The Design of Matched Balanced Orthogonal Multiwavelets

Joël M.H. Karel[1], Sjoerd van Steenkiste[2] and Ralf L.M. Peeters[1]*

[1]Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, Netherlands, [2]Dalle Molle Institute for Artificial Intelligence, Scuola Universitaria Professionale Della Svizzera and Università Della Svizzera Italiana, Lugano-Viganello, Switzerland

The theory of orthogonal multiwavelets offers enhanced flexibility for signal processing applications and analysis by employing multiple waveforms simultaneously, rather than a single one. When implementing them with polyphase filter banks, it has been recognized that balanced vanishing moments are needed to prevent undesirable artifacts to occur, which otherwise compromise the interpretation and usefulness of the multiwavelet analysis. In the literature, several such balanced orthogonal multiwavelets have been constructed and published; but however useful, their choice is still limited. In this work we present a full parameterization of the space of all orthogonal multiwavelets with two balanced vanishing moments (of orders 0 and 1), for arbitrary given multiplicity and degree of the polyphase filter. This allows one to search for matching multiwavelets for a given application, by optimizing a suitable design criterion. We present such a criterion, which is sparsity-based and useful for detection purposes, which we illustrate with an example from electrocardiographic signal analysis. We also present explicit conditions to build in a third balanced vanishing moment (of order 2), which can be used as a constraint together with the earlier parameterization. This is demonstrated by constructing a balanced orthogonal multiwavelet of multiplicity three, having three balanced vanishing moments, but this approach can easily be employed for arbitrary multiplicity.

Keywords: wavelet theory, orthogonal multiwavelets, balanced vanishing moments, matched wavelets, sparsity, parameterization, lossless polyphase filters

INTRODUCTION

Wavelets [1, 2] are a popular signal processing tool, able to provide a time-frequency representation of signals. Though there are various viewpoints on wavelets, here we will restrict ourselves to wavelets from filter banks. Within this class, various desirable properties are possible to achieve, e.g. orthogonality, linear phase, compact support, symmetry and vanishing moments. We will restrict ourselves to orthogonal wavelets with compact support [3]. Vanishing moments induce a degree of smoothness and allow the wavelet transform to be interpreted as a multiscale differential operator, which allows one to measure the regularity of a signal [4]. The orthogonality property does not combine well with the other properties of scalar wavelets; for instance there is no compactly supported orthogonal wavelet other than the Haar wavelet that is also symmetric.

In the orthogonal scalar case, if one uses all free parameters to build in as many vanishing moments as possible, the Daubechies wavelets are obtained. If instead only a limited number of vanishing moments is required, there is freedom left for other properties. One can use this freedom also to design matched wavelets for an application, e.g., to promote a sparse representation of a specific prototype signal. In [5] a parameterization was developed for orthogonal scalar wavelets,

which were matched to a prototype signal. This approach was further expanded in [6, 7] by using a parameterization based on lossless systems. In that parameterization, the polyphase filter associated with the orthogonal wavelet is recursively constructed as the transfer matrix of a lossless system using Schur interpolation theory [8] with rotation matrices and elementary delay operators.

Multiwavelets [9–12] are a generalisation of scalar wavelets, and consist of a tuple of r scalar wavelets. Multiwavelets are more flexible, and can combine properties such as compact support, orthogonality and symmetry. This added flexibility is for example advantageous in multiwavelet denoising, which has been applied in rolling bearing fault detection [13–16], and in the load spectrum of computer numerical control lathe [17]. Recently, the correspondence between multiwavelet shrinkage and nonlinear diffusion was studied [18]. *We will be addressing orthogonal multiwavelets with compact support generated from filter banks with a complexity (filter order) that can be chosen by the user.* Much of the theory of orthogonal scalar wavelets carries over to orthogonal multiwavelets, but some subtle differences are encountered.

An important difference arises when processing a one-dimensional signal with a multiwavelet. Since the filter bank associated with the multiwavelet is a multi-input multi-output system (MIMO), vectorisation of the input signal is required. A natural vectorisation is obtained by decomposing the signal into its phases. However, arbitrary multiwavelet filters do not preserve this structure throughout the filtering operation. As a result the output channels of the low-pass filter become unbalanced. Moreover, the multiwavelet filter does not guarantee the preservation of polynomial signals by the low-pass filter, even when an appropriate vanishing moment is in place. This further complicates a direct interpretation of the multiwavelet analysis. This "balancing problem" [9, 12] is caused by the different spectral behaviour of the components of the scaling function that together with the wavelet function define the multiwavelet. Lack of balancing has hampered the use of multiwavelets by the signal processing community, despite the widely recognized potential. An option to overcome this is to first reconstruct a signal from the low-pass outputs, split it into phases, and feed those to the filter bank, but this is an unattractive approach, since it deteriorates computational performance especially with an increasing number of scales.

A more attractive solution is to impose a balancing condition on the scaling function, to complement the vanishing moment condition on the wavelet function. In the literature, there are constructions of balanced multiwavelets [19, 20] and of balanced multiframelets [21, 22]. However, enforcing such a balancing condition has turned out to become increasingly hard when the order of the vanishing moment grows larger.

The problem of balancing was first formally addressed by Lebrun and Vetterli in [23]. In [24] necessary and sufficient conditions were provided for a multiwavelet to be p-order balanced. These conditions were formalised in [9]. In [11] necessary and sufficient conditions on the zeros of certain filters associated with p-order balanced multiwavelets were developed. Both authors include examples of multiwavelets

that are balanced up to order two or three[1] for multiplicity $r = 2$. These multiwavelets were obtained by solving systems of nonlinear polynomial equations using Gröbner bases.

The balancing conditions given in [11, 24] are highly non-linear, and are difficult to satisfy. These conditions were further characterised in [12]. In [25] the construction of balanced multiwavelets is simplified by using the lifting scheme. In [26] balanced multiwavelets with interpolatory property are discussed, for multiscaling functions of multiplicity $r = 2$.

From the literature it is found that a parameterization for the construction of multiwavelets balanced up to order p with an arbitrary multiplicity r is largely lacking. In [27, 28] a parameterization for the construction of zero-order balanced multiwavelets is derived. This was a large step forward in the construction of balanced multiwavelets, and allows for balanced multiwavelet design. A set of directly applicable filter conditions for the construction of balanced multiwavelets for order $p > 0$ and an arbitrary multiplicity r, is not available in the literature.

In this paper we will develop such a parameterization for orthogonal multiwavelets with compact support based on lossless systems, with an arbitrary multiplicity r. To this end, we will first introduce multiwavelets from filter banks in **Section 2.1**. Next we will discuss parameterizations of multiwavelets as lossless FIR polyphase filters in **Section 2.2**. Then the balancing concept is investigated from a signal processing viewpoint in **Sections 2.3–2.5**, and balancing up to and including order 1 is built directly into the parameterization in **Sections 2.6–2.7**. For order 2, an additional condition is provided in **Section 2.8** in a form which can be used for numerical optimization. The parameterizations describe the free parameters in a form that is suitable to match a multiwavelet filter bank to a prototype signal. For this matching, it is argued in **Section 2.9** that e.g., L_1-norm minimization or L_4-norm maximization can be used, as in [5–7], exploiting conservation of energy due to orthogonality. The effects of the balancing issue is illustrated in **Section 3.1**. Two multiwavelet design examples are provided in **Sections 3.2, 3.3**, to illustrate the approach and techniques discussed. We will show in this paper how all balanced orthogonal multiwavelets of orders 0 and 1 can be obtained for arbitrary multiplicity r and any given polyphase filter order (McMillan degree), which we consider a major step forward in making multiwavelets applicable.

MATERIALS AND METHODS

Multiwavelets and Multiwavelet Filter Banks

In this section we briefly review the theory of multiwavelets and multiwavelet filter banks, to a large extent based on the work in [9, 10]. We will restrict to orthogonal multiwavelets having compact support. Compact support translates into finite impulse response (FIR) filters, whereas orthogonality is captured conveniently when switching from a filter bank

---

[1]In those papers, p-order balancing is defined to correspond to p balanced vanishing moments, of orders 0, 1, ..., $p - 1$. In this paper we will call this 'balanced up to order $p - 1$.'

description to polyphase filtering: it corresponds to the polyphase FIR filter being para-unitary, i.e. lossless. For lossless filters parameterizations are available in the literature, which offer opportunities to build in additional properties. We will exploit these when addressing vanishing moments and balancing conditions.

A multiresolution structure for the space $L^2(\mathbb{R})$ is defined, as usual in wavelet theory, to consist of nested approximation spaces V_m, which are linear subspaces of $L^2(\mathbb{R})$ such that $\ldots, \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots$, with intersection $\cap_{m\in\mathbb{Z}} V_m = \{0\}$ and with completeness $\overline{\cup_{m\in\mathbb{Z}} V_m} = L^2(\mathbb{R})$. For each $m \in \mathbb{Z}$ we define the detail space W_m to be such that $V_{m+1} = V_m \oplus W_m$ and $V_m \perp W_m$. Thus, V_m and W_m are orthogonal complements within the enveloping space V_{m+1}. Furthermore, we assume that this structure is both shift-invariant ($\forall f(t) \in L^2(\mathbb{R}), k, m \in \mathbb{Z}$: $f(t) \in V_m \Leftrightarrow f(t-k) \in V_m$) and scale-invariant ($\forall f(t) \in L^2(\mathbb{R}), m \in \mathbb{Z}$: $f(t) \in V_m \Leftrightarrow f(2t) \in V_{m+1}$).

In orthogonal multiwavelet theory, it is assumed next that there exists a *multiscaling function*. This is a vector of r scaling functions $\Phi(t) = (\phi_{[0]}(t), \phi_{[1]}(t), \ldots, \phi_{[r-1]}(t))^T$ of which the entries, together with their integer translates, generate an orthonormal basis of the approximation space V_0. Thus, for each such multiscaling function we have that $\langle \phi_{[j]}(t-k), \phi_{[p]}(t-\ell) \rangle = \delta_{j,p}\delta_{k,\ell}$ for all integers $k, \ell \in \mathbb{Z}$ and all indices $j, p \in \{0, 1, \ldots, r-1\}$. Here $\delta_{j,k}$ is the Kronecker delta, equal to 1 if $j = k$ and equal to 0 otherwise. In this set-up, the orthonormal basis of V_0 is generated by the r scaling functions jointly rather than by any single one of them. Because of shift and scale-invariance of the multiresolution structure, the orthonormal basis of V_0, for each m, induces an orthonormal basis of V_m which is constituted by the entries of the vectors in $\{2^{m/2}\Phi(2^m t - k)|k \in \mathbb{Z}\}$.

Likewise, it is assumed that there exists an associated *multiwavelet function* $\Psi(t)$ which together with its integer translates generates an orthonormal basis of the detail space W_0. This multiwavelet is a vector of r wavelet functions $\Psi(t) = (\psi_{[0]}(t), \psi_{[1]}(t), \ldots, \psi_{[r-1]}(t))^T$, with the property $\langle \psi_{[j]}(t-k), \psi_{[p]}(t-\ell) \rangle = \delta_{j,p}\delta_{k,\ell}$ for all $k, \ell \in \mathbb{Z}$ and all $j, p \in \{0, 1, \ldots, r-1\}$. Since $V_0 \perp W_0$ it holds that $\langle \phi_{[j]}(t-k), \psi_{[p]}(t-\ell) \rangle = 0$, for all $j, p, k,$ and ℓ. As before for the spaces V_0 and V_m, the orthonormal basis of W_0 induces an orthonormal basis of W_m which consists of the entries of the vectors in $\{2^{m/2}\Psi(2^m t - k)|k \in \mathbb{Z}\}$.

From the fact that $\Phi(t) \in V_0^r$ and $V_0 \subset V_1$, it follows that $\Phi(t)$ can be represented in terms of the induced basis $\{\sqrt{2}\,\Phi(2t-k)|k \in \mathbb{Z}\}$ of V_1. So, there exist unique $r \times r$ matrix coefficients $C_k, k \in \mathbb{Z}$, for which the multiscaling function $\Phi(t)$ satisfies a two-scale vector equation (a refinement equation) called the *dilation equation*:

$$\Phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} C_k \Phi(2t+k). \tag{1}$$

Likewise, we have that $\Psi(t) \in W_0^r$ and $W_0 \subset V_1$, so it can also be represented uniquely in terms of the same induced basis of V_1. This implies that there exist unique $r \times r$ matrix coefficients D_k, $k \in \mathbb{Z}$, which express the multiwavelet $\Psi(t)$ in terms of the

multiscaling function on V_1 by means of a two-scale vector equation called the *wavelet equation*:

$$\Psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} D_k \Phi(2t+k). \tag{2}$$

If we also impose that the multiscaling and multiwavelet functions have *compact support* (i.e., they are nonzero only on a finite interval, which makes them localized in space—a desirable property of wavelets), it follows that only a finite number of basis functions in the induced basis of V_1 can contribute to representing $\Phi(t)$ and $\Psi(t)$. As a consequence, the infinite sums in the dilation and wavelet equation become finite sums, and by shifting $\Phi(t)$ and $\Psi(t)$ appropriately, if necessary, it can be assumed without loss of generality that the index k in each sum runs from $k = 0$ to $k = 2n - 1$ for some integer $n \geq 1$.[2]

The finite matrix coefficient sequences $\{C_k|k = 0, 1, \ldots 2n - 1\}$ and $\{D_k|k = 0, 1, \ldots 2n - 1\}$ are used to define two $r \times r$ polynomial matrices in z^{-1} according to $C(z) = \sum_{k=0}^{2n-1} C_k z^{-k}$ and $D(z) = \sum_{k=0}^{2n-1} D_k z^{-k}$. The function $C(z)$ is the transfer function of a FIR low-pass multiwavelet filter, whereas $D(z)$ is the transfer function of a FIR high-pass multiwavelet filter. Together they make up an orthogonal FIR filter bank, satisfying the Smith-Barnwell orthogonality conditions [29]:

$$C(z)C(z^{-1})^T + C(-z)C(-z^{-1})^T = 2I_r, \tag{3}$$

$$D(z)D(z^{-1})^T + D(-z)D(-z^{-1})^T = 2I_r, \tag{4}$$

$$C(z)D(z^{-1})^T + C(-z)D(-z^{-1})^T = 0. \tag{5}$$

The converse of what has been presented so far, is also largely true: if one starts from an orthogonal FIR filter bank with $r \times r$ FIR filters $C(z)$ and $D(z)$ satisfying **Eqs 3–5**, then this induces a multiresolution structure with orthonormal translation invariant bases generated by the multiscaling and multiwavelet functions satisfying the dilation equation and wavelet equation under relatively mild conditions. What is needed, is what we have assumed earlier, namely that the dilation equation admits a solution and that the spaces eventually span all of $L^2(\mathbb{R})$.

The *cascade algorithm* is a tool to compute $\Phi(t)$ for a given choice of filter coefficients $C_k, k = 0, 1, \ldots, 2n - 1$ by iteration: from a current estimate of $\Phi(t)$ a new estimate is computed by substituting the current estimate into the right-hand side of the dilation equation and reading off the new estimate from the resulting left-hand side. This requires a suitable initialisation,

---

[2]Note that, unlike what is common in most of the literature, we have chosen in our definition of the dilation and wavelet equation above to use functions $\Phi(2t + k)$ rather than $\Phi(2t - k)$, in line with the convention in [7]. This causes no loss of generality, but makes that the multiscaling and multiwavelet functions in $\Phi(t)$ and $\Psi(t)$ are all compactly supported on the *negative* interval $(-(2n - 1), 0]$ as can be quickly derived from (1) and (2). This has the two-fold advantage that we can save on the notation and the number of filters required in the rest of this work, as well as that we can work with causal filters and avoid unnecessary delay or advance operators when switching to a signal processing perspective. Should one switch to using functions $\Phi(2t - k)$ on the right-hand side of the dilation and wavelet equations, then the time-reversed functions $\Phi(-t)$ and $\Psi(-t)$ are the new solutions in the adapted setting.

which is achieved by the Haar multiwavelet, for which (consistent with our convention) the multiscaling function $\Phi(t)$ is given by $\phi_{[i]}(t) = \sqrt{r}$ on the interval $(-\frac{i+1}{r}, -\frac{i}{r}]$ and $\phi_{[i]}(t) = 0$ elsewhere. Clearly, this Haar multiwavelet is an example of an orthogonal multiwavelet (albeit not a spectacular one, as it simply mimics the scalar Haar wavelet). It is easily verified that orthonormality of the induced basis is preserved in each iteration step and that compact support holds too due to the finite number of terms in the summation; so if the cascade algorithm converges it will produce a feasible solution for $\Phi(t)$. If the algorithm happens to not converge, then a multiwavelet interpretation is lacking, but the filter bank may still prove valuable from a signal processing perspective. The cascade algorithm is useful to visualize the multiwavelets which correspond to a chosen filter bank. When discussing balancing conditions, it is instrumental for determining the vector v_0 in **Eq. 39**.

Assume a function (a scalar signal) $s(t) \in V_{m+1}$ to be represented by a sequence of $r \times 1$ vector coefficients $\{s_\ell\}$ in terms of the induced basis of V_{m+1}:

$$s(t) = \sum_{\ell \in \mathbb{Z}} 2^{(m+1)/2} s_\ell^T \Phi(2^{m+1}t - \ell). \tag{6}$$

Since $V_{m+1} = V_m \oplus W_m$ we can decompose $s(t)$ into

$$s(t) = a(t) + b(t), \tag{7}$$

in which $a(t)$ is an approximation signal with coefficient vectors $\{a_\ell\}$ with respect to the induced multiscaling basis in V_m, and $b(t)$ is a detail signal (orthogonal to $a(t)$) with coefficient vectors $\{b_\ell\}$ with respect to the induced multiwavelet basis in W_m:

$$a(t) = \sum_{\ell \in \mathbb{Z}} 2^{m/2} a_\ell^T \Phi(2^m t - \ell), \tag{8}$$

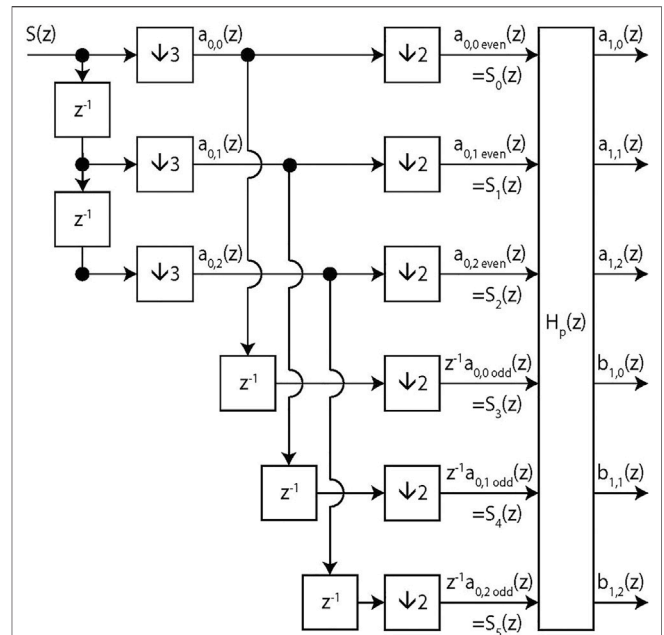$$b(t) = \sum_{\ell \in \mathbb{Z}} 2^{m/2} b_\ell^T \Psi(2^m t - \ell). \tag{9}$$

It now holds that these coefficient vectors can be computed from those of $s(t)$ as follows:

$$a_\ell = \sum_{k=0}^{2n-1} C_k s_{2\ell-k}, \tag{10}$$

$$b_\ell = \sum_{k=0}^{2n-1} D_k s_{2\ell-k}. \tag{11}$$

In digital signal processing terms, the sequences $\{a_\ell\}$ and $\{b_\ell\}$ are obtained from $\{s_\ell\}$ by first filtering $\{s_\ell\}$ with the filters $C(z)$ and $D(z)$ (in parallel), and then dyadically down-sampling the resulting sequences: only the even-indexed coefficient vectors are kept (and relabeled), all the odd-indexed coefficient vectors are discarded.[3]

This process can be reorganized in a more efficient way, by first splitting the coefficient vector sequence $\{s_\ell\}$ into two phases (using down-sampling on the sequence and on a delayed duplicate) and then passing both phases jointly through a

---

[3]The convolution sums in **Eqs 10–11**, involving $s_{2\ell-k}$ rather than $s_{2\ell+k}$, appear because of the convention we adopted for the dilation and wavelet equation.

FIGURE 1 | A single scale of a polyphase wavelet transform with $r = 3$ and a univariate input signal $S(z)$ that is split into $2r$ phases, yielding r approximation components $a_{1,\ell}$ and r detail components $b_{1,\ell}$.

suitable polyphase filter $H_p(z)$ to directly produce $\{a_\ell\}$ and $\{b_\ell\}$. More generally, if we start from a discrete-time representation of the scalar signal $s(t)$ as a sequence of scalar values $\{f_\ell\}$ (for instance obtained by regular sampling) and corresponding z-transform $S(z) = \sum_{\ell \in \mathbb{Z}} f_\ell z^{-\ell}$, then one can split $\{f_\ell\}$ into $2r$ phases and submit these jointly to the $2r \times 2r$ polyphase filter $H_p(z)$. These $2r$ phases are defined by rewriting $S(z)$ as:

$$S(z) = S_0(z^{2r}) + z^1 S_1(z^{2r}) + \cdots + z^{2r-1} S_{2r-1}(z^{2r}), \tag{12}$$

where $S_k(z)$ denotes the z-transform of the k-th phase of $\{f_\ell\}$ ($k = 0, 1, \ldots, 2r - 1$). This matches the situation with a coefficient vector sequence $\{s_\ell\}$ if each s_ℓ contains the phases $0, 1, \ldots, r - 1$ of the scalar coefficient sequence $\{f_\ell\}$ for even indices ℓ and the phases $r, r + 1, \ldots, 2r - 1$ for odd indices ℓ.

The $2r \times 2r$ FIR polyphase filter $H_p(z)$ is constructed as:

$$H_p(z) = \begin{pmatrix} C_{\text{even}}(z) & C_{\text{odd}}(z) \\ D_{\text{even}}(z) & D_{\text{odd}}(z) \end{pmatrix} =$$

$$\begin{pmatrix} c_{0,0}(z) & c_{0,2}(z) & \cdots & c_{0,2r-2}(z) & c_{0,1}(z) & c_{0,3}(z) & \cdots & c_{0,2r-1}(z) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{r-1,0}(z) & c_{r-1,2}(z) & \cdots & c_{r-1,2r-2}(z) & c_{r-1,1}(z) & c_{r-1,3}(z) & \cdots & c_{r-1,2r-1}(z) \\ d_{0,0}(z) & d_{0,2}(z) & \cdots & d_{0,2r-2}(z) & d_{0,1}(z) & d_{0,3}(z) & \cdots & d_{0,2r-1}(z) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{r-1,0}(z) & d_{r-1,2}(z) & \cdots & d_{r-1,2r-2}(z) & d_{r-1,1}(z) & d_{r-1,3}(z) & \cdots & d_{r-1,2r-1}(z) \end{pmatrix} \tag{13}$$

where $C_{\text{even}}(z)$, $C_{\text{odd}}(z)$, $D_{\text{even}}(z)$, and $D_{\text{odd}}(z)$ are $r \times r$ polynomial matrices (in z^{-1}) that split $C(z)$ and $D(z)$ into two phases, such that

$$C(z) = C_{\text{even}}(z^2) + z^{-1} C_{\text{odd}}(z^2) \tag{14}$$

$$D(z) = D_{\text{even}}(z^2) + z^{-1} D_{\text{odd}}(z^2). \tag{15}$$

This architecture is illustrated in **Figure 1** and leads to the same approximation coefficients a_ℓ and detail coefficients b_ℓ as

before, where we used double subscripts to anticipate a multiresolution structure with repeated down-sampling and filtering steps later on.

For the $2r \times 2r$ polyphase filter $H_p(z)$ defined in **Eq. 13** the orthogonality conditions of **Eqs 3–5** translate into the equivalent condition:

$$H_p(z)H_p(z^{-1})^T = I_{2r}. \tag{16}$$

This implies that $H_p(z)$ is a lossless FIR filter, with coefficient matrices $H_k = \begin{pmatrix} C_{2k} & C_{2k+1} \\ D_{2k} & D_{2k+1} \end{pmatrix}$, for $k = 0, 1, \ldots, n - 1$.

Complementary to **Eq. 13**, which expresses $H_p(z)$ in terms of $C(z)$ and $D(z)$, the low-pass and high-pass FIR filters $C(z)$ and $D(z)$ are reobtained from the polyphase filter $H_p(z)$ by:

$$\begin{pmatrix} C(z) \\ D(z) \end{pmatrix} = H_p(z^2) \begin{pmatrix} I_r \\ z^{-1}I_r \end{pmatrix}. \tag{17}$$

Parameterizing Lossless FIR Polyphase Filters

Lossless polyphase filters have a rich structure and have been well studied in the literature. See for instance [30, 31] for an accessible review of lossless filters and their key properties, including application areas and examples of the various uses they have. The basic property of a discrete-time lossless system, from which its name derives, is that the total energy of the output signals equals the total energy of the input signals, regardless of the signals being used. Here, energy is measured by the sum of squares of the values in the (discrete-time) signals. For $H_p(z)$ this precise property is captured by **Eq. 16**. The point is that for $|z| = 1$ it holds that $z^{-1} = z^*$ (i.e., the complex conjugate) and since $H_p(z)$ has real coefficients it follows that $H_p(z^{-1})^T = H_p(z)^*$. This shows that $H_p(z)$ is unitary for all z on the unit circle, which causes the conservation of energy property.

For the construction of arbitrary lossless FIR polyphase filters (for given choices of r and n), the condition of **Eq. 16** is not convenient to work with and impose directly. Equating coefficients for all the entries of the corresponding expressions on the left and on the right, can of course be done but does not provide orthogonality conditions in a form suitable for analytic or numerical computation. Instead, there is a body of literature which describes how the class of all $2r \times 2r$ lossless systems of a fixed given order can be parameterized, by carrying out a recursion with respect to the system order, known as the tangential Schur algorithm; see [8]. This approach is flexible, as it allows the user to build in certain properties: for instance, by making specific choices for some parameters in the general procedure it allows one to parameterize the subclass of lossless FIR filters.

The following theorem was used in [27, 28] to parameterize (real) lossless FIR polyphase filters of a given order; see also [10, 31] for a similar yet slightly different construction. Here, the order k of a lossless FIR polyphase filter $G^{(k)}(z)$ is the *McMillan degree* of this rational matrix. For lossless functions this equals the degree of the denominator of the rational function $\det(G^{(k)}(z))$. For the lossless FIR polyphase filter $H_p(z)$, the order $n - 1$ traditionally is the *maximum lag* of the filter, which relates to the length of the interval on which the multiwavelet lives. In the scalar case $r = 1$ these definitions coincide. In the multiwavelet case $r \geq 2$ they do not.

Generically, the McMillan degree and the maximum lag are the same for the lossless FIR polyphase filters in our class, but for special choices of the unit vectors u_k they may be different: the maximum lag of $H_p(z)$ is less than its McMillan degree $n - 1$ if and only if $u_{k+1}^T u_k = 0$ for some $k \in \{1, \ldots, n - 2\}$. We prefer to adopt the McMillan degree definition of the filter order when addressing parameterizations, to avoid having to consider special cases resulting from this mismatch.[4]

Theorem 3.1. Let $G^{(k)}(z)$ be a real $2r \times 2r$ lossless FIR polyphase filter of order $k \geq 1$. Then there exists a real $2r \times 1$ vector u_k of norm $\|u_k\| = 1$ such that $G^{(k)}(z)$ can be factored as:

$$G^{(k)}(z) = (I_{2r} + (z^{-1} - 1)u_k u_k^T)G^{(k-1)}(z), \tag{18}$$

where $G^{(k-1)}(z)$ is a real $2r \times 2r$ lossless FIR polyphase filter of order $k - 1$.

This theorem shows that a lossless FIR polyphase filter of order k can be reduced in k iteration steps to a lossless FIR polyphase filter $G^{(0)}(z)$ of order 0, which is nothing else than a constant orthogonal matrix. We therefore drop the variable z for this last filter and simply write $G^{(0)}$.

For the polyphase matrix $H_p(z)$ of the filter bank, the recursion makes clear that $G^{(0)}$ is in fact equal to the value of $H_p(z)$ at $z = 1$: $G^{(0)} = H_p(1)$. This will make it more convenient to build in balanced vanishing moments than for the construction described in [10]. Generically, there will exist unit vectors $u_1, \ldots, u_{n-1}$ which allow $H_p(z)$ to be factored as:

$$H_p(z) = (I_{2r} + (z^{-1} - 1)u_{n-1}u_{n-1}^T) \ldots (I_{2r} + (z^{-1} - 1)u_1 u_1^T)G^{(0)}. \tag{19}$$

To make the parameterization explicit, all the vectors u_k of norm 1 as well as the orthogonal matrix $G^{(0)}$ still can be parameterized in terms of scalar parameters. This is easily achieved. A $2r \times 1$ unit vector u_k can be recursively parameterized as $u_k = \begin{pmatrix} \cos(\theta_{k,1}) \\ g_k \sin(\theta_{k,1}) \end{pmatrix}$ with $\theta_{k,1} \in [0, \pi]$ and g_k again a unit vector but of smaller size $(2r - 1) \times 1$. For orthogonal matrices it is well-known how they can be parameterized, for instance, with Givens rotations and Householder matrices.[5] When addressing balanced vanishing moments of several orders, this parameterization will be refined further.

---

[4]To compute the McMillan degree of a lossless FIR polyphase filter $H_p(z) = \sum_{k=0}^{n-1}H_k z^{-k}$ with maximum lag $n - 1$, it is convenient to construct the block-Hankel matrix $\mathcal{H} = \begin{pmatrix} H_1 & H_2 & \ldots & H_{n-2} & H_{n-1} \\ H_2 & H_3 & \ldots & H_{n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ H_{n-2} & H_{n-1} & & & 0 \\ H_{n-1} & 0 & \ldots & 0 & 0 \end{pmatrix}$. The McMillan degree m is equal to $\operatorname{rank}(\mathcal{H})$. The SVD is a particularly suitable tool to determine m numerically, because it holds for a lossless system that all the m non-zero singular values of $\mathcal{H}$ are equal to 1, see [36].

[5]Givens rotations and Householder reflections are commonly used to perform QR-decomposition. When applied to an orthogonal matrix, the upper triangular R is in fact diagonal with all entries on its main diagonal equal to ± 1. The space of orthogonal matrices has two connected components, characterized by the sign of the determinant ± 1, which is something to take into account when parameterizing this space. Givens rotations all have determinant 1 and are used to create zeros one by one; they involve a single angular parameter. Householder reflections all have determinant -1 and are used to create multiple zeros at once; they involve a vector of norm 1 requiring multiple angular parameters. See [37]. We will introduce Householder matrices in **Eq. 45**.
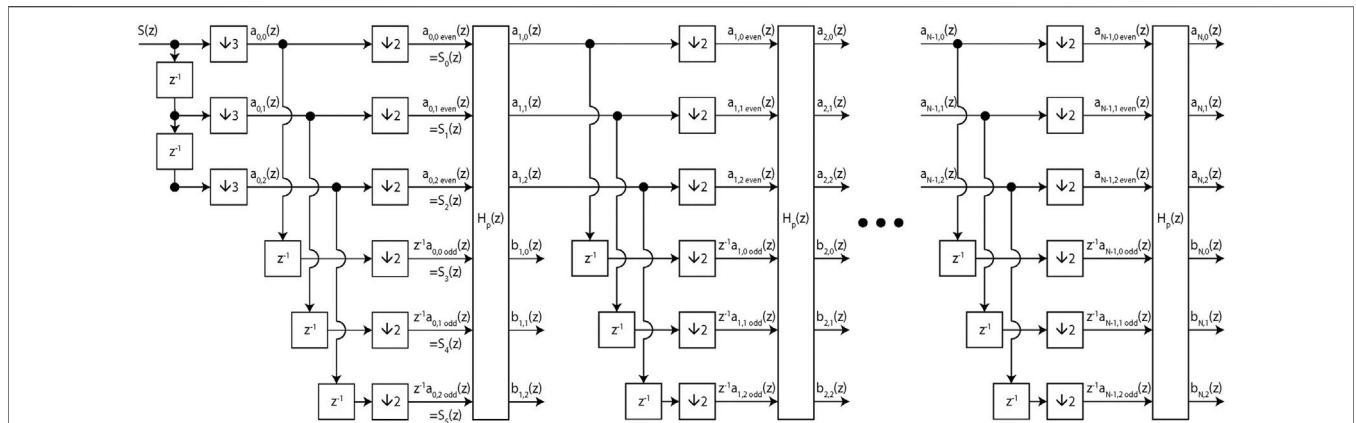
FIGURE 2 | Multiresolution structure of a multiwavelet filter bank with $r = 3$ on a univariate input signal $S(z)$ and N wavelet scales.

The Balancing Problem

When processing a signal with a multiwavelet filter bank, the differences in spectral behavior of the r components of the multiscaling function may result in unbalanced channels of the low-pass filter. The consequences will further escalate in a multiresolution structure, when the filter bank is repeatedly applied to part of the output (the approximation signals) of a previous filtering step. The imbalance will adversely affect performance in compression and detection applications in signal and image processing, due to the fact that high and low frequencies start to mix. This problem is known as the balancing problem and was pointed out and characterized in detail in [9, 12, 24, 32].

The balancing problem is first encountered when processing a single scale of a multiwavelet filter bank with a sampled univariate signal as in **Figure 1**. Assume that the sampled input signal with z-transform $S(z)$ is constant, then all down-sampled signals with z-transforms $S_k(z)$ will be constant too and in fact be equal. A constant signal is supposed to pass unaltered through the polyphase filter, as the multiwavelet function $\Psi(t)$ should pick up details and the multiscaling function $\Phi(t)$ the trend (an approximation to the signal); indeed, a constant function is all about trends and has no details. This can be achieved by imposing a *zero-order vanishing moment* on the filter bank, as is commonly done. This is the multiwavelet counterpart to the admissibility condition for scalar wavelets: $\int_{\mathbb{R}} \Psi(t)dt = 0$.

However, though each of the low-pass outputs $a_{1,\ell}(z)$ will then be constant (due to the vanishing moment), their values may differ between components, since the vanishing moment by itself does not enforce coherence between the low-pass output channels. The signals $a_{1,\ell}$ can therefore no longer be considered as phases from the constant approximation signal $a(t)$. Though in itself this is already undesired as it hampers interpretation of the decomposition by the filter, this is further complicated when the signals are reused unaltered in a multiresolution structure. Spurious frequencies will then be

introduced and severely compromise the usefulness and interpretability of the multiwavelet analysis. If a reconstruction of the approximation signal $a(t)$ is done, possibly involving a number of upsampling and inverse filter steps, then this imbalance could be undone after which the signal can be split in $2r$ phases again as in **Figure 1** for each wavelet scale. But to ensure that the filters will work properly on dyadic halfbands, this will involve extra effort whereas the interpretability of the approximation signals remains unfixed.

From a practical perspective, the approach as illustrated in **Figure 2** is more appealing. Here, intermediate reconstruction and phase splitting between the scales is omitted. The approximation coefficient sequences are directly split into two phases and propagated to the next scale, as is common for scalar wavelets and consistent with the signal processing formulas (10)–(11). Without balancing, the consequence of this setup is that for a constant input signal at scale 2 the low-pass output $a_{2,\ell}$ will no longer even be constant. This happens despite the fact that the zero-order vanishing moment ensures that constant signals are retained in the approximation low-pass output. With balancing, a constant input signal produces equal output signals $a_{1,\ell}$ which remain well-behaved when further propagated to the next scale.

This issue generalizes to higher order vanishing moments which aim to have all polynomials up to a given degree in the approximation spaces. Including balanced vanishing moments of each order $\leq p$ ensures that every polynomial of degree $\leq p$ is annihilated in the high-pass output and retained *as a polynomial of degree $\leq p$* in the low-pass output.

The p-th order vanishing moment condition on the multiwavelet function $\Psi(t)$ is given by:

$$\int_{\mathbb{R}} \Psi(t)t^p dt = 0. \tag{20}$$

It makes clear that polynomials of degree $\leq p$ are fully suppressed in the detail spaces spanned by the multiwavelet function. The general form of the balancing conditions on the
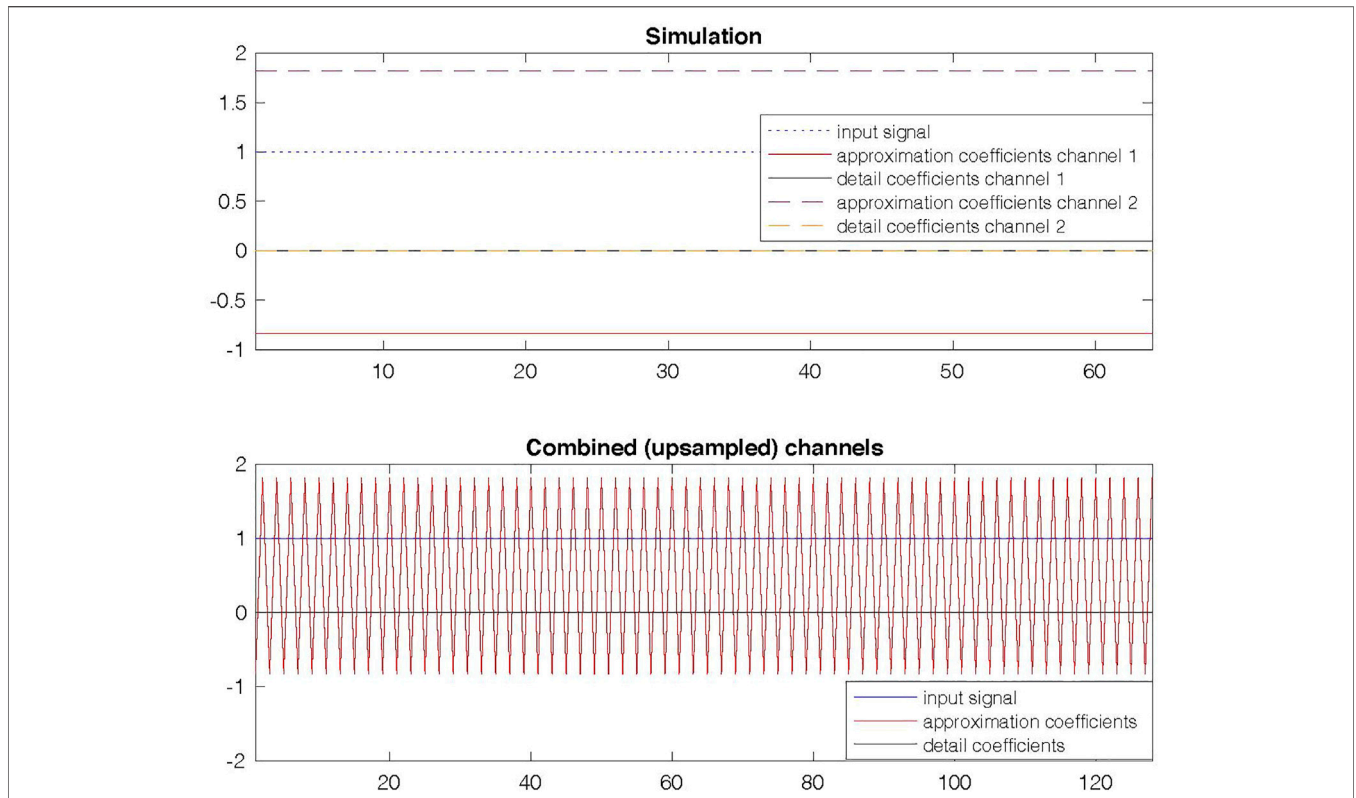
FIGURE 3 | Top: result of feeding a constant signal into an unbalanced multiwavelet filter bank. Bottom: Consequence of interpreting the output of the two low-pass channels as phases of a signal.

multiscaling function $\Phi(t)$ up to order p are given by [11]. In view of our convention for the dilation and wavelet equation they take the form:

$$\int_{\mathbb{R}} \phi_{[0]}(t)\left(t + \frac{0}{r}\right)^p dt = \int_{\mathbb{R}} \phi_{[1]}(t)\left(t + \frac{1}{r}\right)^p dt = \ldots$$
$$= \int_{\mathbb{R}} \phi_{[r-1]}(t)\left(t + \frac{r-1}{r}\right)^p dt. \quad (21)$$

Constructing specific orthogonal multiscaling functions which obey both (20) and (21) up to a given order p can be straightforward, e.g. by using a concatenation of r shifted versions of a Daubechies-p scaling function, where each $\phi_{[j]}(t)$ has been shifted by $-j/r$. Then, due to orthogonality of the shifted Daubechies scaling functions, clearly an orthogonal multiscaling function is obtained, which mimics the scalar setup entirely by having the wavelet forms $\psi_{[j]}(t)$ now capture what is normally achieved by integer translations of a single wavelet. Using such Daubechies-p wavelets ensures vanishing moments up to and including order $p - 1$. Since each component $\phi_{[j]}(t)$ of the multiscaling function is just a time-shifted version of the others, it is also balanced.

However, constructing non-trivial orthogonal multiwavelet functions with different (non-shifted) wave forms and balanced vanishing moments up to a given order p is not straightforward. To address this question, we proceed as follows. (1) First, we derive vanishing moment conditions on the polyphase filter $H_p(z)$. (2)

Next, we derive additional balancing conditions. (3) We investigate how to satisfy all these conditions, up to a chosen order p, by building them into the parameterization of lossless FIR polyphase filters of order n. This is achieved for orders 0 and 1, whereas for order 2 the conditions can be brought into a form which can be used for numerical search.

Vanishing Moment Conditions for Multiwavelets

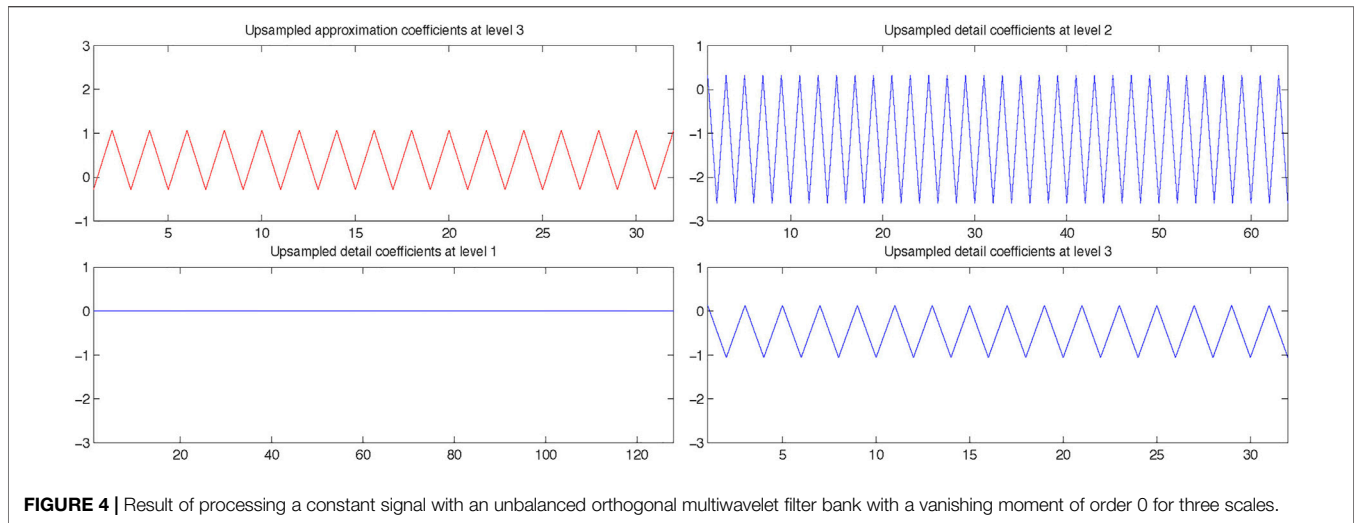Let us introduce the $r \times 1$ vectors v_k and w_k (for $k = 0, 1, 2, \ldots$) by defining:

$$v_k = \int_{\mathbb{R}} \Phi(t)t^k dt, \qquad w_k = \int_{\mathbb{R}} \Psi(t)t^k dt. \quad (22)$$

These are the k-th moments of the multiscaling and multiwavelet functions. From the dilation and wavelet **Eqs 1–2**, we derive the following relationships. First, integration gives

$$v_0 = \frac{1}{2}\sqrt{2}\sum_{k=0}^{2n-1} C_k v_0, \quad (23)$$

$$w_0 = \frac{1}{2}\sqrt{2}\sum_{k=0}^{2n-1} D_k v_0. \quad (24)$$

Here it is used that $\int_{\mathbb{R}} \Phi(2t+k)dt = \frac{1}{2}\int_{\mathbb{R}} \Phi(\tau)d\tau$ by substituting $\tau = 2t + k$.

FIGURE 4 | Result of processing a constant signal with an unbalanced orthogonal multiwavelet filter bank with a vanishing moment of order 0 for three scales.

Next, if we first multiply left and right hand sides of the dilation and wavelet equations by t and then integrate, we obtain in a similar fashion:

$$v_1 = \frac{1}{4}\sqrt{2}\left(\sum_{k=0}^{2n-1} C_k v_1 - \sum_{k=0}^{2n-1} k C_k v_0\right), \quad (25)$$

$$w_1 = \frac{1}{4}\sqrt{2}\left(\sum_{k=0}^{2n-1} D_k v_1 - \sum_{k=0}^{2n-1} k D_k v_0\right). \quad (26)$$

Here it is used that $\int_{\mathbb{R}} \Phi(2t+k)t\,dt = \frac{1}{4}\left(\int_{\mathbb{R}} \Phi(\tau)\tau\,d\tau - k\int_{\mathbb{R}} \Phi(\tau)d\tau\right)$ by writing $t = \frac{1}{2}\left((2t+k)-k\right)$ and substituting $\tau = 2t + k$.

More generally, if we first multiply left and right hand sides of the dilation and wavelet equations by t^2 and then integrate, we obtain:

$$v_2 = \frac{1}{8}\sqrt{2}\left(\sum_{k=0}^{2n-1} C_k v_2 - 2\sum_{k=0}^{2n-1} k C_k v_1 + \sum_{k=0}^{2n-1} k^2 C_k v_0\right), \quad (27)$$

$$w_2 = \frac{1}{8}\sqrt{2}\left(\sum_{k=0}^{2n-1} D_k v_2 - 2\sum_{k=0}^{2n-1} k D_k v_1 + \sum_{k=0}^{2n-1} k^2 D_k v_0\right). \quad (28)$$

Here we wrote $t^2 = \frac{1}{4}\left((2t+k)-k\right)^2$ and substituted $\tau = 2t + k$ to get $\int_{\mathbb{R}} \Phi(2t+k)t^2\,dt = \frac{1}{8}\left(\int_{\mathbb{R}} \Phi(\tau)\tau^2\,d\tau - 2k\int_{\mathbb{R}} \Phi(\tau)\tau\,d\tau + k^2\int_{\mathbb{R}} \Phi(\tau)d\tau\right)$. It is clear how this generalizes to arbitrary order p, using $t^p = \frac{1}{2^p}(\tau - k)^p$.

The vanishing moment conditions up to and including order p are:

$$w_0 = 0, \quad w_1 = 0, \quad \ldots, \quad w_p = 0, \quad (29)$$

because w_k represents the (vector) wavelet coefficient of the polynomial signal t^k for the untranslated multiwavelet $\Psi(t)$, and these wavelet coefficients should all vanish. The vanishing moment conditions state that the filter coefficients C_k and D_k ($k = 0, 1, \ldots, 2n - 1$) should be such as to allow these equations to be

satisfied for some vectors $v_0, v_1, \ldots, v_p$, representing the moments of the multiscaling function.

For orders $p = 0$, $p = 1$, and $p = 2$, the vanishing moment conditions above can be rewritten in terms of the filters $C(z)$ and $D(z)$. Focusing on the equations involving $C(z)$ (the other equations with $D(z)$ are handled entirely analogously) we see from differentiation, that

$$C(z) = \sum_{k=0}^{2n-1} C_k z^{-k}, \qquad C'(z) = -\sum_{k=0}^{2n-1} k C_k z^{-(k+1)},$$

$$C''(z) = \sum_{k=0}^{2n-1} k(k+1) C_k z^{-(k+2)}, \quad (30)$$

whence

$$\sum_{k=0}^{2n-1} C_k = C(1), \qquad \sum_{k=0}^{2n-1} k C_k = -C'(1),$$

$$\sum_{k=0}^{2n-1} k^2 C_k = C''(1) + C'(1). \quad (31)$$

The relationship (17) links the filters $C(z)$ and $D(z)$ to the polyphase filter $H_p(z)$. Differentiation after multiplication by any constant vector v gives:

$$\begin{pmatrix} C'(z) \\ D'(z) \end{pmatrix} v = 2z H_p'(z^2)\begin{pmatrix} v \\ z^{-1}v \end{pmatrix} - z^{-2} H_p(z^2)\begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (32)$$
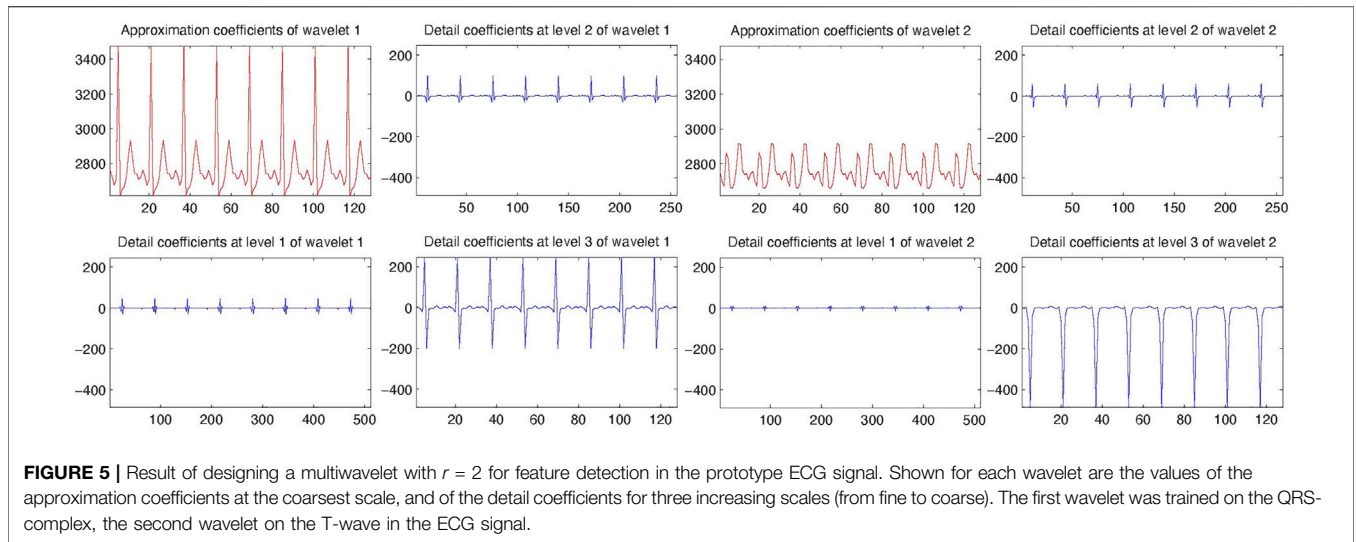
which produces at $z = 1$:

$$\begin{pmatrix} C'(1) \\ D'(1) \end{pmatrix} v = 2 H_p'(1)\begin{pmatrix} v \\ v \end{pmatrix} - H_p(1)\begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (33)$$

An additional differentiation step yields:

$$\begin{pmatrix} C''(z) \\ D''(z) \end{pmatrix} v = 4z^2 H_p''(z^2)\begin{pmatrix} v \\ z^{-1}v \end{pmatrix} + 2 H_p'(z^2)\begin{pmatrix} v \\ z^{-1}v \end{pmatrix}$$

$$- 4z^{-1} H_p'(z^2)\begin{pmatrix} 0 \\ v \end{pmatrix} + 2z^{-3} H_p(z^2)\begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (34)$$

FIGURE 5 | Result of designing a multiwavelet with $r = 2$ for feature detection in the prototype ECG signal. Shown for each wavelet are the values of the approximation coefficients at the coarsest scale, and of the detail coefficients for three increasing scales (from fine to coarse). The first wavelet was trained on the QRS-complex, the second wavelet on the T-wave in the ECG signal.

which gives at $z = 1$:

$$\begin{pmatrix} C''(1) \\ D''(1) \end{pmatrix} v = 4H_p''(1) \begin{pmatrix} v \\ v \end{pmatrix} + 2H_p'(1) \begin{pmatrix} v \\ -v \end{pmatrix} + 2H_p(1) \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{35}$$

These results are combined and summarized in the following theorem.

Theorem 3.2. Let the vectors v_k ($k = 0, 1, 2$) be defined as in **Eq. 22**. The polyphase filter $H_p(z)$ imposes a vanishing moment of order 0 on the multiwavelet if it holds that:

$$\begin{pmatrix} v_0 \\ 0 \end{pmatrix} = \frac{1}{2} \sqrt{2} H_p(1) \begin{pmatrix} v_0 \\ v_0 \end{pmatrix}. \tag{36}$$

A vanishing moment of order 1 occurs if:

$$\begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \frac{1}{4} \sqrt{2} \left(2H_p'(1) \begin{pmatrix} v_0 \\ v_0 \end{pmatrix} + H_p(1) \begin{pmatrix} v_1 \\ v_1 - v_0 \end{pmatrix} \right). \tag{37}$$

A vanishing moment of order 2 occurs if:

$$\begin{pmatrix} v_2 \\ 0 \end{pmatrix} = \frac{1}{8} \sqrt{2} \left(4H_p''(1) \begin{pmatrix} v_0 \\ v_0 \end{pmatrix} + 4H_p'(1) \begin{pmatrix} v_0 + v_1 \\ v_1 \end{pmatrix} \right. $$
$$\left. + H_p(1) \begin{pmatrix} v_2 \\ v_2 - 2v_1 + v_0 \end{pmatrix} \right). \tag{38}$$

To have balanced vanishing moments of order up to and including p, requires extra conditions on the vectors v_k, $k = 0, 1, 2,$ ..., p. This is addressed next.

Balanced Vanishing Moment Conditions for Multiwavelets

The balancing conditions can be computed from **Eq. 21** by integration.

For order $p = 0$ this gives that all entries of v_0 are equal, say with value α_0. To find α_0, the cascade algorithm is useful. For the Haar

multiwavelet in the initialisation step, the scaling function $\phi_{[j]}(t)$ is positive and constant on the interval $(-\frac{j+1}{r}, \frac{j}{r}]$ of width $\frac{1}{r}$; hence, as the basis is orthonormal, this constant value equals $\sqrt{r}$. The value $(v_0)_{[j]}$ therefore initially equals $\frac{1}{\sqrt{r}}$. At the initialisation step, the vector v_0 already has the required form and therefore it is preserved during the cascade algorithm iterations. It follows that $\alpha_0 = \frac{1}{\sqrt{r}}$, and so:

$$v_0 = \frac{1}{\sqrt{r}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{39}$$

For order $p = 1$ it is obtained that $(v_1)_{[j]} + \frac{j}{r}(v_0)_{[j]}$ should have the same value, say $\frac{\lambda}{\sqrt{r}}$ for all $j = 0, 1, \ldots, r - 1$. Imposing balancing of order 1 only makes sense when balancing of order 0 is also imposed. Then, with the earlier result $(v_0)_{[j]} = \frac{1}{\sqrt{r}}$, it is obtained that:

$$v_1 = -\frac{1}{r\sqrt{r}} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ r - 1 \end{pmatrix} + \frac{\lambda}{\sqrt{r}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{40}$$

By inspecting what happens for balanced orthogonal multiwavelets generated from scalar orthogonal wavelets with vanishing moments through shifting, just as explained earlier for the Haar multiwavelet, we find that the value of λ varies between multiwavelets. It therefore is left as a free parameter.

Next, for order $p = 2$ it is obtained that $(v_2)_{[j]} + 2\frac{j}{r}(v_1)_{[j]} + (\frac{j}{r})^2 (v_0)_{[j]}$ should have the same value, say $\frac{\mu}{\sqrt{r}}$ for all $j = 0, 1, \ldots, r - 1$. With the earlier results $(v_0)_{[j]} = \frac{1}{\sqrt{r}}$ and $(v_1)_{[j]} = -\frac{j}{r\sqrt{r}} + \frac{\lambda}{\sqrt{r}}$ we obtain:

$$v_2 = \frac{1}{r^2\sqrt{r}} \begin{pmatrix} 0 \\ 1^2 \\ \vdots \\ (r-1)^2 \end{pmatrix} - \frac{2\lambda}{r\sqrt{r}} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ r - 1 \end{pmatrix} + \frac{\mu}{\sqrt{r}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{41}$$

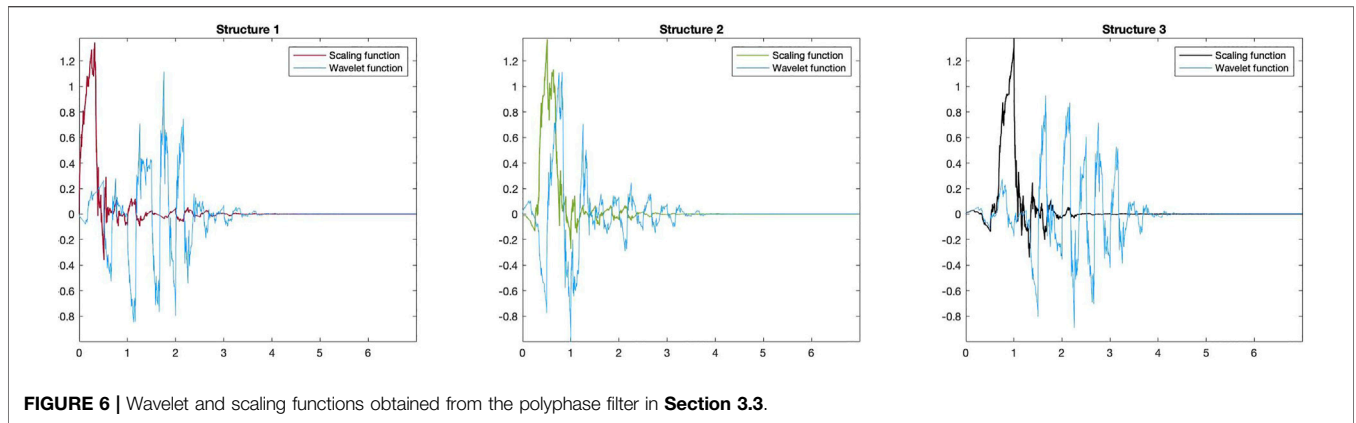The parameter μ is again left free.

FIGURE 6 | Wavelet and scaling functions obtained from the polyphase filter in **Section 3.3**.

From a signal processing point of view, if a signal $s(t)$ is a polynomial in t of degree $\leq p$ and passed through the multiwavelet filter bank starting from its sampled sequence, then the output channels retain the interpretation of phases of sampled signals.

These additional balancing conditions on the vectors v_0, v_1 and v_2 can now be combined with the previous vanishing moment conditions. This gives the following result.

Theorem 3.3. (a) The polyphase filter $H_p(z)$ imposes a balanced vanishing moment of order 0 on the multiwavelet structure if it holds that:

$$\sqrt{2}\begin{pmatrix}1\\\vdots\\1\\0\\\vdots\\0\end{pmatrix} = H_p(1)\begin{pmatrix}1\\\vdots\\1\\1\\\vdots\\1\end{pmatrix}. \tag{42}$$

(b) Balanced vanishing moments of orders 0 and 1 occur if in addition to the condition of (a) it holds that there exists a constant λ such that:

$$2\sqrt{2}\begin{pmatrix}0\\\vdots\\r-1\\0\\\vdots\\0\end{pmatrix} - \lambda r\sqrt{2}\begin{pmatrix}1\\\vdots\\1\\0\\\vdots\\0\end{pmatrix} = -2rH_p'(1)\begin{pmatrix}1\\\vdots\\1\\1\\\vdots\\1\end{pmatrix} + H_p(1)\begin{pmatrix}0\\\vdots\\r-1\\r\\\vdots\\2r-1\end{pmatrix}. \tag{43}$$

(c) Balanced vanishing moments of orders 0, 1 and 2 occur if in addition to the conditions of (a) and (b) it holds that there exists a constant μ such that:

$$4\sqrt{2}\begin{pmatrix}0\\1^2\\\vdots\\(r-1)^2\\0\\\vdots\\0\end{pmatrix} - 4\lambda r\sqrt{2}\begin{pmatrix}0\\1\\\vdots\\r-1\\0\\\vdots\\0\end{pmatrix} + (3\mu - 2\lambda^2)r^2\sqrt{2}\begin{pmatrix}1\\1\\\vdots\\1\\0\\\vdots\\0\end{pmatrix} =$$

$$= 4r^2H_p''(1)\begin{pmatrix}1\\1\\\vdots\\1\\1\\\vdots\\1\end{pmatrix} - 4rH_p'(1)\begin{pmatrix}0\\1\\\vdots\\r-1\\r\\\vdots\\2r-1\end{pmatrix} + 4r^2H_p'(1)\begin{pmatrix}1\\1\\\vdots\\1\\1\\\vdots\\1\end{pmatrix} + H_p(1)\begin{pmatrix}0\\1^2\\\vdots\\(r-1)^2\\r^2\\\vdots\\(2r-1)^2\end{pmatrix}. \tag{44}$$

The proof of this result is by direct computation, using **Eqs 39–41** together with Thm. 3.2. Part (a) of this theorem was shown earlier in the work of [9] and used before in [27]. Parts (b) and (c) are novel characterizations, using the lossless polyphase filter and its derivatives at $z = 1$.

Parameterization of Lossless FIR Polyphase Filters With a Balanced Vanishing Moment of Order 0

A parametrization of lossless FIR polyphase filters with a balanced vanishing moment of order 0 has previously been described in [27, 28]. We present it here for completeness and also because some of the techniques and notation will be reused when we address balancing of order 1. In fact, the parameterization in **Eq. 19** is slightly different from the parameterization used in the literature cited above, but the basic ideas are similar.

From part (a) of Thm. 3.3 we have that for a zero-order balanced vanishing moment, the orthogonal matrix $H_p(1)$ must map the vector $(1 \ldots 1 \quad 1 \ldots 1)^T$ into the vector $\sqrt{2}(1 \ldots 1 \quad 0 \ldots 0)^T$. This can be achieved with Householder transformation matrices. Let R_1 be the Householder matrix which maps $(1 \ldots 1 \quad 0 \ldots 0)^T$ to $\sqrt{r}(1\ 0 \ldots 0 \quad 0 \ldots 0)^T$ and let R_2 be the Householder matrix which maps $(1 \ldots 1 \quad 1 \ldots 1)^T$ to $\sqrt{2r}(1\ 0 \ldots 0 \quad 0 \ldots 0)^T$. A Householder matrix $R_{v,w}$ which maps a vector v into another vector w with $\|w\| = \|v\|$ is orthogonal and symmetric and of the form

$$R_{v,w} = I - 2\frac{(w-v)(w-v)^T}{(w-v)^T(w-v)}. \tag{45}$$

Clearly $R_{v,w} = R_{w,v} = R_{v,w}^T = R_{v,w}^{-1}$. For the matrices R_1 and R_2 we have the explicit forms

$$R_1 = I_{2r} - \frac{1}{r - \sqrt{r}}\begin{pmatrix}1-\sqrt{r}\\1\\\vdots\\1\\0\\\vdots\\0\end{pmatrix}\begin{pmatrix}1-\sqrt{r}\\1\\\vdots\\1\\0\\\vdots\\0\end{pmatrix}^T, \tag{46}$$

$$R_2 = I_{2r} - \frac{1}{2r - \sqrt{2r}} \begin{pmatrix} 1-\sqrt{2r} \\ 1 \\ \vdots \\ 1 \\ \hline 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1-\sqrt{2r} \\ 1 \\ \vdots \\ 1 \\ \hline 1 \\ \vdots \\ 1 \end{pmatrix}^T . \tag{47}$$

Note that the zero-order condition can be rewritten as

$$\sqrt{2} R_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = R_1 H_p(1) R_2 R_2 \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \tag{48}$$

and in this way turns into

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = R_1 H_p(1) R_2 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{49}$$

The matrix $R_1 H_p(1) R_2$ is again orthogonal, and it therefore necessarily is of the form

$$R_1 H_p(1) R_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix}, \tag{50}$$

with Q an orthogonal matrix of size $(2r - 1) \times (2r - 1)$. The converse clearly also holds true, which motivates the following result.

Theorem 3.4. All lossless FIR polyphase filters $H_p(z)$ of order $n - 1$ with a balanced vanishing moment of order 0 are obtained as:

$$H_p(z) = (I_{2r} + (z^{-1} - 1) u_{n-1} u_{n-1}^T) \cdots (I_{2r} + (z^{-1} - 1) u_1 u_1^T) R_1 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} R_2, \tag{51}$$

where R_1 and R_2 are the fixed Householder matrices given in **Eqs 46**, **47**, where Q ranges over the set of $(2r - 1) \times (2r - 1)$ orthogonal matrices, and where $u_1, u_2, \ldots, u_{n-1}$ ranges over the set of $n - 1$ unit vectors of size $2r \times 1$.

Explicit parameterization of the orthogonal matrix Q can be done with Givens rotations and Householder matrices as indicated earlier, see **Section 2.2**. There we also showed how a parameterization of the unit vectors $u_1, \ldots, u_{n-1}$ can be obtained. We see that manipulating the vectors u_k (or increasing their number) does not affect the zero order balancing property as it is fully implied by the structure of **Eq. 51**. This is conveniently exploited when considering an additional balanced vanishing moment of order 1 below.

Parameterization of Lossless FIR Polyphase Filters With Balanced Vanishing Moments of Orders 0 and 1

To build an extra balanced vanishing moment of order 1 into the lossless FIR polyphase filters $H_p(z)$, retaining the balanced vanishing moment of order 0, the idea is to start from the parameterization in Thm. 3.4 and to refine the structure of the orthogonal matrix Q and the unit vectors $u_1, \ldots, u_{n-1}$ to meet the condition of **Eq. 43**. The latter condition involves an additional scalar parameter λ, which can be freed up by premultiplication of all the terms in the equation by the Householder matrix R_1 of **Eq. 46** as we will now show. In view of the structure of $H_p(1) = G^{(0)}$ given by **Eq. 50**, it is useful to first work out the following two matrix-vector products (the horizontal lines are just for clarity):

$$R_1 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ r-1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{r}(r-1)/2 \\ \begin{pmatrix} 1 \\ \vdots \\ r-1 \end{pmatrix} - \frac{r+\sqrt{r}}{2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad R_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ r-1 \\ r \\ \vdots \\ 2r-1 \end{pmatrix} = \begin{pmatrix} \sqrt{2r}(2r-1)/2 \\ \begin{pmatrix} 1 \\ \vdots \\ r-1 \\ r \\ \vdots \\ 2r-1 \end{pmatrix} - \frac{2r+\sqrt{2r}}{2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{pmatrix}. \tag{52}$$

The matrix $H_p'(1)$ results from differentiation of **Eq. 19** followed by substitution of $z = 1$, which gives:

$$H_p'(1) = -\left(\sum_{k=1}^{n-1} u_k u_k^T \right) G^{(0)} = -\left(\sum_{k=1}^{n-1} u_k u_k^T \right) R_1 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} R_2. \tag{53}$$

Using all of this, condition (43) takes the form:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} \left(\begin{pmatrix} \sqrt{2r}(2r-1)/2 \\ \begin{pmatrix} 1 \\ \vdots \\ r-1 \\ r \\ \vdots \\ 2r-1 \end{pmatrix} - \frac{2r+\sqrt{2r}}{2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{pmatrix} + 2r\sqrt{2r} \left(\sum_{k=1}^{n-1} (R_1 u_k)(R_1 u_k)^T \right) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) =$$

$$= \begin{pmatrix} \sqrt{2r}(r-1) \\ 2\sqrt{2} \begin{pmatrix} 1 \\ \vdots \\ r-1 \end{pmatrix} - \sqrt{2r}(\sqrt{r}+1) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} \lambda r\sqrt{2r} \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{54}$$

The parameter λ only appears in the top row of this vector equation. Selecting this row admits computation of λ as:

$$\lambda = -2 \sum_{k=1}^{n-1} \left((R_1 u_k)_1 \right)^2 - \frac{1}{2r}. \tag{55}$$

Here, $(R_1 u_k)_1$ denotes the first entry in the unit vector $R_1 u_k$. It is convenient to parameterize these vectors $R_1 u_k$ (rather than the unit vectors u_k) as

$$R_1 u_k = \begin{pmatrix} \cos(\theta_k) \\ \tilde{u}_k \sin(\theta_k) \end{pmatrix} \tag{56}$$

for some scalar $\theta_k \in [0, \pi]$ and with $\tilde{u}_k$ a unit vector of size $(2r-1) \times 1$. Then

$$\lambda = -2\sum_{k=1}^{n-1} \cos^2(\theta_k) - \frac{1}{2r}, \qquad (57)$$

whereas, upon division by $\sqrt{2r}$, the remaining part of the balancing condition attains the form

$$Q\left(\frac{1}{\sqrt{2r}}\begin{pmatrix}1\\ \vdots\\ r-1\\ r\\ \vdots\\ 2r-1\end{pmatrix} - \frac{\sqrt{2r}+1}{2}\begin{pmatrix}1\\ \vdots\\ 1\\ 1\\ \vdots\\ 1\end{pmatrix}\right) + 2r\left(\sum_{k=1}^{n-1}\tilde{u}_k\sin(\theta_k)\cos(\theta_k)\right) = \frac{2}{\sqrt{r}}\begin{pmatrix}1\\ \vdots\\ r-1\\ 0\\ \vdots\\ 0\end{pmatrix} - (\sqrt{r}+1)\begin{pmatrix}1\\ \vdots\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}. \qquad (58)$$

Introducing the following $(r-1) \times 1$ vectors h_r, for positive integers r:

$$h_r = \frac{2}{r\sqrt{r}}\begin{pmatrix}1\\ \vdots\\ r-1\end{pmatrix} - \frac{1+\sqrt{r}}{r}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} \qquad (59)$$

allows us to summarize these findings concisely in the following theorem.

Theorem 3.5. All lossless FIR polyphase filters $H_p(z)$ of order $n-1$ with two balanced vanishing moments, of orders 0 and 1, are obtained as:

$$H_p(z) = (I_{2r} + (z^{-1}-1)u_{n-1}u_{n-1}^T)\cdots(I_{2r} + (z^{-1}-1)u_1 u_1^T)R_1\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & & &\\ \vdots & & Q &\\ 0 & & &\end{pmatrix}R_2, \qquad (60)$$

where R_1 and R_2 are the fixed Householder matrices given in **Eqs 46, 47**, where Q is $(2r-1) \times (2r-1)$ orthogonal, and where u_k ($k = 1, \ldots, n-1$) are unit vectors of size $2r \times 1$ parameterized as

$$u_k = R_1\begin{pmatrix}\cos(\theta_k)\\ \tilde{u}_k\sin(\theta_k)\end{pmatrix}, \qquad (61)$$

with scalar $\theta_k \in [0, \pi]$, unit vectors $\tilde{u}_k$ of size $(2r-1) \times 1$, and such that the following condition is satisfied:

$$Qh_{2r} = \begin{pmatrix}h_r\\ 0\\ \vdots\\ 0\end{pmatrix} - \sum_{k=1}^{n-1}\tilde{u}_k\sin(2\theta_k), \qquad (62)$$

with the vectors h_r and h_{2r} defined as in **Eq. 59**.

To complete the parameterization of Theorem 3.5, we now show how all tuples of orthogonal Q, unit vectors $\tilde{u}_k$ and scalars θ_k can be obtained which make up all the solutions of **Eq. 62**.

For h_r it is straightforward to compute its norm $\|h_r\|$ as

$$\|h_r\| = \sqrt{\frac{1}{3} - \frac{1}{3r^2}}. \qquad (63)$$

For $r = 1$, we encounter the scalar orthogonal wavelet case, and the vector h_r is in fact empty (of size 0, ×, 1). Balancing is not meaningful then. Indeed the condition reduces to the scalar equation $\frac{-1}{2}Q = -\sum_{k=1}^{n-1}\tilde{u}_k\sin(2\theta_k)$ in which Q as well as $\tilde{u}_1, \ldots, \tilde{u}_{n-1}$ are all ± 1. Note that the scalars $\tilde{u}_k$ can all be fixed to 1, because u_k and therefore also $\tilde{u}_k$ need only be parameterized up to a sign. For Q it is remarked in general that the space of real

orthogonal matrices has two connected components, characterized by the determinant which is either 1 or -1. For $r = 1$, restricting to just one of these components corresponds to using a sign convention for $\psi(t)$, which can be performed by fixing a sign for the last row of $H_p(z)$ and therefore the sign of Q. Choosing $Q = 1$ and taking slight differences in parameterization into account, the resulting condition $\sum_{k=1}^{n-1}\sin(2\theta_k) = \frac{1}{2}$ is entirely consistent, unsurprisingly, with the vanishing moment condition of order 1 reported in [6, 10, 28] or in Proposition 2 in [7].

Focusing on the multiwavelet case with $r \geq 2$, it is noted that $\|h_r\|$ increases monotonically from $\frac{1}{2} = 0.50$ (at $r = 2$) to $\frac{1}{\sqrt{3}} \approx 0.58$ (for $r \to \infty$). It is also noted that $\|Qh_{2r}\| = \|h_{2r}\| = \sqrt{\frac{1}{3} - \frac{1}{12r^2}}$ because Q is orthogonal.

Writing $g_k = -\tilde{u}_k\sin(2\theta_k)$, we see that this is an arbitrary vector of norm ≤ 1, with its direction determined by the unit vector $\tilde{u}_k$ and its norm determined by $\sin(2\theta_k)$. With a sum of m such vectors g_k one can precisely cover all points in a $(2r-1)$-dimensional hyperball of radius m centered at the origin.

The idea of the parameterization, is to build the vectors g_k one by one, such that the condition remains feasible, i.e., the remaining vectors-to-be-constructed can still be assigned values to meet **Eq. 62**. Suppose $g_1, \ldots, g_{\ell-1}$ have been chosen and denote (for $\ell = 1, 2, \ldots, n$):

$$q_\ell := \begin{pmatrix}h_r\\ 0\\ \vdots\\ 0\end{pmatrix} + g_1 + \ldots + g_{\ell-1}. \qquad (64)$$

Then, for $\ell = 1, 2, \ldots, n-2$, the vector g_ℓ must be chosen such that the vector $q_{\ell+1} = q_\ell + g_\ell$ is in (or on) the hyperball (centered at the origin) of radius

$$r_\ell := \|h_{2r}\| + n - 1 - \ell. \qquad (65)$$

This still allows the right-hand side of **Eq. 62**, viz. the expression $q_{\ell+1} + g_{\ell+1} + \cdots + g_{n-1} = q_n$, to eventually land on the hypersphere of vectors with norm $\|h_{2r}\|$. For the left-hand side expression Qh_{2r}, note that the norm equals $\|h_{2r}\|$ for every orthogonal Q, and that a suitable orthogonal matrix Q can always be constructed to make Qh_{2r} equal to that landing point (a Householder matrix would do). In fact, all such orthogonal Q can be constructed along the same lines as before when constructing $G^{(0)}$ for balancing of order 0, as we will demonstrate below in **Eq. 74**. Note that if one cannot land on that hypersphere the construction is infeasible, since the norm of Qh_{2r} is fixed; so this precisely characterizes feasibility. (Note, also, that points inside the hypersphere of radius $\|h_{2r}\|$ can always be brought to a point on the hypersphere in one step, by adding a single vector g_k, since $\|h_{2r}\| < 1$; therefore one need not consider a constraint on the inside of the hypersphere.)

For $\ell = 1, 2, \ldots, n-2$, we proceed as follows. If $\|q_\ell\| \leq r_\ell - 1$, then g_ℓ can be any vector of norm ≤ 1, so there are no particular constraints on $\tilde{u}_\ell$ (of norm 1) or θ_ℓ. Note that $r_\ell - 1 = \|h_{2r}\| + n - 2 - \ell \geq \|h_{2r}\| \geq \frac{1}{2}$, so this is a non-empty set of vectors including the case $q_\ell = 0$.

If $\|q_\ell\| > r_\ell - 1$, then let $q_\ell^\perp$ denote an arbitrary vector of unit norm and orthogonal to q_ℓ. The space of such vectors is easily parameterized.[6] Next, note that each vector g_ℓ can be written uniquely as:

$$g_\ell = \alpha_\ell q_\ell + \beta_\ell q_\ell^\perp, \qquad (66)$$

with scalar coefficients α_ℓ (still unrestricted) and $\beta_\ell \geq 0$. For g_ℓ to be feasible it must hold that: (1) $\|g_\ell\| \leq 1$, (2) $\|q_\ell + g_\ell\| \leq r_\ell$. Because q_ℓ and $q_\ell^\perp$ are orthogonal to each other, setting $\beta_\ell = 0$ gives the maximum range for α_ℓ. Requirement (1) implies: $\alpha_\ell \in [-\frac{1}{\|q_\ell\|}, \frac{1}{\|q_\ell\|}]$. Requirement (2) implies: $\alpha_\ell \in [-1 - \frac{r_\ell}{\|q_\ell\|}, -1 + \frac{r_\ell}{\|q_\ell\|}]$. Combining these two intervals, using $\|q_\ell\| > r_\ell - 1$, it follows that both requirements are met if and only if:

$$-\frac{1}{\|q_\ell\|} \leq \alpha_\ell \leq -1 + \frac{r_\ell}{\|q_\ell\|}. \qquad (67)$$

Once α_ℓ is chosen, β_ℓ can be chosen non-negative but such that the two conditions remain satisfied. This gives:

$$0 \leq \beta_\ell \leq \min\left\{ \sqrt{1 - \alpha_\ell^2 \|q_\ell\|^2}, \sqrt{r_\ell^2 - (1 + \alpha_\ell)^2 \|q_\ell\|^2} \right\}. \qquad (68)$$

For $\ell = n - 1$, choosing the final vector g_{n-1} (of norm ≤ 1) is special, in the sense that it must satisfy the equality $\|q_{n-1} + g_{n-1}\| = r_{n-1} = \|h_{2r}\|$ rather than an inequality as before.

If $\|q_{n-1}\| \leq 1 - r_{n-1}$, then q_n can land on every point of the hypersphere with radius r_{n-1} by choosing g_{n-1} as the difference between that point and q_{n-1}.

If $\|q_{n-1}\| > 1 - r_{n-1}$, then the hypersphere centered at the origin with radius r_{n-1} and the hypersphere centered at q_{n-1} with radius 1 intersect at points $q_n = q_{n-1} + g_{n-1} = (1 + \alpha_{n-1}) q_{n-1} + \beta_{n-1} q_{n-1}^\perp$ if the following two conditions hold:

$$\alpha_{n-1}^2 \|q_{n-1}\|^2 + \beta_{n-1}^2 = 1, \qquad (69)$$

$$(1 + \alpha_{n-1})^2 \|q_{n-1}\|^2 + \beta_{n-1}^2 = r_{n-1}^2. \qquad (70)$$

Subtracting the first equation from the second leaves a linear equation for α_{n-1}, which gives:

$$\alpha_{n-1} = -\frac{1}{2}\left(1 + \frac{1 - r_{n-1}^2}{\|q_{n-1}\|^2} \right). \qquad (71)$$

This determines the minimal value for α_{n-1} to give solutions on the hypersphere for q_n. The maximal value for α_{n-1} is obtained with $\beta_{n-1} = 0$ and equals $\alpha_{n-1} = -1 + \frac{r_{n-1}}{\|q_{n-1}\|}$. This implies that all feasible $g_{n-1} = \alpha_{n-1} q_{n-1} + \beta_{n-1} q_{n-1}^\perp$ have

$$-\frac{1}{2}\left(1 + \frac{1 - r_{n-1}^2}{\|q_{n-1}\|^2} \right) \leq \alpha_{n-1} \leq -1 + \frac{r_{n-1}}{\|q_{n-1}\|} \qquad (72)$$

and

$$\beta_{n-1} = \sqrt{r_{n-1}^2 - (1 + \alpha_{n-1})^2 \|q_{n-1}\|^2}. \qquad (73)$$

---

[6] To parameterize all such $q_\ell^\perp$, if $q_\ell \neq 0$, one may proceed as follows. Define the Householder matrix P_ℓ to map the first unit vector $(1 \ 0 \ \ldots \ 0)^T$ to the vector $q_\ell/\|q_\ell\|$. Then the columns $2, 3, \ldots, 2r - 1$ of P_ℓ constitute an orthonormal basis for the space of vectors $q_\ell^\perp$. All the vectors $q_\ell^\perp$ are produced as $P_\ell y$ with y ranging over all the unit vectors with their first entry equal to 0. Such vectors y are recursively generated analogous to **Eq. 56**.

Finally, to find all feasible orthogonal Q once the right-hand side expression q_n is constructed to have $\|q_n\| = r_{n-1} = \|h_{2r}\|$, let R_3 be the Householder matrix which maps q_n to $\|h_{2r}\| (1 \ 0 \ \ldots \ 0)^T$ and let R_4 be the Householder matrix which maps h_{2r} to $\|h_{2r}\| (1 \ 0 \ \ldots \ 0)^T$. Then all feasible orthogonal matrices Q are obtained as

$$Q = R_3 \begin{pmatrix} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q} & \\ 0 & & & \end{pmatrix} R_4 \qquad (74)$$

with $\tilde{Q}$ an arbitrary orthogonal matrix of size $(2r - 2) \times (2r - 2)$.

Balanced Vanishing Moments of Orders 0, 1 and 2

Incorporating an additional balanced vanishing moment of order $p = 2$, requires further refinement of the current parameterization such that condition (44) is also satisfied. As we have not developed a way to incorporate this into a parameterization, we will only show how the condition can be reworked into a form which makes it suitable to be added as a constraint to the parametrization in terms of the parameters used before.

A first step to achieve this, is to free up and remove the parameter μ by premultiplication of all terms in the condition by the Householder matrix R_1. The effect of this is that μ only appears in the term $(3\mu - 2\lambda^2) r^2 \sqrt{2} R_1 (1 \ 1 \ \ldots \ 1 \ 0 \ \ldots \ 0)^T = (3\mu - 2\lambda^2) r^2 \sqrt{2r} (1 \ 0 \ \ldots \ 0 \ 0 \ \ldots \ 0)^T$, of which just the first entry is nonzero. Therefore, as before for λ for balancing of order 1, the parameter μ can be left free and computed afterwards from the equation which corresponds to the first entries on the left and right-hand sides.

The actual conditions for balancing of order 2, are then captured by the remaining $2r - 1$ equations in which μ no longer appears. The parameter λ which does show up in them, should be replaced by its value in terms of the chosen parameters as given in **Eq. 57**. This gives the constraints in a form which can readily be used for numerical optimization, using a routine for constrained optimization which admits nonlinear constraints.

Design Criteria for Matching

For the scalar orthogonal wavelet design case, sparsity of a prototype signal was already advocated in [7]. This is appealing for detection and compression purposes. There it was also shown that maximization of sparsity of the vector of all approximation and detail coefficients $w = \{w_k\}$ in the setting with a critically sampled orthogonal wavelet transform, boils down to either: 1) maximization of the variance of the sequence of absolute values $|w_k|$ of the coefficients, or 2) maximization of the variance of the squared wavelet coefficients. This is due to the fact that Parseval's identity holds [4], which makes that the sum of squares of all the wavelet and detail coefficients equals the sum of squares of all the values of the digital signal being processed. This result does obviously not change for orthogonal multiwavelet design. Hence [[7], Theorem 1], holds in the current setting too:

Theorem 3.6. Let w be the vector of the approximation coefficients at the coarsest scale and the detail coefficients at

all scales, resulting from the processing of a signal s by means of an orthogonal multiwavelet filter bank across multiple scales. Then:

1) Maximization of the variance of the vector of absolute values $\{|w_k|\}$ is equivalent to minimization of the L_1-norm of the vector w.
2) Maximization of the variance of the energy vector $\{|w_k|^2\}$ is equivalent to maximization of the L_4-norm of the vector w.

For practical applications the latter criterion, L_4-maximization, is more appealing than the former since it can be combined with weighing both in scale and location [27]. This allows for designing a matched multiwavelet, where each of the components of $\Phi(t)$ is forced to focus on a single aspect in a signal (see **Section 3.2** for an example).

If instead of a critically sampled multiwavelet transform an undecimated multiwavelet transform [33] is used, this has advantages for detection purposes, since due to the redundant representation time-invariance is achieved. The lack of down-sampling, however, causes an abundance of coefficients at coarser scales. To preserve conservation of energy properties for orthogonal multiwavelets, this abundance can be counteracted by dyadic discounting of energies towards coarser scales as in [7]. This makes it possible to use adapted versions of the L_1-minimization and L_4-maximization criteria for undecimated multiwavelet transforms promoting sparsity.

Another criterion one might consider is entropy. One can quantify the effectiveness of compression algorithms in terms of entropy [34]. As suggested in [35], entropy can be used to select an orthogonal wavelet basis, and a similar approach can serve as a design criterion. This is certainly appealing for compression purposes.

ILLUSTRATIVE EXAMPLES

Experiment to Illustrate the Balancing Problem

The effects of having unbalanced vanishing moments are easily illustrated by considering what happens if a constant signal is fed into an unbalanced orthogonal multiwavelet filter bank with multiplicity $r = 2$ and of order $n - 1 = 3$, using the construction scheme in Theorem 3.1, but without balancing. To this end we took $Q = I_{2r-1} = I_3$ and we changed the Householder matrix R_1 to produce a non-constant vector v_0.

When a constant input signal is fed into an orthogonal multiwavelet filter bank with an order 0 unbalanced vanishing moment, as described in **Section 3.1**, the results (depending on the exact choice of coefficients) are as shown in **Figure 3**. What can be observed in the top figure, is that the detail coefficients are all zero for both channels (consistent with the imposed vanishing moments), and the approximation coefficients are constant per channel, but the values are different between the channels. In terms of the notation in **Eqs 8**, **9**, the detail vector coefficients b_k are all zero, the approximation vector coefficients a_k are all constant, but their entries $(a_k)_1$ and $(a_k)_2$ are different.

In the bottom figure of **Figure 3** it is displayed what the resulting signal looks like if the outputs are nevertheless interpreted as the two phases of a signal.

If the outputs are now processed further with the multiwavelet filter bank for a few more scales, this has the effect that at later scales the constant nature of the original signal is lost and the detail coefficients become nonzero, as demonstrated in **Figure 4**.

Example of Multiwavelet Design for ECG Feature Detection

In this section we introduce an example for balanced multiwavelet design. We will employ L_4-maximization with weighted masking to match the multiwavelet filter bank to a prototype signal. The example evolves around creating a representation for a prototype ECG signal as is commonly encountered in the field of cardiology. The goal of this design is for feature detection to distinguish and detect different complexes in the ECG signal.

Feature detection is an application that makes efficient use of the different components of the multiwavelet. The idea of feature detection via multiwavelets is to design each component of the multiwavelet to detect a specific feature in the signal. Orthogonality between the components of the scaling- and wavelet function allows the multiwavelet to detect up to r orthogonal features, which cannot all be accurately detected by a scalar wavelet at the same time. Since features in a signal need not be orthogonal but can be overlapping, making them hard to separate, the multiwavelet approach is fundamentally different from template based approaches. When training orthogonal multiwavelets on different features, the goal is to have them pick up orthogonal aspects from the features that help to distinguish between them.

For this application the variance maximizing L_4-objective function is no longer measured across all the wavelet coefficients. Instead, a time-scale mask is employed for each component of the multiwavelet [28], and the value of the objective function of a component is only measured on the wavelet coefficients that are in the time-scale mask. In this way, maximization of the criterion forces energy into the masked areas, helping the multiwavelets to focus on the user-selected signal features covered by the time-scale masks. (This same idea is less conveniently pursued with L_1-minimization, because minimization will force energy out of the masked areas, but while the energy leaks away this does not mean that useful multiwavelets are promoted.)

In this example, the design procedure employs a first-order balanced orthogonal multiwavelet, with focus on the approximation coefficients rather than the detail coefficients, and aiming to detect features by looking at the peaks of the approximation coefficients at masked areas at the coarsest scale $l = 3$.

It can be seen from **Figure 5** that the first scaling function mainly represents the QRS-complex on which it was trained, whereas the second scaling function manages to capture the T-wave on which it was trained. By thresholding the approximation coefficients of the first channel the location of the QRS-complex can be obtained. By thresholding the second

channel the location of the T-peak is obtained. Both channels can detect the feature that they were designed for independently from one another. This is remarkable, as the T-wave overlaps in spectrum and waveform with the QRS-complex, which carries far more energy. Though the second wavelet also picks up some energy from the QRS-complex, it still allows to detect the T-peak location with simple thresholding. This demonstrates that a matched multiwavelet is a powerful signal processing tool to detect and discriminate between different features in a signal.

Example of an Orthogonal Multiwavelet With Compact Support, Multiplicity 3 and Balanced Vanishing Moments of Orders 0, 1, and 2

In this example we show an orthogonal multiwavelet filter with multiplicity $r = 3$, involving a lossless FIR polyphase filter $H_p(z)$ of order $n - 1 = 3$, with balanced vanishing moments of orders $p = 0$, $p = 1$, and $p = 2$ as a concrete example of multiwavelet design with three balanced vanishing moments. To obtain such a filter, we utilized the parameterization that was introduced in **Section 2.7** for multiwavelets with balanced vanishing moments of orders 0 and 1, and we followed the approach of **Section 2.8**. For $r = 3$ and $n - 1 = 3$, we first chose random parameters to obtain an initial orthogonal multiwavelet only lacking a second order vanishing moment. From there, constrained optimization was used to adapt the parameters to satisfy condition (44), producing an orthogonal multiwavelet with three balanced vanishing moments.

For this example, the FIR polyphase filter $H_p(z) = H_0 + H_1 z^{-1} + H_2 z^{-2} + H_3 z^{-3}$ has the following coefficient matrices H_k:

$$H_0 = \begin{pmatrix} 0.5698 & 0.8165 & 0.0101 & 0.0248 & 0.0105 & -0.0342 \\ -0.0317 & -0.0070 & 0.7024 & 0.6980 & 0.1093 & -0.0555 \\ 0.0167 & 0.0051 & -0.0687 & 0.0488 & 0.5488 & 0.8164 \\ -0.0351 & 0.1002 & 0.1538 & -0.2415 & 0.0724 & -0.0026 \\ 0.0577 & -0.0455 & -0.4025 & 0.2809 & 0.6175 & -0.3845 \\ 0.0170 & 0.0280 & -0.0517 & 0.0043 & 0.1349 & -0.0645 \end{pmatrix}$$

(75)

$$H_1 = \begin{pmatrix} 0.0551 & -0.0351 & 0.0089 & -0.0026 & -0.0170 & 0.0190 \\ -0.0205 & 0.0109 & 0.0101 & -0.0165 & 0.0270 & -0.0225 \\ 0.1108 & -0.0792 & 0.0464 & -0.0487 & 0.0339 & -0.0205 \\ -0.4528 & 0.2714 & 0.2526 & -0.3512 & 0.4382 & -0.3106 \\ -0.3402 & 0.2485 & -0.0335 & 0.0410 & -0.0668 & 0.0795 \\ -0.0069 & -0.0331 & -0.4161 & 0.3877 & -0.0384 & -0.1749 \end{pmatrix}$$

(76)

$$H_2 = \begin{pmatrix} -0.0283 & 0.0203 & -0.0104 & 0.0109 & -0.0080 & 0.0053 \\ 0.0266 & -0.0187 & 0.0077 & -0.0076 & 0.0042 & -0.0024 \\ 0.0136 & -0.0092 & 0.0020 & -0.0012 & -0.0011 & 0.0013 \\ 0.3095 & -0.2110 & 0.0525 & -0.0384 & -0.0127 & 0.0188 \\ -0.1124 & 0.0831 & -0.0559 & 0.0630 & -0.0558 & 0.0394 \\ 0.4275 & -0.3234 & 0.2539 & -0.2942 & 0.2794 & -0.2014 \end{pmatrix}$$

(77)

$$H_3 = \begin{pmatrix} -0.0042 & 0.0029 & -0.0011 & 0.0010 & -0.0004 & 0.0002 \\ 0.0012 & -0.0008 & 0.0003 & -0.0003 & 0.0001 & -0.0001 \\ -0.0020 & 0.0014 & -0.0005 & 0.0005 & -0.0002 & 0.0001 \\ -0.0360 & 0.0251 & -0.0093 & 0.0086 & -0.0037 & 0.0017 \\ -0.0362 & 0.0252 & -0.0093 & 0.0087 & -0.0037 & 0.0018 \\ 0.1926 & -0.1343 & 0.0495 & -0.0462 & 0.0196 & -0.0093 \end{pmatrix}$$

(78)

Associated with this polyphase filter are the wavelet and scaling functions as shown in **Figure 6**.

By definition, the Householder matrices R_1 and R_2 are given by:

$$R_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) & -\frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) & 0 & 0 & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) & \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5774 & 0.5774 & 0.5774 & 0 & 0 & 0 \\ 0.5774 & 0.2113 & -0.7887 & 0 & 0 & 0 \\ 0.5774 & -0.7887 & 0.2113 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(79)

$$R_2 = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{1}{5}\left(4 - \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) \\ \frac{1}{\sqrt{6}} & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & \frac{1}{5}\left(4 - \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) \\ \frac{1}{\sqrt{6}} & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & \frac{1}{5}\left(4 - \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) \\ \frac{1}{\sqrt{6}} & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & \frac{1}{5}\left(4 - \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) \\ \frac{1}{\sqrt{6}} & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & -\frac{1}{5}\left(1 + \frac{1}{\sqrt{6}}\right) & \frac{1}{5}\left(4 - \frac{1}{\sqrt{6}}\right) \end{pmatrix} =$$

$$= \begin{pmatrix} 0.4082 & 0.4082 & 0.4082 & 0.4082 & 0.4082 & 0.4082 \\ 0.4082 & 0.7184 & -0.2816 & -0.2816 & -0.2816 & -0.2816 \\ 0.4082 & -0.2816 & 0.7184 & -0.2816 & -0.2816 & -0.2816 \\ 0.4082 & -0.2816 & -0.2816 & 0.7184 & -0.2816 & -0.2816 \\ 0.4082 & -0.2816 & -0.2816 & -0.2816 & 0.7184 & -0.2816 \\ 0.4082 & -0.2816 & -0.2816 & -0.2816 & -0.2816 & 0.7184 \end{pmatrix}$$

(80)

For the example, the orthogonal matrix Q and the unit vectors u_1, u_2 and u_3 which appear in the parameterization of $H_p(z)$ have the following values:

$$Q = \begin{pmatrix} 0.6826 & 0.3298 & 0.3193 & -0.2805 & -0.4946 \\ 0.7291 & -0.2987 & -0.2422 & 0.2729 & 0.4959 \\ 0.0378 & 0.3017 & -0.7704 & 0.3463 & -0.4406 \\ 0.0139 & -0.7987 & 0.0961 & 0.1939 & -0.5613 \\ -0.0280 & 0.2703 & 0.4864 & 0.8303 & -0.0153 \end{pmatrix}$$

(81)

$$u_1 = \begin{pmatrix} -0.0191 \\ -0.0221 \\ -0.1660 \\ 0.0116 \\ 0.6349 \\ -0.7539 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0.0565 \\ -0.0586 \\ -0.0355 \\ -0.7850 \\ 0.1836 \\ -0.5850 \end{pmatrix}, \quad u_3 = \begin{pmatrix} -0.0209 \\ 0.0060 \\ -0.0102 \\ -0.1808 \\ -0.1816 \\ 0.9663 \end{pmatrix}$$

(82)

The values of the parameters λ and μ encountered for this particular balanced multiwavelet of order 2, are given by:

$$\lambda = -0.1966, \qquad \mu = 0.0387. \tag{83}$$

The resulting unit vectors $R_1 u_1$, $R_1 u_2$, and $R_1 u_3$ are:

$$R_1 u_1 = \begin{pmatrix} -0.1196 \\ 0.1152 \\ -0.0287 \\ 0.0116 \\ 0.6349 \\ -0.7539 \end{pmatrix}, \quad R_1 u_2 = \begin{pmatrix} -0.0217 \\ 0.0482 \\ 0.0713 \\ -0.7850 \\ 0.1836 \\ -0.5850 \end{pmatrix}, \quad R_1 u_3 = \begin{pmatrix} -0.0145 \\ -0.0028 \\ -0.0189 \\ -0.1808 \\ -0.1816 \\ 0.9663 \end{pmatrix}$$

(84)

These vectors $R_1 u_k$ (k = 1, 2, 3) are partitioned as $R_1 u_k = \begin{pmatrix} \cos(\theta_k) \\ \tilde{u}_k \sin(\theta_k) \end{pmatrix}$ with $\theta_k = \arccos((R_1 u_k)_1) \in [0, \pi]$ and $\|\tilde{u}_k\| = 1$. It holds that

$$\theta_1 = 1.6907, \qquad \theta_2 = 1.5925, \qquad \theta_3 = 1.5853. \tag{85}$$

The vectors g_k, which feature in the parameterization for balancing of order 1, are defined as $g_k = -\tilde{u}_k \sin(2\theta_k)$ and given by:

$$g_1 = \begin{pmatrix} 0.0276 \\ -0.0069 \\ 0.0028 \\ 0.1518 \\ -0.1803 \end{pmatrix}, \quad g_2 = \begin{pmatrix} 0.0021 \\ 0.0031 \\ -0.0341 \\ 0.0080 \\ -0.0254 \end{pmatrix}, \quad g_3 = \begin{pmatrix} -0.0001 \\ -0.0005 \\ -0.0053 \\ -0.0053 \\ 0.0281 \end{pmatrix}$$

(86)

The vectors h_3 and h_6 are defined as follows:

$$h_3 = -\frac{1}{9}\begin{pmatrix} 3 + \sqrt{3} \\ 3 - \sqrt{3} \end{pmatrix}, \qquad h_6 = -\frac{1}{18}\begin{pmatrix} 3 + 2\sqrt{6} \\ 3 + \sqrt{6} \\ 3 \\ 3 - \sqrt{6} \\ 3 - 2\sqrt{6} \end{pmatrix}, \tag{87}$$

having norms: $\|h_3\| = \sqrt{8/27} \approx 0.5443$ and $\|h_6\| = \sqrt{35/108} \approx 0.5693$. It holds that

$$Q h_6 = \begin{pmatrix} h_3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + g_1 + g_2 + g_3. \tag{88}$$

It is verified by direct computation upon substitution into **Eqs 16, 51, 42, 43, 44, 62** that the orthogonality conditions as well as the three balanced vanishing moment conditions for order $p = 0$,

1, 2 are all properly satisfied, and entirely consistent with the proposed parameterization.

DISCUSSION

The results in **Section 3.1** show, from a signal processing perspective, the importance of having balanced vanishing moments for multiwavelets. Not having this means that polynomials do not yield zero detail coefficients in subsequent scales, even if the required vanishing moments are built in. The example in **Section 3.2** shows that if the available remaining freedom is used to promote sparsity, this has a benefit for detection purposes. Other design objectives could also be employed. When employing weighing over scales, this also opens the opportunity to use overcomplete representations for shift-invariance. Masking in the time-scale plane allows to focus on selected parts of a prototype signal. When one would switch to a criterion based on the information value, this can be valuable for compression purposes.

In **Section 3.3** a concrete example was worked out for the design of orthogonal multiwavelets with compact support and three balanced vanishing moments (orders 0, 1, and 2) and multiplicity r = 3. Previous examples of orthogonal multiwavelets in the literature with several balanced vanishing moments either had balancing only up to order 1, or multiplicity r = 2; see [9, 11]. The Gröbner basis approach used there is a limiting factor for finding matching balanced multiwavelets to applications, because the approach does not allow for excess degrees of freedom. In this paper we advocated a different approach by building an explicit parameterization which allows free parameters to be tuned for various purposes. We have shown how *all* balanced orthogonal multiwavelets of orders 0 and 1 can be obtained for arbitrary multiplicity r and any given polyphase filter order (McMillan degree), which we consider a major step forward in making multiwavelets applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JK, SS, and RP together conceived and explored the concept of parameterizing balanced orthogonal multiwavelets with built-in vanishing moment properties. Numerical simulations were implemented and performed by SS. JK, and RP provided the background on multiwavelets, balancing, and parameterisation of lossless systems, and provided cardiac signal data. JK and RP composed the final article based on an initial draft by SS.

REFERENCES

1. Grossmann A, Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM J Math Anal* (1984) 15:723–36. doi:10.1137/0515056

2. Grossmann A. Wavelet Transforms and Edge Detection. In: S Albeverio, P Blanchard, M Hazewinkel, L Streit, editors. *Stochastic Processes in Physics an Engineering*. Dordrecht, The Netherlands: D. Reidel Publishing Company (1988). p. 149–57. ISBN: 90-277-2659-0. doi:10.1007/978-94-009-2893-0_7

3. Daubechies I. Orthonormal Bases of Compactly Supported Wavelets. *Comm Pure Appl Math* (1988) 41:909–96. doi:10.1002/cpa.3160410705

4. Mallat S. *A Wavelet Tour of Signal Processing: The Sparse Way*. Amsterdam: Academic Press (2009). doi:10.1016/B978-0-12-374370-1.X0001-8

5. Karel JMH, Peeters RLM, Westra RL, Moermans KMS, Haddad SAP, Serdijn WA. Optimal Discrete Wavelet Design for Cardiac Signal Processing. In: Proceeding: 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society (EMBC 2005); 1-4 September, 2005; Shanghai, China. IEEE (2005). p. 2769–72. doi:10.1109/IEMBS.2005.1617046

6. Peeters R, Karel J. Data Driven Design of an Orthogonal Wavelet with Vanishing Moments. In: 21st International Symposium on Mathematical Theory of Networks and Systems; Groningen, Netherlands, 2014 7–11 July (2014). p. 1665–72.

7. Karel J, Peeters R. Orthogonal Matched Wavelets with Vanishing Moments: A Sparsity Design Approach. *Circuits Syst Signal Process* (2018) 37:3487–514. doi:10.1007/s00034-017-0716-1

8. Hanzon B, Olivi M, Peeters RLM. Balanced Realizations of Discrete-Time Stable All-Pass Systems and the Tangential Schur Algorithm. *Linear Algebra Appl* (2006) 418:793–820. doi:10.1016/j.laa.2006.03.027

9. Lebrun J, Vetterli M. High-order Balanced Multiwavelets: Theory, Factorization, and Design. *IEEE Trans Signal Process* (2001) 49:1918–30. doi:10.1109/78.942621

10. Strang G, Nguyen T. *Wavelets and Filter Banks*. Wellesley, MA.: Wellesley-Cambridge Press (1996). ISBN: 978-0961408879.

11. Selesnick IW. Multiwavelet Bases with Extra Approximation Properties. *IEEE Trans Signal Process* (1998) 46:2898–908. doi:10.1109/78.726804

12. Chui CK, Jiang Q. Balanced Multi-Wavelets in $\mathbb{R}s$. *Math Comp* (2004) 74:1323–45. doi:10.1090/S0025-5718-04-01681-3

13. Sun H, Zi Y, He Z. Wind Turbine Fault Detection Using Multiwavelet Denoising with the Data-Driven Block Threshold. *Appl Acoust* (2014) 77: 122–9. doi:10.1016/j.apacoust.2013.04.016

14. Chen J, Li Z, Pan J, Chen G, Zi Y, Yuan J, et al. Wavelet Transform Based on Inner Product in Fault Diagnosis of Rotating Machinery: A Review. *Mech Syst Signal Process* (2016) 70-71:1–35. doi:10.1016/j.ymssp.2015.08.023

15. Chen J, Wan Z, Pan J, Zi Y, Wang Y, Chen B, et al. Customized Maximal-Overlap Multiwavelet Denoising with Data-Driven Group Threshold for Condition Monitoring of Rolling Mill Drivetrain. *Mech Syst Signal Process* (2016) 68-69:44–67. doi:10.1016/j.ymssp.2015.07.022

16. Hong L, Liu X, Zuo H. Compound Faults Diagnosis Based on Customized Balanced Multiwavelets and Adaptive Maximum Correlated Kurtosis Deconvolution. *Measurement* (2019) 146:87–100. doi:10.1016/j.measurement.2019.06.022

17. He J, Yang Z, Chen C, Li G, Li Z, Jia Y. Improved Multiwavelet Denoising with Neighboring Coefficients of Cutting Force for Application in the Load Spectrum of Computer Numerical Control Lathe. *Adv Mech Eng* (2018) 10: 1–11. doi:10.1177/1687814018754674

18. Alkhidhr H, Jiang Q. Correspondence between Multiwavelet Shrinkage and Nonlinear Diffusion. *J Comput Appl Math* (2021) 382:113074. doi:10.1016/j.cam.2020.113074

19. Jiang Q. On the Design of Multifilter banks and Orthonormal Multiwavelet Bases. *IEEE Trans Signal Process* (1998) 46:3292–303. doi:10.1109/78.735304

20. Chui CK, Jiang Q. Multivariate Balanced Vector-Valued Refinable Functions. In: W Haussmann, K Jetter, M Reimer, J Stöckler, editors. *Modern Developments in Multivariate Approximation*. Springer Basel AG: Basel, Vol. 145 (2003). p. 71–102. doi:10.1007/978-3-0348-8067-1_4

21. Han B, Lu R. Compactly Supported Quasi-Tight Multiframelets with High Balancing Orders and Compact Framelet Transforms. *Appl Comput Harmonic Anal* (2021) 51:295–332. doi:10.1016/j.acha.2020.11.005

22. Han B, Lu R. Multivariate Quasi-Tight Framelets with High Balancing Orders Derived from Any Compactly Supported Refinable Vector Functions. *Sci China Math* (2021):1–30. [Epub ahead of print]. doi:10.1007/s11425-020-1786-9

23. Lebrun J, Vetterli M. Balanced Multiwavelets. In: 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing; 21-24 April, 1997; Munich, Germany, 3 (1997). p. 2473–6. doi:10.1109/ICASSP.1997.599579

24. Lebrun J, Vetterli M. Balanced Multiwavelets Theory and Design. *IEEE Trans Signal Process* (1998) 46:1119–25. doi:10.1109/78.668561

25. Bacchelli S, Cotronei M, Lazzaro D. An Algebraic Construction of K-Balanced Multiwavelets via the Lifting Scheme. *Numer Algorithms* (2000) 23:329–56. doi:10.1023/A:1019120621646

26. Li B, Peng L. Balanced Multiwavelets with Interpolatory Property. *IEEE Trans Image Process* (2011) 20:1450–7. doi:10.1109/TIP.2010.2092439

27. Peeters RLM, Karel JMH, Westra RL, Haddad SAP, Serdijn WA. Multiwavelet Design for Cardiac Signal Processing. In: Proceeding: 28th Annual International Conference of the IEEE Enigneering in Medicine and Biology Society (EMBC 2006); 30 August-3 September, 2006; New York City (2006). p. 1682–5. doi:10.1109/IEMBS.2006.259733

28. Karel J. *A Wavelet Approach to Cardiac Signal Processing for Low-Power Hardware Applications*. Universitaire Pers Maastricht (2009). ISBN: 978-90-5278-887-6.

29. Smith M, Barnwell T. A Procedure for Designing Exact Reconstruction Filter banks for Tree-Structured Subband Coders. *IEEE Int Conf Acoust Speech Signal Process* (1984) 9:421–4. doi:10.1109/ICASSP.1984.1172486

30. Vaidyanathan P. Theory and Design of M-Channel Maximally Decimated Quadrature Mirror Filters with Arbitrary M, Having the Perfect-Reconstruction Property. *IEEE Trans Acoust Speech Signal Process* (1987) 35:476–92. doi:10.1109/TASSP.1987.1165155

31. Vaidyanathan PP, Doganata Z. The Role of Lossless Systems in Modern Digital Signal Processing: a Tutorial. *IEEE Trans Educ* (1989) 32:181–97. doi:10.1109/13.34150

32. Selesnick IW. Balanced Multiwavelet Bases Based on Symmetric FIR Filters. *IEEE Trans Signal Process* (2000) 48:184–91. doi:10.1109/78.815488

33. Nason GP, Silverman BW. The Stationary Wavelet Transform and Some Statistical Applications. In: A Antoniadis G Oppenheim, editors. *Lecture Notes in Statistics: Wavelets and Statistics*. New York, NY: Springer-Verlag, Vol. 103 (1995). p. 281–99. doi:10.1007/978-1-4612-2544-7_17

34. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J* (1948) 27:379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

35. He H, Tan Y, Wang Y. Optimal Base Wavelet Selection for Ecg Noise Reduction Using a Comprehensive Entropy Criterion. *Entropy* (2015) 17: 6093–109. doi:10.3390/e17096093

36. Peeters R, Olivi M, Hanzon B. Balanced Realization of Lossless Systems: Schur Parameters, Canonical Forms and Applications. In: Proceedings of the 15th IFAC SYSID 2009; 6-8 July, 2009; Saint-Malo (2009). p. 273–83. doi:10.3182/20090706-3-fr-2004.00045

37. Golub GH, Van Loan CF. *Matrix Computations*. 3rd ed. Baltimore, MD: Johns Hopkins University Press (1996). ISBN: 0-8018-5413-X.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.