} = \sum_{k\in \mathrm{Z}} \begin{bmatrix} \boldsymbol{H}_{2k} & \boldsymbol{H}_{2k+1} \\ \boldsymbol{G}_{2k} & \boldsymbol{G}_{2k+1} \end{bmatrix} z^{-k}. \qquad (9)$$

If the recurrence relations (1) are uniquely satisfied by compactly supported scaling functions $\varphi(x)$ and $\tilde{\varphi}(x)$, it it is straightforward to show that the conditions of biorthogonality will be satisfied if and only if the perfect reconstruction condition

$$\boldsymbol{P}(z)\tilde{\boldsymbol{P}}^{\dagger}(z)=\boldsymbol{I} \qquad (10)$$

is satisfied [20] for any $z\in\mathbb{C}\setminus\{0\}$.

The perfect reconstruction condition has a lot of applications. For example, it shows that to given primal symbols $\boldsymbol{H}(z)$ and $\boldsymbol{G}(z)$ there exists a unique pair of dual symbols $\tilde{\boldsymbol{H}}(z)$ and $\tilde{\boldsymbol{G}}(z)$ and, once the primal symbols are known, the dual ones can be computed by inverting and transposing the polyphase matrix and recombining the symbols from its entries. It also implies that, if both $\boldsymbol{P}(z)$ and $\tilde{\boldsymbol{P}}(z)$ are to be of finite degree, they must be unimodular.

In this paper we will focus only on the problem of generating symbols that satisfy the perfect reconstruction conditions (10). Characterizing the conditions under which recurrence relations satisfying (10) converge is addressed in [3, 2, 9, 10, 18]).

4.4 Proof of Lemma 2

To prove Lemma 2, we must show that there exists a finite-degree matrix $\boldsymbol{S}(z)$ and a unimodular matrix $\boldsymbol{T}(z)$ that satisfy

$$\boldsymbol{P}_B(z) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}(z) \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{S}(z) & \boldsymbol{I} \end{bmatrix} \boldsymbol{P}_A(z). \tag{11}$$

We have

$$\boldsymbol{P}_A(z) = \begin{bmatrix} \boldsymbol{H}_e(z) & \boldsymbol{H}_o(z) \\ \boldsymbol{G}_{A,e}(z) & \boldsymbol{G}_{A,o}(z) \end{bmatrix}, \qquad \boldsymbol{P}_B(z) = \begin{bmatrix} \boldsymbol{H}_e(z) & \boldsymbol{H}_o(z) \\ \boldsymbol{G}_{B,e}(z) & \boldsymbol{G}_{B,o}(z) \end{bmatrix}.$$

Since both sets satisfy the perfect reconstruction conditions, that is,

$$\boldsymbol{P}_A(z)\tilde{\boldsymbol{P}}_A^{\dagger}(z) = \boldsymbol{P}_B(z)\tilde{\boldsymbol{P}}_B^{\dagger}(z) = \boldsymbol{I}$$

and the upper half of $\boldsymbol{P}_A(z)$ and $\boldsymbol{P}_B(z)$ are the same, we have

$$\boldsymbol{P}_B(z)\tilde{\boldsymbol{P}}_A^{*}(z^{-1}) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{G}_{B,e}(z)\tilde{\boldsymbol{H}}_{A,e}^{\dagger}(z) + \boldsymbol{G}_{B,o}(z)\tilde{\boldsymbol{H}}_{A,o}^{\dagger}(z) & \boldsymbol{G}_{B,e}(z)\tilde{\boldsymbol{G}}_{A,e}^{\dagger}(z) + \boldsymbol{G}_{B,o}(z)\tilde{\boldsymbol{G}}_{A,o}^{\dagger}(z) \end{bmatrix}. \tag{12}$$

Now $\boldsymbol{P}_B(z)\tilde{\boldsymbol{P}}_A^{*}(z^{-1})$ must be unimodular because $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ are both unimodular. Since $\det(\boldsymbol{P}_B(z)\tilde{\boldsymbol{P}}_A^{\dagger}(z))$ is equal to the determinant of its lower right block, we have

$$\boldsymbol{T}(z) = \boldsymbol{G}_{B,e}(z)\tilde{\boldsymbol{G}}_{A,e}^{\dagger}(z) + \boldsymbol{G}_{B,o}(z)\tilde{\boldsymbol{G}}_{A,o}^{\dagger}(z)$$

is also unimodular. Let

$$\boldsymbol{S}(z) = \boldsymbol{T}(z)^{-1}(\boldsymbol{G}_{B,e}(z)\tilde{\boldsymbol{H}}_{A,e}^{\dagger}(z) + \boldsymbol{G}_{B,o}(z)\tilde{\boldsymbol{H}}_{A,o}^{\dagger}(z)).$$

Factoring the matrix on the right hand side of (12) and multiplying both sides of the equation by $\tilde{\boldsymbol{P}}_A(z)$, we obtain (11). Now,

$$\tilde{\boldsymbol{P}}_B(z) = (\boldsymbol{P}_B^{\dagger}(z))^{-1} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{T}^{\dagger}(z))^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{S}^{\dagger}(z) \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \tilde{\boldsymbol{P}}_A(z).$$

The rest of the statement is obtained by extracting the polyphase components and recombining them to form the symbols.

We point out here that it is also true that any matrix polynomial $\boldsymbol{S}$ and any unimodular matrix polynomial $\boldsymbol{T}$ define via equations (8) symbols $\boldsymbol{G}_B(z)$, $\tilde{\boldsymbol{H}}_B(z)$, $\tilde{\boldsymbol{G}}_B(z)\}$ such that the corresponding polyphase matrices $\boldsymbol{P}_B(z)$ (created from $\boldsymbol{H}(z)$ and $\boldsymbol{G}_B(z)$) and $\tilde{\boldsymbol{P}}_B(z)$ satisfy the perfect reconstruction condition.

4.5 New Degrees of Freedom

To better understand the new degrees of freedom we obtain from going from scalar wavelets to multiwavelets, we now consider the effects of different types of *simple lifting steps*, that is, lifting steps involving using a single polyphase component of a signal to predict another polyphase component. Each simple lifting step has the effect of multiplying the polyphase matrix from the left hand side by a matrix with unit diagonal and a single nonzero off-diagonal entry. There are four classes of simple lifting steps which correspond to adding linear combinations of wavelets/scaling functions to a wavelet/scaling function. The location of the non-zero off-diagonal term determines the class.

The effect of a simple lifting step with the off-diagonal term in the lower left/upper right quadrant is to add a linear combination of primal/dual scaling functions to a primal/dual wavelet. These types of lifting step are simple extensions of scalar lifting steps.

The simple lifting steps with their non-zero terms in the lower right or upper left quadrants have no analog in the scalar wavelet case. A lifting step with nonzero term in the lower right quadrant adds a linear combination of the translates of a primal/dual wavelet to a different primal/dual wavelet. A lifting step with nonzero term in the upper left quadrant corresponds to predicting one set of scaling function coefficients using a different set of coefficients. The effect on the functions of such a prediction is much more complicated than in the wavelet case because the scaling functions appear on both sides of the recurrence relations. In general, all scaling functions are modified as a result of this type of lifting step, and the new wavelets are the solutions of a different set of recurrence relations. Simply adding a linear combination of the translates of one of the scaling functions to another scaling function (or any change of basis in the approximation spaces, for that matter) can be expressed as a *two scale similarity transform* in which the scaling symbol is multiplied by mutually related factors from the left and the right hand side. We discuss the two-scale similarity transform in section 6.

4.6 Factoring Multiwavelet Transforms into Lifting Steps

We can decompose *any* compactly supported biorthogonal multiwavelet transform into a finite sequence of simple lifting steps and a rescaling step. Sometimes it may be useful to allow besides lifting steps and rescaling also permutations in the factorization. They are not necessary, but they may help to reduce the overall number of factors. A constructive proof can be obtained by extending the derivation of [4] to polyphase matrices larger than 2×2 or by adapting the factorization used to place a matrix in Smith-McMillan form, which is well known in engineering circles, see e.g. [28].

In general, there are numerous ways to factorize a polyphase matrix into simple lifting steps. We can narrow the range of choices by introducing a canonical form that groups together lifting steps of the same kind. The result is that we can obtain any set of biorthogonal multiscaling functions and multiwavelets by starting from a basis defined by a polyphase matrix of the form $\begin{bmatrix} \boldsymbol{T}_1(z) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}$, where $\boldsymbol{T}_1(z)$ is unimodular. We first successively apply lifting steps to primal and dual wavelet functions, each time using only the scaling functions to alter them, not the other wavelets. When we reach the desired multiscaling functions, we finish by combining wavelet functions only, among themselves. Having a factorization with less redundancy and a rigid pattern of lifting steps is helpful particularly when a factorization based parameterization is used in filter design by numerical optimization as in [11] or [27].

Theorem 3 *A polyphase matrix $\boldsymbol{P}(z)$ corresponding to compactly supported biorthogonal multiwavelets can be expressed in the form*

$$\begin{aligned} \boldsymbol{P}(z) \quad = \quad & \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}_2(z) \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{S}_K(z) & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{W}_K(z) \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \cdots \\ & \cdots \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{S}_1(z) & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{W}_1(z) \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_1(z) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}. \end{aligned} \tag{13}$$

Here, $\boldsymbol{S}_k(z)$ and $\boldsymbol{W}_k(z)$, $k = 1, \ldots, K$ are of finite degree. The diagonal blocks $\boldsymbol{T}_m(z)$, $m = 1, 2$, are unimodular and therefore

$$\boldsymbol{T}_m(z) = \boldsymbol{D}_m(z)\boldsymbol{L}_{m,N_m}(z)\boldsymbol{U}_{m,N_m}(z) \cdots \boldsymbol{L}_{m,1}(z)\boldsymbol{U}_{m,1}(z), \tag{14}$$

where $\boldsymbol{L}_{m,n}(z)$ and $\boldsymbol{U}_{m,n}(z)$ are also of finite degree, lower and upper triangular, respectively, with ones on the diagonal and $\boldsymbol{D}_m(z)$ are diagonal matrices with monomials on the diagonal.

Proof. We first note that by grouping adjacent upper/lower triangular simple lifting factors together, we can factor any unimodular matrix $\boldsymbol{P}(z)$ into

$$\boldsymbol{P}(z) = \boldsymbol{D}(z)\boldsymbol{L}_K(z)\boldsymbol{U}_K(z)\cdots\boldsymbol{L}_1(z)\boldsymbol{U}_1(z)$$

where $\boldsymbol{U}_k(z)$ and $\boldsymbol{L}_k(z)$ are upper and lower triangular polynomial matrices, respectively, with ones on the diagonal and $\boldsymbol{D}(z)$ is diagonal with monomials on the diagonal. We next partition each of the triangular factors into four $r \times r$ blocks and factor out the diagonal blocks of each as follows:

$$\begin{bmatrix} \boldsymbol{L}_{1,1}(z) & \boldsymbol{0} \\ \boldsymbol{L}_{2,1}(z) & \boldsymbol{L}_{2,2}(z) \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_{1,1}(z) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{L}_{2,1}(z) & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{L}_{2,2}(z) \end{bmatrix}.$$

Notice that the diagonal blocks are unimodular (i.e., they have a polynomial inverse). We can move the block diagonal matrices to the ends of the factorization using the relation

$$\begin{bmatrix} \boldsymbol{A}(z) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}(z) \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{S}(z) & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{B}(z)\boldsymbol{S}(z)\boldsymbol{A}(z)^{-1} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}(z) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}(z) \end{bmatrix}$$

and similar relations for upper triangular matrices. The diagonal matrices with monomials on the diagonal can be moved in the same way.

We conclude with the observation that when the primal polyphase matrix is factored into lifting steps, obtaining the dual polyphase matrix (and, from there, the dual filters and functions) is easy. From the perfect reconstruction condition (10), the dual polyphase matrix is a paraconjugated transpose of the inverse of the given polyphase matrix. The inverse of each simple lifting factor is the factor itself, only its nonzero off-diagonal entry has the opposite sign. Via the decomposition into simple lifting factors, formulas for the inverses of factors in (13) and (14) can be derived.

5 Adding Vanishing Moments with Lifting

5.1 Moments of wavelet functions

One particularly attractive property of lifting is that it provides a simple means to construct multiwavelet bases with prescribed approximation order. We may start, for example, with the Dirac basis $\tilde{\varphi}_1(x) = \delta(x)$, $\tilde{\varphi}_2(x) = \delta(x - \frac{1}{2})$, $\tilde{\psi}_1(x) = \delta(x - \frac{1}{4})$, $\tilde{\psi}_2(x) = \delta(x - \frac{3}{4})$, for which the polyphase matrix is given by

$$\tilde{P}(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (15)$$

and successively add vanishing moments to the dual and primal wavelets until the requisite order of approximation is reached. The following result is central to our construction procedure.

Theorem 4 *Let $\mathbf{\Phi}$, $\mathbf{\Psi}$, $\tilde{\mathbf{\Phi}}$, and $\tilde{\mathbf{\Psi}}$ be multiscaling and multiwavelet functions giving rise to a biorthogonal multiresolution analysis for L^2. Suppose the multiwavelet $\mathbf{\Psi}$ has n vanishing moments, i.e., $\int \mathbf{\Psi}(x)x^p dx = \mathbf{0}, p = 0, \ldots, n-1$. If we apply a lifting step to modify $\tilde{\mathbf{\Psi}}$, then the resultant $\mathbf{\Psi}^{new}$ has also n vanishing moments.*

Proof. The idea is a simple one. Because the original primal wavelets have n vanishing moments, the dual scaling functions must have approximation order n. The lifting step changes $\mathbf{\Psi}_k(x)$ as well as the primal wavelets and scaling functions, but it does not change the dual scaling functions. Hence, after the lifting step the dual scaling functions still have approximation order n, which means that the new primal wavelets must have n vanishing moments.

Because of this theorem, we can add vanishing moments to the primal and dual wavelets in stages. We first add moments to the dual wavelets, then to the primals, then to the duals, and so on. We require the following expression for the moments so that we can cancel them out.

Lemma 5 *Let $\boldsymbol{M}_{\Phi,n} = \int \mathbf{\Phi}(x)x^n dx$ and $\boldsymbol{M}_{\Psi,n} = \int \mathbf{\Psi}(x)x^n dx$. Then we have*

$$\boldsymbol{M}_{\Phi,n} = \sum_{j=0}^{n} \sum_{k \in Z} H_k \binom{n}{j} k^j \boldsymbol{M}_{\Phi,n-j} \tag{16}$$

$$\boldsymbol{M}_{\Psi,n} = \sum_{j=0}^{n} \sum_{k \in Z} G_k \binom{n}{j} k^j \boldsymbol{M}_{\Phi,n-j} \tag{17}$$

Constructing a basis with specified numbers of vanishing moments is now a straightforward procedure. We choose an initial wavelet $w_0(x)$ to modify, where $w_0(x)$ is either a primal or dual wavelet. We then choose a set of k translates of scaling and wavelet functions $w_1(x), \ldots, w_k(x)$ that we will use to modify the function $w_0(x)$ via a lifting step. Our new function will be

$$w_0^{new} = w_0(x) - \sum_{i=1}^{k} c_i w_i(x). \tag{18}$$

The coefficients c_i are chosen so that the modified function has k vanishing moments, i.e. so that $\int w_0^{new}(x)x^n dx = 0$ for $0 \leq n < k$. The coefficients

c_i satisfy the linear system $\boldsymbol{Mc} = \boldsymbol{m}_0$, where M is the $k \times k$ matrix with entries $M_{i,j} = \int w_j x^i dx$ and $\boldsymbol{m}_0$ is the vector with entries $m_{0,i} = \int w_0 x^i dx$.

We repeat this lifting procedure for primal and dual wavelets until we obtain a basis with a prescribed number of vanishing moments. We can generate a variety of bases by varying the order in which we add the vanishing moments and by varying the functions we use for moment cancelation in each lifting step. We can start from a Dirac basis and generate a new basis altogether or we can modify an existing basis. More general procedure of increasing the approximation order within lifting framework is considered in [12]

As an example of this lifting procedure, we take an initial Dirac basis and use lifting to add four vanishing moments to the primal and dual wavelets. In our first lifting step we use both wavelets and scaling functions to add 4 vanishing moments to the wavelet $\tilde{\psi}_1(x)$:

$$\tilde{\psi}_1^{new}(x) = \tilde{\psi}_1^{old}(x) + \frac{1}{6}\tilde{\psi}_2(x+1) - \frac{2}{3}\tilde{\varphi}_1(x) - \frac{2}{3}\tilde{\varphi}_2(x) + \frac{1}{6}\tilde{\psi}_2(x) \tag{19}$$

The use of wavelets in the lifting step is possible because we are constructing multiwavelets rather than scalar wavelets.

For the second lifting step we use scaling functions only to add 4 vanishing moments to the wavelet $\tilde{\psi}_2(x)$:

$$\tilde{\psi}_2^{new}(x) = \tilde{\psi}_2^{old}(x) + \frac{1}{16}\tilde{\varphi}_1(x) - \frac{9}{16}\tilde{\varphi}_2(x) - \frac{9}{16}\tilde{\varphi}_1(x-1) + \frac{1}{16}\tilde{\varphi}_2(x-1) \tag{20}$$

We cannot use $\tilde{\psi}_1^{new}$(x) in this step because we have already cancelled its first 4 moments, and we will obtain a singular system of equations to solve.

We add four vanishing moments to the primal wavelets using the primal scaling functions as follows:

$$\begin{aligned} \psi_1^{new}(x) &= \psi_1^{old}(x) + \tfrac{427499}{18951168}\varphi_2(x+1) - \tfrac{285847}{1018880}\varphi_1(x) - \\ &\qquad \tfrac{25339447}{94755840}\varphi_2(x) + \tfrac{802601}{31585280}\varphi_1(x-1) \\ \psi_2^{new}(x) &= \psi_2^{old}(x) + \tfrac{278449}{3740160}\varphi_1(x) - \tfrac{19774217}{78543360}\varphi_2(x) - \\ &\qquad \tfrac{1967613}{8727040}\varphi_1(x-1) + \tfrac{1090837}{15708672}\varphi_2(x-1) \end{aligned}$$

As we discuss below, there are important advantages to using scaling functions only for the lifting of the primal wavelets.

The result of our lifting is a biorthogonal basis that has 4 vanishing moments in both the primal and dual wavelets. Our use of both wavelets and scaling functions in the lifting steps has given us better control of the

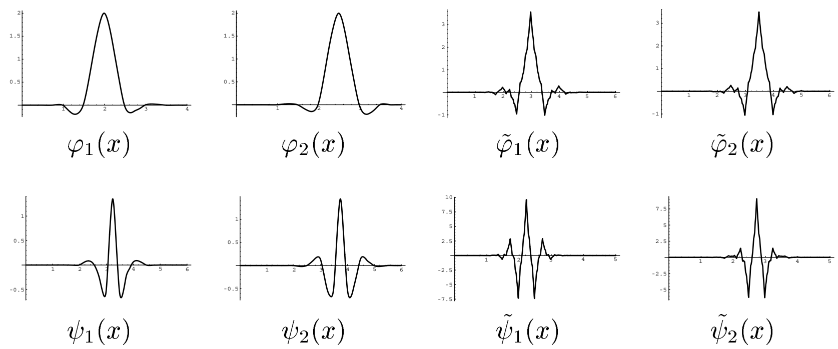

Figure 3: Biorthogonal scaling functions and wavelets constructed via lifting

tradeoff between analysis filter length and number of vanishing moments than in the wavelet case. The dilation coefficients of the constructed basis can be found at `http://www.mcs.drexel.edu/~vstrela`.

The basis we have constructed is very similar to the order 4 Deslauriers-Dubuc wavelet basis. The sole difference is that in the wavelet case the first lifting step uses only scaling functions to add vanishing moments to the dual wavelet. The result of our using the wavelets for the first lifting step is that we shorten one of the analysis filters. The filter for $\tilde{\psi}_1(x)$ has length 5 and that for $\tilde{\psi}_2(x)$ has length 7, whereas the analysis filter in the wavelet case has length 7. The shorter filter length leads to a comparable reduction in the length of the support of the wavelets. The length of the supports of $\tilde{\psi}_1(x)$ and $\tilde{\psi}_2(x)$ (estimated numerically) are 3.6 and 4.2, respectively, compared to the width 4.4 Deslauriers-Dubuc wavelet. Note, however, that the practical import of this reduction of support is tempered by the fact that most of the reduction in support takes place in a region where the wavelet $\tilde{\psi}$ is nearly zero. We obtain an alternative measure of the spread of the functions by the variance $\frac{1}{\|f\|^2}\int(x-\mu)^2|f(x)|^2dx$, where $\mu = \frac{1}{\|f\|^2}\int x|f|^2dx$. We find that our reduction in support comes at the price of a slight increase in the variance.

5.2 Prefiltering

In most applications one is given a set of function samples to transform rather than a set of scaling function coefficients as is assumed by the transform. The working assumption in the scalar wavelet case is usually that the scaling function coefficients are approximately equal to the given function samples. Such a map from function samples to scaling function coefficients is much less natural in the multiwavelet case. When the multiscaling functions differ significantly, a direct mapping of samples to coefficients introduces high-frequency artifacts. The problem of approximating multiscaling function coefficients from a given set of samples has been the object of a number of recent papers [32, 24, 13, 31, 29, 7, 1]. Lifting provides a way to construct families of multiwavelets for which such preprocessing reduces to a simple polyphase split.

Consider the scaling function coefficients for a smooth function $f(x)$. Expanding $f(x)$ into a Taylor series, we obtain

$$\int f(x)\tilde{\varphi}(x-k)dx = \int [f(k) + xf'(k) + \frac{1}{2!}x^2 f''(k) + \ldots]\varphi(x)dx \quad (21)$$

$$= \sum \frac{1}{j!} f^{(j)}(k) M_{\tilde{\varphi},j} \quad (22)$$

where $M_{\tilde{\varphi},j}$ is moment j of $\tilde{\varphi}$, given by $M_{\tilde{\varphi},j} = \int x^j \tilde{\varphi}(x)dx$. When $\tilde{\varphi}(x)$ has zero moments for $j = 1, \ldots, n$, we can approximate $\langle \tilde{\varphi}(x-k), f(x)\rangle$ by the sample $f(k)$ provided f is sufficiently smooth. Such a scaling function is said to have the *Coiflet property.*

The following lemma allows us to construct multiwavelets with scaling functions having the Coiflet property.

Lemma 6 *Let $\mathbf{\Phi}(x)$, $\mathbf{\Psi}(x)$, $\tilde{\mathbf{\Phi}}(x)$, and $\tilde{\mathbf{\Psi}}(x)$ be multiscaling and multiwavelet functions that give rise to a biorthogonal multiresolution analysis for L^2. Suppose the wavelets $\tilde{\mathbf{\Psi}}(x)$ have n vanishing moments. If we modify $\mathbf{\Psi}(x)$ via a lifting step that involves only translates of the scaling functions $\mathbf{\Phi}(x)$, we do not change the first n moments of the functions $\tilde{\mathbf{\Phi}}(x)$.*

Proof. Lifting $\mathbf{\Psi}$ using only translates of scaling functions results in a modified function of the form $\mathbf{\Psi}^{new}(x) = \mathbf{\Psi}(x) + \sum_k \mathbf{A}_k \mathbf{\Phi}(x-k)$. From the perfect reconstruction condition (10) it follows that $\tilde{\mathbf{H}}_k^{new} = \tilde{\mathbf{H}}_k + \sum_l \mathbf{A}_l^* \tilde{\mathbf{G}}_{k+2l}$. From Lemma 5 we have

$$\mathbf{M}_{\tilde{\Phi},p}^{new} = \frac{\sqrt{2}}{2} \sum_{j=0}^{p} \sum_k (\tilde{\mathbf{H}}_k + \sum_l \mathbf{A}_l \tilde{\mathbf{G}}_{k+2l}) \binom{p}{j} k^j \mathbf{M}_{\tilde{\Phi},p-j}^{new} \quad (23)$$

Suppose $p \leq n-1$. Using (17) we can express the terms in $\tilde{\mathbf{G}}_k$ as linear combinations of the moments $\mathbf{M}_{\tilde{\Psi},i}$, $i \leq j \leq p \leq n-1$. Because these moments are all zero, we have that $\mathbf{M}_{\Phi,p}^{new}$ and $\mathbf{M}_{\Phi,p}$ satisfy the same equations and hence must be equal.

6 The Two-Scale Similarity Transform

As we discussed in section 4.5, one of the important new freedoms we obtain in constructing multiwavelets rather than wavelets is that we can use scaling function coefficients to predict other scaling function coefficients. Unlike when we predict wavelet coefficients, the new scaling functions are *not* simply a linear combination of the old. The result of the new prediction is to modify the recurrence relations for the new scaling functions, and there is no simple way to express the effects of this modifiction.

We now consider the process of creating new scaling functions from a linear combination of the old ones. Let $\mathbf{\Phi}(x)$ be a scaling vector satisfying the recurrence relation (1). We want to examine the modified scaling vector $\mathbf{\Phi}^{new}(x) = \sum_k \boldsymbol{M}_k \mathbf{\Phi}(x-k)$. When $\boldsymbol{M}(z)$ is invertible for all $|z|=1$, the new scaling vector satisfies

$$\begin{aligned}\hat{\mathbf{\Phi}}^{new}(\omega) = \boldsymbol{M}(e^{i\omega})\hat{\mathbf{\Phi}}(\omega) = \tfrac{\sqrt{2}}{2}\boldsymbol{M}(e^{i\omega})\boldsymbol{H}(e^{i\frac{\omega}{2}})\hat{\mathbf{\Phi}}(\tfrac{\omega}{2}) = \\ \tfrac{\sqrt{2}}{2}\boldsymbol{M}(e^{i\omega})\boldsymbol{H}(e^{i\frac{\omega}{2}})\boldsymbol{M}^{-1}(e^{i\frac{\omega}{2}})\hat{\mathbf{\Phi}}^{new}(\tfrac{\omega}{2}).\end{aligned} \tag{24}$$

From this we obtain

$$\boldsymbol{H}^{new}(z) = \frac{1}{2}\boldsymbol{M}(z^2)\boldsymbol{H}(z)\boldsymbol{M}^{-1}(z). \tag{25}$$

This modification of $\boldsymbol{H}(z)$ is called a *two-scale similarity transform* (TST), and such transforms are studied in detail in [22, 19].

Matrices $\boldsymbol{M}(z)$ that are invertible for all $|z|=1$ do not change the approximation order of the scaling vector [22]. However, one can show if $\boldsymbol{M}(z)$ is singular in a special way, something more interesting happens: one can use a two-scale similarity transform to increase the approximation order of the scaling vector [17]. We emphasize that the scaling vectors that arise from this kind of degenerate TST do *not* correspond to linear combinations of the old functions, however. The result follows:

Theorem 7 *Suppose that the matrix symbol $\boldsymbol{H}_p(z)$ provides approximation order p, and that $\boldsymbol{H}_p(1)\boldsymbol{r}_p = \boldsymbol{r}_p$. Let $\boldsymbol{M}_p(z)$ be a matrix polynomial such that* $\det(\boldsymbol{M}_p(z)) = \text{const}\cdot(1-z)$ *and $\boldsymbol{M}_p(1)\boldsymbol{r}_p = \mathbf{0}$. Then the symbol $\boldsymbol{H}_{p+1}(z)$,*

$$\boldsymbol{H}_{p+1}(z) = \frac{1}{2}\boldsymbol{M}_p(z^2)\boldsymbol{H}_p(z)\boldsymbol{M}_p^{-1}(z) \tag{26}$$

provides approximation order $p+1$.

For scalar wavelets, the number of factors of $(1+z)$ in the symbol for the scaling function corresponds to the approximation order of the scaling function. The matrices $\boldsymbol{M}$ in the degenerate two-scale transforms are analogs of these factors of $(1+z)$.

Strela have developed an alternative method for constructing biorthogonal multiwavelet bases using degenerate two-scale similarity transforms [21]. In the TST construction method, approximation order conditions are automatically satisfied, but we must solve a system of equations to ensure that the perfect reconstruction condition is satisfied. In contrast, for each step of lifting, the perfect reconstruction condition is satisfied, but we must solve a system of equations to obtain a particular approximation order.

The two-scale similarity transform method generalizes a method for scalar wavelet construction described by Daubechies in [3]. Like lifting, two-scale similarity method is based on the factorization of a matrix polynomial. However, rather than focusing primarily on spatial-domain considerations, we instead focus on the distribution of a set of matrix factors related to the factors of $(1+z)$ in the scalar case. The basic idea is to construct an initial set of scaling functions with a high approximation order. We then find a set of dual scaling functions that satisfy the perfect reconstruction conditions. We next shift some of the approximation orders from the primal scaling functions to the duals via a procedure called *balancing*. The result is a primal and dual scaling vector with specified approximation orders. Finally, we find a set of primal and dual wavelets that satisfy the perfect reconstruction conditions.

Theorem 7 provides an easy way to construct a symbol with a given approximation order p:

$$\boldsymbol{H}_p(z) = \frac{1}{2^p}\boldsymbol{M}_p(z^2)\ldots\boldsymbol{M}_1(z^2)\boldsymbol{H}_0(z)\boldsymbol{M}_1^{-1}(z)\boldsymbol{M}_p^{-1}(z)$$

After this is done a dual symbol $\tilde{\boldsymbol{H}}_0(z)$ has to be found that satisfies the perfect reconstruction condition (10). We then perform balancing, transferring moments from the primal symbol to the dual in order to ensure the approximation properties of the dual basis. Finally, we solve for a set of primal and dual wavelet symbols that satisfy the perfect reconstruction condition.

Figure 4 shows the dual and primal scaling functions and wavelets constructed using a two-scale similarity transform method that parallels the construction of Section 5.1. The details may be found online at `http://www.mcs.drexel.edu/~vstrela`. Unlike the functions we constructed in our lifting example in Section 5.1, all the two-scale similarity transform functions are either symmetric or antisymmetric. Special symmetric lifting factors described in [26] would have had to be used to achieve similar symmetry properties by lifting. It can be proved that the Sobolev exponents for $\boldsymbol{\Phi}(x)$ and $\tilde{\boldsymbol{\Phi}}(x)$ are $s_{\max} = 1.1309$ and $s_{\max} = 3.5$, respectively. The obtained scaling vectors are slightly smoother than those

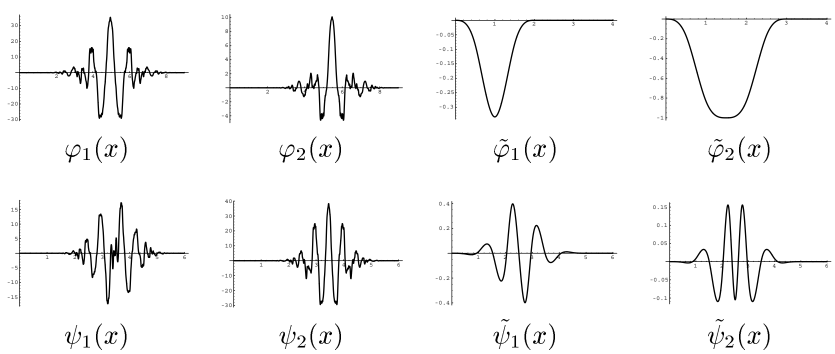

Figure 4: Biorthogonal scaling functions and wavelets constructed via two-scale similarity method

constructed in Section 5, which have primal and dual Sobolev exponents $s_{\max} = 1.1157$ and $s_{\max} = 2.7080$. The synthesis filter in the two-scale case is shorter than in the lifting case, but the analysis filter is longer.

Although the effects of lifting on the scaling vectors is more difficult to describe than in the two-scale case, the overall construction procedure is much simpler.

7 Conclusion

Lifting provides a simple way of constructing biorthogonal multiwavelet bases with a given approximation order. Equally importantly, it provides valuable intuition as to the source of the new degrees of freedom available in the construction process: with multiwavelets, we can use wavelet coefficients to predict other wavelet coefficients, and we can use scaling function coefficients to predict other scaling function coefficients. Moreover, as in the wavelet case, lifting steps provide a universal construction procedure for compactly supported multiwavelet bases.

References

[1] K. Attakitmongcol, D. Hardin, and D. Wilkes, Multiwavelet prefilters II: Optimal orthogonal prefilters, preprint, (1999).

[2] A. Cohen, I. Daubechies, and G. Plonka, Regularity of refinable function vectors, J. Fourier Anal. Appl. 3 (1997), 295-324.

[3] I. Daubechies, "Ten Lectures on Wavelets", SIAM, Philadelphia, 1992.

[4] I. Daubechies and W. Sweldens, Factoring wavelet transforms into lifting steps, J. Fourier Anal. Appl. 4 (1998).

[5] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, Fractal functions and wavelet expansions based on several scaling functions, J. Approx Theory 78 (1994), 373-401.

[6] S. S. Goh, Q. Jiang, and T. Xia, Construction of biorthogonal multiwavelets using the lifting scheme, preprint (1998).

[7] D. P. Hardin and D. W. Roach, Multiwavelet prefilters I: Orthogonal prefilters preserving approximation order $p \leq 2$, preprint (1997).

[8] C. Heil and D. Colella, Matrix refinement equations: Existence and uniqueness, J. Fourier Anal. Appl. 2 (1996), 363-377.

[9] Q. Jiang, On the regularity of matrix refinable functions, SIAM J. Math. Anal., to appear.

[10] Q. Jiang and Z. Shen, On existence and weak stability of matrix refinable functions, Constr. Approx., to appear.

[11] J. Kautsky and R. Turcajová, Adaptive wavelets for signal analysis, Lecture Notes in Computer Science 970 (1995), 906-911.

[12] F. Keinert, Raising multiwavelet approximation order through lifting, preprint (1999).

[13] J. M. Lebrun and M. Vetterli, Balanced Multiwavelet Theory and Design, IEEE Trans. SP 46 (1998), 1119–1125.

[14] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, Trans. PAMI 11 (1989), 674-693.

[15] J. Marasovich, "Biorthogonal Multiwavelets", Ph.D. thesis, Mathematics Department, Vanderbilt University, 1996.

[16] G. Plonka, Approximation order provided by refinable function vectors, Constr. Approx. 13 (1997), 221-244.

[17] G. Plonka and V. Strela, Construction of multi-scaling functions with approximation and symmetry, SIAM J. of Math. Anal. 29 (1998), 481-510.

[18] Z. Shen, Refinable function vectors, SIAM J. Math. Anal., to appear.

[19] G. Strang, V. Strela, and D.-X. Zhou, compactly supported refinable functions with infinite masks, in "Contemporary Mathematics" (L. Baggett and D. Larson, Eds.), AMS, 1999.

[20] G. Strang and T. Nguyen, "Wavelets and Filter Banks", Cambridge-Wellesley Press, 1996.

[21] V. Strela, A note on Construction of biorthogonal multi-scaling functions, in "Contemporary Mathematics" (A. Aldoubi and E. B. Lin, Eds.), pp., 149-157 AMS, 1998.

[22] V. Strela, Multiwavelets: Regularity, orthogonality and symmetry via two-scale similarity transform, Studies in Appl. Math. 98 (1997), 335-354.

[23] V. Strela, "Multiwavelets: Theory and Applications", PhD thesis, MIT, 1996.

[24] V. Strela and A. T. Walden, Signal and image denoising via wavelet thresholding: Orthogonal and biorthogonal, scalar and multiple wavelet transforms, Imperial College, Statistics Section, Technical Report TR-98-01 (1998).

[25] W. Sweldens, The lifting scheme: acustom-design construction of biorthogonal wavelets, Appl. Comput. Harmonic Anal. 3 (1996), 186-200.

[26] R. Turcajová , Construction of symmetric biorthogonal multiwavelets by lifting, Proc. SPIE 3813 (1999).

[27] R. Turcajová and V. Strela, Two ways of constructing Hermite spline multiwavelets, in preparation.

[28] P. P. Vaidyanathan, "Multirate Systems and Filter Banks", Prentice Hall, 1993.

[29] M. J. Vrehl and A. Aldroubi, Projection based prefiltering for multiwavelet transform, IEEE Trans. SP 46 (1998), 3088–3092.

[30] M. Vetterli and C. Herley, Wavelets and filter banks: Theory and design, IEEE Trans. Acoust. Speech Signal Process. 40 (1992), 2207–2232.

[31] X.-G. Xia, A new prefilter design for discrete multiwavelet transform, IEEE Trans. SP 46 (1998), 1558–1570.

[32] X.-G. Xia, J. Geronimo, D. Hardin, and B. Suter, Design of prefilters for discrete multiwavelet transforms, IEEE Trans. SP 44 (1996), 25–35.