

Efficient Design of Orthonormal Wavelet Bases for Signal Representation

Jian-Kang Zhang, Timothy N. Davidson, *Member, IEEE*, and Kon Max Wong, *Fellow, IEEE*

Abstract—The efficient representation of a signal as a linear combination of elementary “atoms” or building blocks is central to much signal processing theory and many applications. Wavelets provide a powerful, flexible, and efficiently implementable class of such atoms. In this paper, we develop an efficient method for selecting an orthonormal wavelet that is matched to a given signal in the sense that the squared error between the signal and some finite resolution wavelet representation of it is minimized. Since the squared error is not an explicit function of the design parameters, some form of approximation of this objective is required if conventional optimization techniques are to be used. Previous approximations have resulted in nonconvex optimization problems, which require delicate management of local minima. In this paper, we employ an approximation that results in a design problem that can be transformed into a convex optimization problem and efficiently solved. Constraints on the smoothness of the wavelet can be efficiently incorporated into the design. We show that the error incurred in our approximation is bounded by a function that decays to zero as the number of vanishing moments of the wavelet grows. In our examples, we demonstrate that our method provides wavelet bases that yield substantially better performance than members of standard wavelet families and are competitive with those designed by more intricate nonconvex optimization methods.

Index Terms—Convex optimization, signal adapted wavelet design, signal representation.

I. INTRODUCTION

A SIGNAL expansion expresses a signal as linear combination of elementary atoms or building blocks, and, as such, signal expansions are central to much signal processing theory and many applications. Recently, the wavelet transform has emerged as a powerful and efficient tool for signal expansion and has shown potential in several applications, e.g., [1]. For example, in computer vision, wavelet-based multiresolution has been used for motion estimation and stereo vision problems [2], and multiresolution techniques have found applications in transient and edge detection [3], [4], in computer graphics, pitch estimation [5], medical imaging [6], and communications [7], [8]. One of the features of the wavelet transform is that it can be viewed as a family of transforms indexed by the wavelet function (or the scaling function). When selecting a wavelet function, one may pursue a “universal” approach in which one seeks a wavelet that provides “good” performance over a broad

class of signals, e.g., [9]. An alternative approach is to seek a “signal-adapted” wavelet that is “matched” to a particular signal or to a narrow class of signals, e.g., [10], [11]. In recent years, several authors have developed systematic and efficient design algorithms for signal-adapted filterbanks [12]–[22]. The goal of the present paper is to develop a correspondingly systematic and efficient design algorithm for signal-adapted wavelet bases. In particular, for a given signal of interest, we will address the problem of finding the wavelet (or, more precisely, the scaling function) that minimizes the squared error between the signal and some finite resolution wavelet representation of itself [10], [11]. A simple version of this problem can be formally stated as follows: *Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$ and an even integer L , find an orthonormal scaling function $\varphi(t)$ supported on $[0, L - 1]$ that minimizes $\|f(t) - P_0(f, \varphi)\|_{L_2(\mathbb{R})}^2$, where $\|x(t)\|_{L_2(\mathbb{R})}^2 = \int_{-\infty}^{\infty} |x(t)|^2 dt$ is the energy of $x(t)$, and $P_0(f, \varphi)$ represents the projection of the signal $f(t)$ onto the multiresolution subspace $V_0(\varphi) = \text{span}\{\varphi(t - n)\}_{n \in \mathbb{Z}}$, i.e.,*

$$P_0(f, \varphi) = \sum_{n \in \mathbb{Z}} \langle f(t), \varphi(t - n) \rangle \varphi(t - n)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. While this problem is stated for a deterministic signal $f(t)$, we will point out in (12) that our approach extends directly to stationary stochastic signals with a given power spectral density. The above problem can be viewed as a sampling problem [23] in which the “generating function” $\varphi(t)$ is required to satisfy the constraints of an orthonormal multiresolution analysis and is designed so that the squared error between the original signal $f(t)$ and the reconstructed signal $P_0(f, \varphi)$ is minimized.

At first glance, this problem appears to be infinite-dimensional (the “parameter to be optimized” is a function, $\varphi(t)$). However, the set of orthonormal scaling functions of support $[0, L - 1]$ is determined [24, Ch. 6] by L parameters $h[\ell]$ through a fixed point of the “dilation equation” $\varphi(t) = \sqrt{2} \sum_{\ell=0}^{L-1} h[\ell] \varphi(2t - \ell)$. Although this results in a finite-dimensional optimization problem, there is no explicit relationship between $h[\ell]$ and $\varphi(t)$. (The dilation equation is an implicit relationship.) This makes it difficult to determine the gradients of our objective with respect to the parameters, which would ordinarily be used in the solution of the problem. Several authors have recognized the fact that this represents a rather awkward optimization problem and have developed approximations to the cost function that are explicit functions of the design parameters [10], [11]. (The method in [10] is applicable to a more general setting than the bandlimited

Manuscript received November 20, 2002; revised August 7, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yuan-Pei Lin.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, L8S 4K1, Canada (e-mail: jkzhang@mail.ece.mcmaster.ca; davidson@mcmaster.ca; wong@mail.ece.mcmaster.ca).

Digital Object Identifier 10.1109/TSP.2004.828923

setting considered in [11] and herein.) These approximations will be described in more detail in Section II, but in order to motivate our design approach, we point out two drawbacks of the approximations in [10] and [11]. First, although the gradients of the approximated objective with respect to design parameters can be explicitly determined, the resulting optimization problem is not convex, and hence, delicate management of local minima will be required. (The importance of this management of local minima was observed in both [10] and [11].) Second, the established methods do not explicitly control the error incurred in the approximation of the objective, leading to concerns that the optimal solution might incur considerable approximation error.

In this paper, we address both these drawbacks by deriving an approximation of the objective that is not only explicit in the design parameters but can be transformed into a convex optimization problem from which a globally optimal solution can be efficiently found. (The convex optimization problem we obtain is a semidefinite program.) Our derivation includes an explicit bound on the error incurred in the approximation of the objective, and this bound can be incorporated into the convex optimization problem. Furthermore, constraints on the “smoothness” of the scaling function can be easily incorporated into the formulation. Therefore, using our new method, good wavelet bases matched to a given application can be efficiently designed. In our examples, we demonstrate that our method provides wavelet bases that yield substantially lower squared error than members of the standard wavelet families and are competitive with those designed by more intricate nonconvex optimization methods.

Notation

In order to distinguish between continuous functions, discrete sequences, and their Fourier transforms, we will use the following notational conventions. These conventions have been influenced by both the engineering and mathematics literature. Functions of a real variable will have the variable enclosed in parentheses, (e.g., $f(t), t \in \mathbb{R}$), and functions of a discrete variable will have the variable enclosed in brackets (e.g., $h[\ell], \ell \in \mathbb{Z}$). We will use the notation $\hat{f}(\omega)$ to denote the continuous-time Fourier transform (CTFT) of $f(t)$, that is, $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt$. We will let $H(\omega)$ denote the discrete time Fourier transform (DTFT) of $h[\ell]$, that is, $H(\omega) = \sum_{\ell} h[\ell]e^{-j\omega\ell}$. Given our choice of notation for the CTFT, approximations will be denoted by a tilde. Given a matrix \mathbf{X} , the (i, j) th element of \mathbf{X} will be denoted by $[\mathbf{X}]_{ij}$.

II. BACKGROUND

We begin with a brief review of the basics of wavelet theory [24]–[26]. Let $h[\ell], 0 \leq \ell \leq L-1$ be the impulse response of a finite impulse response (FIR) filter. Furthermore, let $h[\ell]$ be self-orthogonal at even shifts, i.e., let

$$\sum_{\ell} h[\ell]h[\ell - 2n] = \delta[n] \quad (1)$$

where $\delta[n]$ is the Kronecker delta. Now, let $\varphi(t)$ be a scaling function generated by this filter via a fixed point of the dilation equation

$$\varphi(t) = \sqrt{2} \sum_{\ell=0}^{L-1} h[\ell]\varphi(2t - \ell) \quad (2)$$

and let $\psi(t)$ be the corresponding wavelet function

$$\psi(t) = \sqrt{2} \sum_{\ell=0}^{L-1} (-1)^{\ell} h[L-1-\ell]\varphi(2t - \ell). \quad (3)$$

Then, the set of functions $\{\psi_{j,n}(t) = 2^{j/2}\psi(2^j t - n)\}_{j,n \in \mathbb{Z}}$ forms a tight frame [27] for the space of all finite energy functions $L_2(\mathbb{R})$. In order to ensure that this set of functions forms an orthonormal basis for $L_2(\mathbb{R})$, we need to enforce more constraints on $h[\ell]$ in addition to those in (1). For example, a necessary condition is that $\sum_{\ell} h[\ell] = \pm\sqrt{2}$. A set of necessary and sufficient conditions for the set to form an orthonormal basis is available in [28], and another is available in [29], but for the purposes of this paper, we prefer to use Daubechies' sufficient condition [24], [30] because it can be easily reformulated as a convex constraint.

Lemma 1 [24, p. 182], [30]: Let the scaling filter $h[\ell]$ of length L satisfy the orthonormality condition in (1), and let $h[\ell]$ be such that

$$H(\omega) = \left(\frac{1 + e^{-j\omega}}{2} \right)^N Q(\omega) \quad (4)$$

for some $N > 0$ and some $Q(\omega) = \sum_{\ell=0}^{L_q-1} q[\ell]e^{-j\omega\ell}$ with $Q(0) = \sqrt{2}$. If

$$|Q(\omega)| < 2^{N-1/2} \quad \text{for all } \omega \in [-\pi, \pi] \quad (5)$$

then the scaling function in (2) is orthonormal, and the set $\{\psi_{j,n}(t) = 2^{j/2}\psi(2^j t - n)\}_{j,n \in \mathbb{Z}}$, where $\psi(t)$ was defined in (3), forms an orthonormal basis for $L_2(\mathbb{R})$.

We point out that a necessary condition for an $H(\omega)$ in the form of (4) to satisfy the orthonormality condition in (1) is that $N \leq L/2$ [24, p. 171]. Given an $h[\ell]$ that satisfies Lemma 1, any finite energy signal $f(t) \in L_2(\mathbb{R})$ can be decomposed using the following wavelet series:

$$f(t) = \sum_{j,n \in \mathbb{Z}} d[j, n]\psi_{j,n}(t) \quad (6)$$

where the indices j and n denote the scale and translation of the wavelet, respectively. The projection coefficients are simply obtained by the inner product

$$d[j, n] = \int_{-\infty}^{\infty} f(t)\psi_{j,n}(t) dt.$$

In the design of the decomposition in (6), we are often interested in the components of $f(t)$ at and below the J th scale. That is, $\sum_{j \leq J} \sum_n d[j, n]\psi_{j,n}(t)$. From the theory of multiresolution analysis [24, Ch. 5], [26, p. 221], this is simply the projection of $f(t)$ onto $V_J = \text{span}\{2^J\varphi(2^J t - n)\}_{n \in \mathbb{Z}}$, which is the space spanned by the scaling function at the J th scale. To describe that projection, we let $\varphi_{j,n}(t) = 2^{j/2}\varphi(2^j t - n)$ and

$a[j, n] = \int_{-\infty}^{\infty} f(t)\varphi_{j,n}(t) dt$. Then, the projection of $f(t)$ onto V_J is

$$P_J(f, \varphi) = \sum_n a[J, n]\varphi_{J,n}(t). \quad (7)$$

Therefore, the decomposition in (6) can be partitioned into components at or below the J th scale and those at scales above (i.e., finer than) the J th scale

$$f(t) = P_J(f, \varphi) + \sum_{j>J} \sum_n d[j, n]\psi_{j,n}(t). \quad (8)$$

As stated in the Introduction, the design problem in which we are interested is the following: For a given function $f(t)$, determine the scaling function $\varphi(t)$ that minimizes the squared error between $P_J(f, \varphi)$ and $f(t)$

$$\mathcal{E}_J(f, \varphi) = \|f(t) - P_J(f, \varphi)\|_{L_2(\mathbb{R})}^2 \quad (9)$$

where $\|x(t)\|_{L_2(\mathbb{R})}^2 = \int |x(t)|^2 dt$. For functions $f(t)$ that are bandlimited to $[-\pi, \pi]$, the Nyquist sampling interval is 1 s, and hence, the natural scale partition in (8) is $J = 0$. [In that case, the time shift in the function $\varphi_{J,n}(t)$ in (7) is 1 s.] For such signals, Parseval's relation can be used to obtain the following simple expression for $\mathcal{E}_0(f, \varphi)$ [11]

$$\mathcal{E}_0(f, \varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 (1 - |\hat{\varphi}(\omega)|^2) d\omega. \quad (10)$$

In many applications, it may be desirable to minimize $\mathcal{E}_0(f, \varphi)$ over a set of scaling functions that possesses certain properties. In particular, it may be desirable to impose lower bounds on the number of vanishing moments of the corresponding wavelet and on the number of its derivatives that are continuous. The number of vanishing moments, which are denoted by N , is the largest non-negative integer n such that $\int_{-\infty}^{\infty} t^n \psi(t) dt = 0$, that is, the order of locally polynomial signals that are orthogonal to the mother wavelet $\psi(t)$. These polynomial signals thus lie in the corresponding "scaling" subspace V_0 , and hence, N is sometimes referred to as the approximation order of the scaling function. The number of vanishing moments of the wavelet is easily controllable. In fact, any smooth orthogonal wavelet for which the corresponding scaling filter $h[\ell]$ has a zero of multiplicity N at $\omega = \pi$ [i.e., satisfies (4)] has N vanishing moments. In contrast, it can be rather awkward to parameterize all filters that generate wavelets with at least a specified number of continuous derivatives. However, we can obtain a convenient parameterization of a class of scaling filters that generate wavelets with N vanishing moments and at least $M \geq 0$ continuous derivatives by simply replacing (5) in Lemma 1 by the following constraint [24, p. 216]:

$$|Q(\omega)| < 2^{N-M-1/2} \quad \text{for all } \omega \in [-\pi, \pi]. \quad (11)$$

Note that when $M = 0$, (11) reduces to (5). As shown in Appendix A, a necessary condition on M for a $Q(\omega)$ satisfying Lemma 1 and (11) to exist is $M < N/2$.

We are now ready to formally state our design problem.

Problem 1: Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$, an integer L that denotes the length of the scaling filter, and integers $0 < N < L/2$ and $0 \leq M < N/2$, find

a (orthonormal) scaling function $\varphi(t)$ that minimizes $\mathcal{E}_0(f, \varphi)$, subject to (1), (2), (4), and (11).

Before we attempt to solve Problem 1, we point out the following fact: If $f(t)$ is a stochastic signal with a power spectral density $S_f(\omega)$, which is zero outside the interval $[-\pi, \pi]$, then

$$\mathbb{E}\{\mathcal{E}_0(f, \varphi)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_f(\omega)(1 - |\hat{\varphi}(\omega)|^2) d\omega \quad (12)$$

where \mathbb{E} denotes expectation. Although the focus of the present paper is on the design of orthonormal wavelet bases matched to a bandlimited deterministic signal, the common algebraic structure of (10) and (12) indicates that our approach can be extended in a straightforward manner to the design of wavelet bases matched to a bandlimited stochastic signal.

As we pointed out in the Introduction, Problem 1 is a rather awkward optimization problem because the objective is not an explicit function of the parameters $h[\ell]$, which uniquely define $\varphi(t)$ according to (2). Therefore, several authors have suggested bounding or approximating the objective by an explicit function of $h[\ell]$. The approach taken in [11] exploits the Fourier transform of the scaling equation (2), namely

$$\hat{\varphi}(\omega) = \frac{1}{\sqrt{2}} H\left(\frac{\omega}{2}\right) \hat{\varphi}\left(\frac{\omega}{2}\right) \quad (13)$$

which allows us to write $\hat{\varphi}(\omega) = \hat{\varphi}(0) \prod_{k=1}^{\infty} (1/\sqrt{2}) H(\omega/2^k)$, where $\hat{\varphi}(0) = 1$. Approximating this infinite product by a finite one, we have that

$$\hat{\varphi}(\omega) \approx \tilde{\varphi}^{(K)}(\omega) = \prod_{k=1}^K \frac{1}{\sqrt{2}} H\left(\frac{\omega}{2^k}\right) \quad (14)$$

where, consistent with the notational conventions in the Introduction, we have used a tilde to denote the approximation of $\hat{\varphi}(\omega)$. Using (14), we can approximate Problem 1 in the following way.

Problem 2: Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$ and integers $L \geq 2$, $0 < N \leq L/2$, $0 \leq M < N/2$, and $K > 0$, find a (orthonormal) scaling filter $h[\ell]$, $0 \leq \ell \leq L - 1$, which minimizes

$$\tilde{\mathcal{E}}_0^{(K)}(f, h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 (1 - |\tilde{\varphi}^{(K)}(\omega)|^2) d\omega \quad (15)$$

subject to (1), (4), and (11).

Gopinath *et al.* [11] proposed a similar approximation of Problem 1 but did not include the constraint in (11).

Once Problem 2 has been solved, the corresponding scaling function can be found by solving (2) using, for example, the cascade method [24, Sect. 6.5] or the iterated filter method [25, Sect. 4.4.2]. An advantage of Problem 2 is that the objective is now an explicit function of the parameters $h[\ell]$. However, Problem 2 remains an awkward problem to solve because it is not convex in $h[\ell]$. Any algorithm used to design $h[\ell]$ directly will require delicate management of local minima—a point emphasized in [10] and [11]. Furthermore, as K is increased so that (15) becomes a better approximation for (10), the objective in (15) becomes a more complicated function of $h[\ell]$, which can make the calculation of derivatives of the objective quite cumbersome. In addition, while searching for

the scaling filter that minimizes $\tilde{\mathcal{E}}_0^{(K)}(f, h)$, one ought to guard against amplifying the approximation error $\hat{\varphi}(\omega) - \tilde{\varphi}^{(K)}(\omega)$ in (14). (This point was overlooked in [11].) These drawbacks motivate us to find a design method that provides a compromise between the complexity of the optimization problem and the accuracy of the approximation while allowing us to explicitly bound approximation error. A candidate method is provided in the next section.

III. TRANSFORMATION OF PROBLEM

The purpose of this section is to derive an alternative approximation of Problem 1 that results in a convex optimization problem that can be efficiently solved while allowing explicit control of the approximation error.

A. Approximation of $\hat{\varphi}(\omega)$

We begin the derivation with a proposition that describes an approximation of $\hat{\varphi}(\omega)$ and provides a bound on the approximation error. For convenience, we first introduce some notation. Let $\Phi(\omega) = \sum_{k \in \mathbb{Z}} \varphi(k) e^{-jk\omega}$ be the *discrete-time* Fourier transform (DTFT) of the integer samples of the scaling function $\varphi(t)$, and let

$$\Phi_m = \max_{\omega \in [-\pi, \pi]} |\Phi(\omega)|, \quad \text{and} \quad Q_m = \max_{\omega \in [-\pi, \pi]} |Q(\omega)| \quad (16)$$

where $Q(\omega)$ was defined in Lemma 1.

Proposition 1: If $h[\ell]$ satisfies the conditions of Lemma 1 with $N > 0$, then

$$\hat{\varphi}(\omega) = \Phi(\omega) + \mu(\omega) \quad (17)$$

where

$$|\mu(\omega)| \leq \frac{Q_m \Phi_m}{(1 - 2^{-N})\sqrt{2}} \left| \frac{\omega}{4} \right|^N. \quad (18)$$

Furthermore

$$|\hat{\varphi}(\omega)|^2 = 1 + \mu_1(\omega) \quad (19)$$

where

$$|\mu_1(\omega)| \leq \frac{Q_m^2}{2(1 - 2^{-2N})} \left| \frac{\omega}{4} \right|^{2N}. \quad (20)$$

The proof of Proposition 1 is given in Appendix B. In essence, Proposition 1 reflects the lowpass characteristic of the scaling function. It also reflects the fact that the zero moments control the accuracy with which $\Phi(\omega)$ approximates $\hat{\varphi}(\omega)$ near low frequencies, i.e., the greater the approximation order of the scaling function, the higher the accuracy at low frequencies. An advantage of Proposition 1 over some other approximations of $\hat{\varphi}(\omega)$ is that the error term can be controlled quantitatively via (18) or (20).

B. Formulations

To develop our convex formulation of the wavelet design problem, we need to transform the objective and constraints into convex functions of the design parameters. We do so in Sections III-B1 and III-B2, respectively. The formulation itself is provided in Section III-B3.

1) *Transformation of the Objective Function:* Employing Lemma 1 and Proposition 1, we can approximate the original nonconvex objective function by a convex one. Applying Proposition 1 to the right-hand side of (13) leads to

$$|\hat{\varphi}(\omega)|^2 = \frac{1}{2} \left| H \left(\frac{\omega}{2} \right) \right|^2 + \mu_2(\omega) \quad (21)$$

where

$$|\mu_2(\omega)| \leq \frac{Q_m^2}{2(1 - 2^{-2N})} \left| \frac{\omega}{8} \right|^{2N} \quad (22)$$

and we have used the fact that $|H(\omega)| \leq \sqrt{2}$. Substituting (21) into (10) yields

$$\begin{aligned} \mathcal{E}_0(f, \varphi) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \left(1 - \frac{1}{2} \left| H \left(\frac{\omega}{2} \right) \right|^2 \right) d\omega + \Delta(f, \varphi) \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \left| H \left(\frac{\omega}{2} + \pi \right) \right|^2 d\omega + \Delta(f, \varphi) \end{aligned} \quad (23)$$

where we have used the fact that $|H(\omega)|^2 + |H(\pi + \omega)|^2 = 2$, [30]. (An expression for $\Delta(f, \varphi)$ appears in (77) in Appendix E.) The approximation error term satisfies

$$|\Delta(f, \varphi)| \leq \frac{Q_m^2 M_N(f)}{2^{4N+1}(2^{2N} - 1)} = \beta_N(f) Q_m^2 \quad (24)$$

with

$$M_N(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega^{2N} |\hat{f}(\omega)|^2 d\omega \quad (25)$$

$$\beta_N(f) = \frac{M_N(f)}{2^{4N+1}(2^{2N} - 1)}. \quad (26)$$

Note that $M_N(f)$ is the $2N$ th moment of the power spectrum of $f(t)$, and, as such, is a measure of the ‘‘spread’’ of $\hat{f}(\omega)$.

Since $f(t)$ is bandlimited to $[-\pi, \pi]$, it is naturally bandlimited to $[-2\pi, 2\pi]$. Hence, it can be represented by samples taken at integer multiples of $1/2$, i.e.,

$$f(t) = \sum_{k \in \mathbb{Z}} f \left(\frac{k}{2} \right) \frac{\sin \pi(2t - k)}{\pi(2t - k)}. \quad (27)$$

Let

$$r_h[k] = \sum_{m \in \mathbb{Z}} h[m] h[m + k] \quad (28)$$

$$b_j[k] = \frac{1}{2^j} \sum_{m \in \mathbb{Z}} f \left(\frac{m}{2^j} \right) f \left(\frac{m + k}{2^j} \right). \quad (29)$$

As shown in Appendix C, we can use (27)–(29) to rewrite the objective (23) as

$$\mathcal{E}_0(f, \varphi) = \frac{b_1[0]}{2} + \sum_{k=1}^{L-1} (-1)^k b_1[k] r_h[k] + \Delta(f, \varphi) \quad (30)$$

where we have used the fact that $r_h[0] = 1$ and the symmetry of $r_h[k]$ and $b_1[k]$. The first term of (30) is a constant that depends only on $f(t)$. The second term in (30) is a linear, and hence convex, function of the autocorrelation sequence $r_h[k]$. However, the third term in (30) is an implicit function of the filter coefficients, and its presence makes (30) a difficult function to

minimize. However, we can alleviate this difficulty by replacing $\Delta(f, \varphi)$ by its upper bound in (24). Hence

$$\mathcal{E}_0(f, \varphi) \leq \frac{b_1[0]}{2} + \sum_{k=1}^{L-1} (-1)^k b_1[k] r_h[k] + \beta_N(f) \lambda \quad (31)$$

where $\lambda \geq Q_m^2$. Since we know $f(t)$, for a given N , we can calculate $M_N(f)$ in (25) and, hence, $\beta_N(f)$ in (26) analytically:

$$M_N(f) = \frac{\pi^{2N} b_0[0]}{2N+1} + 2 \sum_{\ell=1}^N (-1)^{\ell-1} \pi^{2(N-\ell)} P_{2N}^{2\ell-1} \sum_{k=1}^{\infty} (-1)^k \frac{b_0[k]}{k^{2\ell}} \quad (32)$$

where the symbol P_m^n represents the number of permutations taking n elements from m elements. The details of the derivation of (32) are provided in Appendix D. The corresponding expression for $\beta_N(f)$ can be obtained by substituting (32) into (26). In practice, the infinite sum in (32) will be approximated, leading to an approximate value for $\beta_N(f)$. Amalgamating these results, we obtain the following approximation of Problem 1.

Problem 3: Given a particular signal $f(t)$, which is bandlimited to $[-\pi, \pi]$ and integers $L \geq 2$, $0 < N \leq L/2$, and $0 \leq M < N/2$, find a (orthonormal) scaling filter $h[\ell]$, $0 \leq \ell \leq L-1$, and a scalar $\lambda > 0$ that achieve

$$\min_{h[\ell], \lambda} \frac{b_1[0]}{2} + \sum_{k=1}^{L-1} (-1)^k b_1[k] r_h[k] + \beta_N(f) \lambda \quad (33a)$$

subject to (1), (4)

$$|Q(\omega)|^2 \leq \lambda \quad \text{for all } \omega \quad (33b)$$

$$\lambda \leq 2^{2(N-M-1/2)}. \quad (33c)$$

If the term $\beta_N(f)$ is removed from (33a), then Problem 3 is equivalent to Problem 2 with $K = 1$. However, Problem 3 has the advantage that it can be transformed into a convex optimization problem, as we will show below. Furthermore, by including the term $\beta_N(f) \lambda$ in (33a), we can explicitly bound the difference between the value of $\mathcal{E}_0(f, \varphi)$ in (10) achieved by the scaling function corresponding to the optimal solution to Problem 3 and the true optimal value of $\mathcal{E}_0(f, \varphi)$. The latter value is, of course, the optimal objective value for Problem 1. This bound is derived in Section III-C and decays to zero with increasing N .

2) *Transformation of the Constraints:* The key observation in the transformation of Problem 3 into a convex optimization problem is that the objective and constraints can be written as convex functions of the autocorrelation of the filter (and λ).¹

This observation is straightforward in the case of the objective because (33a) is already expressed as a linear (and hence convex) function of $r_h[k]$ and λ . (Since the term $b_1[0]/2$ does not depend on $h[\ell]$ or λ , it can be removed from the objective.) The case of the constraints is a little more subtle, and we now deal with each one in turn.

a) *Regularity Constraint:* In order to ensure that the resulting wavelet has N zero moments, we require that the scaling filter satisfy the N th-order regularity condition, i.e., it has a zero

¹Similar observations have led to efficient design algorithms for other FIR filter design problems [31]–[33], including the design of signal-adapted filterbanks [16], [21].

of multiplicity N at $\omega = \pi$ or, equivalently, that $H(\omega)$ satisfies (4). In the time domain, this condition is equivalent to

$$\sum_{\ell=0}^{L-1} (-1)^\ell \ell^n h[\ell] = 0 \quad \text{for } n = 0, 1, \dots, N-1. \quad (34)$$

Requiring $H(\omega)$ to have N zeros at $\omega = \pi$ is equivalent to requiring its autocorrelation, which has a frequency response $|H(\omega)|^2$, to have $2N$ zeros at π . This condition is equivalent to

$$\sum_{k=-L+1}^{L-1} (-1)^k k^n r_h[k] = 0 \quad \text{for } n = 0, 1, \dots, 2N-1. \quad (35)$$

However, since $r_h[k]$ is symmetric, this is automatically true for odd n . Hence, (35) is equivalent to

$$\sum_{k=0}^{L-1} (-1)^k k^{2n} r_h[k] = 0 \quad \text{for } n = 0, 1, \dots, N-1 \quad (36)$$

which is N linear equality constraints on $r_h[k]$.

b) *Orthonormality Constraint:* In our design problem, we require that the scaling function is orthonormal. A necessary condition for this is that the scaling filter is orthonormal. This is equivalent to (1). Since $r_h[k] = \sum_{\ell \in \mathbb{Z}} h[\ell] h[\ell+k]$, it is easy to express (1) in term of variables $r_h[k]$:

$$r_h[2n] = \delta[n] \quad \text{for } n = 0, 1, \dots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (37)$$

where the symbol $\lfloor x \rfloor$ denotes the largest integer not exceeding x , and we have used the symmetry of $r_h[k]$. Equation (37) consists of another $\lfloor (L-1)/2 \rfloor + 1$ linear equality constraints on $r_h[k]$. In order to guarantee that the generated scaling function is orthonormal and has at least $M \geq 0$ continuous derivatives, we enforce Daubechies' sufficient condition (11) on $Q(\omega)$, which is equivalent to

$$R_q(\omega) < 2^{2N-2M-1} \quad \text{for all } \omega \quad (38)$$

where $r_q[k] = \sum_{\ell \in \mathbb{Z}} q[\ell] q[\ell+k]$, and hence, $R_q(\omega) = \sum_n r_q[k] e^{-j\omega k} = |Q(\omega)|^2$. Using (4), inequality (38) is equivalent to

$$R_h(\omega) < 2^{N-2M-1} (1 + \cos \omega)^N \quad \text{for all } \omega \quad (39)$$

which is an infinite set of linear inequality constraints on $r_h[k]$: one for each ω .

c) *Normalization Constraint:* We know from the theory of wavelets [24, Sect. 6.2] that the normalization condition $\sum_{\ell=0}^{L-1} h[\ell] = \sqrt{2}$ and the orthonormality condition (1) on the scaling filter ensure that the scaling function generated by $h[\ell]$ belongs to $L_2(\mathbb{R})$. This implies that $R_h(0) = 2$, from which we can obtain $\sum_{\ell=0}^{L-1} h[\ell] = \pm\sqrt{2}$. This is essentially equivalent to the original normalization condition because when performing spectral factorization, we can require that $\sum_{\ell=0}^{L-1} h[\ell] > 0$. In addition, note that the condition $R_h(0) = 2$, which is a single linear equality constraint on $r_h[k]$, is the consequence of the first equation in (36) and the orthonormality condition (1) on the scaling filter.

3) *Reformulations:* To complete the reformulation of Problem 3 in terms of $r_h[k]$ instead of $h[\ell]$, we must add the additional semi-infinite linear inequality constraint $R_h(\omega) \geq 0$ for all $\omega \in [0, \pi]$, which is a necessary and sufficient condition for $r_h[k]$ to be factorizable in the form $r_h[k] = \sum_{\ell} h[\ell]h[\ell+k]$, [34]. Combining the above results, our design problem can be cast as follows.

Formulation 1: Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$ and integers $L \geq 2$, $0 < N \leq L/2$, and $0 \leq M < N/2$, find the autocorrelation sequence $r_h[k]$ achieving

$$\min_{r_h[k], \lambda} \sum_{k=1}^{L-1} (-1)^k b_1[k] r_h[k] + \beta_N(f) \lambda \quad (40a)$$

over $r_h[k]$, $0 \leq k \leq L-1$ and λ subject to the following constraints:

$$r_h[2k] = \delta[k] \quad \text{for } k = 0, 1, \dots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (40b)$$

$$\sum_{k=1}^{L-1} (-1)^k r_h[k] = -\frac{1}{2} \quad (40c)$$

$$\sum_{k=1}^{L-1} (-1)^k k^{2m} r_h[k] = 0 \quad \text{for } m = 1, \dots, N-1 \quad (40d)$$

$$R_h(\omega) \geq 0 \quad \text{for all } \omega \in [0, \pi] \quad (40e)$$

$$R_h(\omega) \leq 2^{-N} \lambda (1 + \cos \omega)^N \quad \text{for all } \omega \in [0, \pi] \quad (40f)$$

$$0 < \lambda < 2^{2(N-M-1/2)}. \quad (40g)$$

The constraints (40c), (40d), and (40e) guarantee that the low-pass filter $H(\omega)$ has N multiplicity of zeros at $\omega = \pi$. The constraints (40b), (40e), (40f), and (40g) together ensure that the scaling function generated by $h[\ell]$ is orthonormal and has at least M continuous derivatives.

Formulation 1 has a linear objective and an infinite number of linear constraints [constraints (40e) and (40f) each generate one linear constraint for every $\omega \in [0, \pi]$]. Although one could attempt to solve this problem by discretizing constraints (40e) and (40f) and applying linear programming techniques to that discretized problem, such an approach is numerically awkward in practice because (40e) requires that $R_h(\omega) \geq 0$ for all ω , but (40d) requires that $R_h(\omega)$ has $2N$ zeros at $\omega = \pi$, that is, that $R_h(\omega) = 0$ and $(\partial^m R_h(\omega)/\partial \omega^m)|_{\omega=\pi} = 0$ for $1 \leq m \leq 2N$. This causes the formulation to become numerically ill-conditioned as N increases. This problem can be alleviated by enforcing (40d) analytically and reformulating the problem in terms of $r_q[k]$, which is the autocorrelation sequence of $q[\ell]$ in (4).² Using (4), we have that

$$R_h(\omega) = \left(\frac{1 + e^{-j\omega}}{2} \right)^N \left(\frac{1 + e^{j\omega}}{2} \right)^N R_q(\omega). \quad (41)$$

²While this paper was being finalized, some independent and concurrent work on a related design problem appeared in which this reformulation in terms of $r_q[k]$ was also used [35].

This is equivalent to

$$r_h[k] = 2^{-2N} \sum_{n=-N}^N \binom{2N}{n+N} r_q[k-n] \quad (42)$$

for $k = 0, 1, \dots, L-1$. (Recall that $r_h[k]$ and $r_q[k]$ are symmetric.) Therefore, the orthonormality constraint (40b) on the original filter $h[\ell]$ is now equivalent to

$$\sum_{n=-N}^N \binom{2N}{n+N} r_q[2k-n] = 2^{2N} \delta[k] \quad (43)$$

for $k = 0, 1, \dots, \lfloor (L-1)/2 \rfloor$. Let

$$m[k] = \begin{cases} \binom{2N}{k+N} & \text{if } k = 0, \pm 1, \dots, \pm N \\ 0 & \text{otherwise.} \end{cases}$$

Then, (43) can be rewritten as

$$m[2k] r_q[0] + \sum_{n=1}^{L_q-1} (m[2k+n] + m[2k-n]) r_q[n] = 2^{2N} \delta[k] \quad (44)$$

for $k = 0, 1, \dots, \lfloor (L-1)/2 \rfloor$. Now substituting (42) into the right-hand side of (31) yields

$$\frac{b_1[0]}{2} + \sum_{k=0}^{L_q-1} s[k] r_q[k] + \beta_N(f) \lambda \quad (45)$$

where

$$s[0] = 2^{-2N} \sum_{n=1}^{L-1} (-1)^n m[n] b_1[n] \quad (46)$$

$$s[k] = 2^{-2N} \sum_{n=1}^{L-1} (-1)^n (m[n+k] + m[n-k]) b_1[n] \quad (47)$$

for $k = 1, 2, \dots, L_q-1$, and we have used the symmetry of $r_q[k]$. Therefore, Formulation 1 can be restated in terms of $r_q[k]$ as follows.

Formulation 2: Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$ and integers $L \geq 2$, $0 < N < L/2$, and $0 \leq M < N/2$, set $L_q = L - N$, and find the autocorrelation sequence $r_q[k]$ of the filter $q[\ell]$ that achieves

$$\min_{r_q[k], \lambda} \sum_{k=0}^{L_q-1} s[k] r_q[k] + \beta_N(f) \lambda \quad (48a)$$

over $r_q[k]$, $n = 0, 1, \dots, L_q-1$, and λ , subject to the following constraints:

$$r_q[0] + 2 \sum_{k=1}^{L_q-1} r_q[k] = 2 \quad (48b)$$

$$\begin{aligned} m[2k] r_q[0] + \sum_{n=1}^{L_q-1} (m[2k+n] + m[2k-n]) r_q[n] \\ = 2^{2N} \delta[k] \quad \text{for } n = 0, 1, \dots, \left\lfloor \frac{L-1}{2} \right\rfloor \end{aligned} \quad (48c)$$

$$R_q(\omega) \geq 0 \quad \text{for all } \omega \in [0, \pi] \quad (48d)$$

$$R_q(\omega) \leq \lambda \quad \text{for all } \omega \in [0, \pi] \quad (48e)$$

$$0 < \lambda < 2^{2(N-M-1/2)}. \quad (48f)$$

Like Formulation 1, Formulation 2 has a linear objective and linear constraints. The difference is that (40b) and (40c) in Formulation 1 have been imposed analytically in Formulation 2, rather than numerically, and hence, Formulation 2 is far better conditioned than Formulation 1. However, Formulation 2 remains awkward to solve because the constraints in (48d) and (48e) are still semi-infinite. They produce one (linear) constraint on $r_q[k]$ for each relevant frequency. One way to handle semi-infinite constraints of this type is to *approximate* them using discretization techniques [16]. An alternative approach is to apply the positive real and bounded real lemmas (e.g., [33] and [36]–[38]) to *precisely* transform the semi-infinite constraints in (48d) and (48e) into finite linear matrix inequalities. In particular, (48d) holds if and only if there exists an $L_q \times L_q$ symmetric positive semidefinite matrix \mathbf{X}_1 such that

$$\sum_n [\mathbf{X}_1]_{n,n+k} = r_q[k] \quad 0 \leq k \leq L_q - 1 \quad (49)$$

and (48e) holds if and only if there exists an $L_q \times L_q$ symmetric positive semidefinite matrix \mathbf{X}_2 such that

$$\sum_n [\mathbf{X}_2]_{n,n+k} = \lambda \delta[k] - r_q[k] \quad 0 \leq k \leq L_q - 1. \quad (50)$$

A symmetric matrix is said to be positive semidefinite if all its eigenvalues are non-negative. This will be denoted by $\mathbf{X} \succeq \mathbf{0}$. Using these two relationships, we obtain the following semidefinite program formulation of Problem 3.

Formulation 3: Given a particular signal $f(t)$ that is bandlimited to $[-\pi, \pi]$ and integers $L \geq 2$, $0 < N \leq L/2$, and $0 \leq M < N/2$, set $L_q = L - N$, and find the autocorrelation sequence $r_q[k]$, scalar $\lambda > 0$, and the $L_q \times L_q$ positive semidefinite symmetric matrices \mathbf{X}_1 and \mathbf{X}_2 , which achieve

$$\min_{r_q[k], \lambda, \mathbf{X}_1, \mathbf{X}_2 \succeq \mathbf{0}} \sum_{k=0}^{L_q-1} s[k] r_q[k] + \beta_N(f) \lambda \quad (51a)$$

$$\text{subject to (48b), (48c), (48f), (49), and (50).} \quad (51b)$$

Formulation 3 consists of a linear objective function, subject to linear equality constraints [(48b), (48c), (49), and (50)], linear inequalities (48f), and the semidefinite constraints $\mathbf{X}_1, \mathbf{X}_2 \succeq \mathbf{0}$. Hence, it is a (convex) semidefinite programming (SDP) problem [39] and can be solved in a highly efficient manner using interior point methods [40]. Several generic SDP solvers are available, including the MATLAB-based SeDuMi package [41]. (There are also early indications [38], [42], [43] that special-purpose SDP solvers that exploit the specific algebraic structure of (51) may be able to solve Formulation 3 substantially faster than a generic solver.) Once the optimal autocorrelation $r_q[k]$ has been found, we can find an optimal $q[\ell]$ using standard spectral factorization techniques [31], [44]. We can then find an optimal $h[\ell]$ using the following time domain equivalent of (4):

$$h[\ell] = 2^{-N} \sum_{n=0}^N \binom{N}{n} q[\ell - n]. \quad (52)$$

TABLE I
PERFORMANCE OF OUR DESIGN METHOD IN EXAMPLE 1 FOR VARIOUS N AND M . IN THIS EXAMPLE $L = 20$, AND WE PROVIDE THE PERCENTAGE REDUCTION IN THE SQUARE ROOT OF SQUARED PROJECTION ERROR OVER THE STANDARD DAUBECHIES LENGTH 20 FILTER

N	M	Optimized objective (33a)	Square root of achieved squared error, $\sqrt{\mathcal{E}_0(f_1, \varphi)}$	% Improvement
2	0	0.0563	0.1679	29%
4	0	0.0337	0.1772	26%
	1	0.0335	0.1777	25%
6	0	0.0373	0.1910	20%
	1	0.0375	0.1920	19%
	2	0.0422	0.2045	14%
8	0	0.0700	0.2030	15%
	1	0.0587	0.2045	14%
	2	0.0555	0.2150	9%

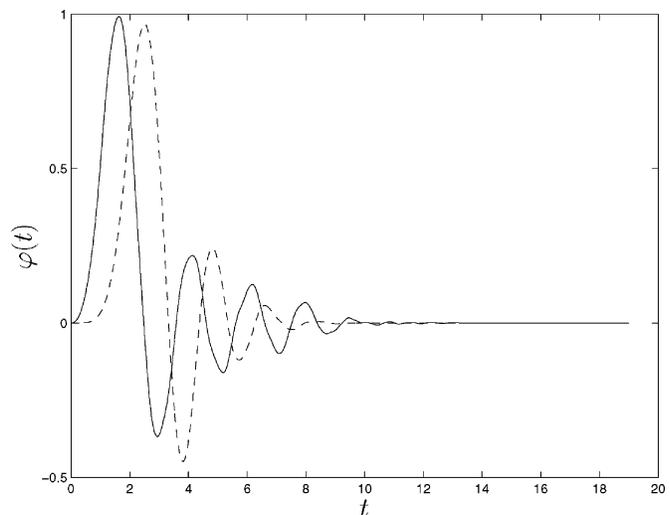


Fig. 1. Our scaling function for $f_1(t)$ (solid) and that generated by the standard length-20 Daubechies filter (dashed).

C. Bound on the Approximation Error

Our efficient design method for signal-adapted wavelet design is based on minimizing an upper bound (31) on the objective $\mathcal{E}_0(f, \varphi)$ in Problem 1. The following proposition, which is proved in Appendix E, shows that the value of $\mathcal{E}_0(f, \varphi)$ achieved by solving Problem 3 is guaranteed to be close to the true optimal value.

Proposition 2: Let $\varphi_1^*(t)$ denote an optimal solution to Problem 1, and let $\varphi_3^*(t)$ denote the scaling function corresponding to an optimal solution to Problem 3. Then

$$0 \leq \mathcal{E}_0(f, \varphi_3^*) - \mathcal{E}_0(f, \varphi_1^*) \leq \frac{M_N(f)}{4^{N+M+1}(4^N - 1)} \quad (53)$$

where N is the number of vanishing moments of the corresponding wavelet function.

Proposition 2 states that if N is chosen so that the right-hand side of (53) is small, then the solution to Problem 3 achieves a value of $\mathcal{E}_0(f, \varphi)$ that is close to the true optimal value. Furthermore, as long as the signal $f(t)$ is such that $\lim_{N \rightarrow \infty} M_N(f)/16^N = 0$, the bound on the right-hand side of (53) approaches zero as N increases. For such signals, a solution to Problem 3 is asymptotically optimal for Problem 1 as N increases. The condition $\lim_{N \rightarrow \infty} M_N(f)/16^N = 0$

TABLE II
OUR $h[\ell]$ FOR $f_1(t)$ WITH $L = 20$, $N = 4$ AND $M = 0$, ALONG WITH THE
STANDARD DAUBECHIES LENGTH 20 SCALING FILTER

Our $h[\ell]$	Daubechies $h[\ell]$
0.1366	0.0267
0.5159	0.1882
0.7011	0.5272
0.3162	0.6883
-0.1680	0.2812
-0.2131	-0.2498
0.0534	-0.1959
0.1476	0.1274
-0.0245	0.0931
-0.1048	-0.0714
0.0141	-0.0295
0.0820	0.0332
-0.0202	0.0036
-0.0588	-0.0107
0.0309	0.0014
0.0237	0.0020
-0.0221	-0.0007
-0.0010	-0.0001
0.0058	0.0001
-0.0015	-0.0000

is satisfied for a broad class of signals. In particular, for (bandlimited) signals with bounded power spectra, i.e., for signals for which $\hat{f}_m = \max_{\omega \in [-\pi, \pi]} |\hat{f}(\omega)| < \infty$, we have that

$$0 \leq M_N(f) \leq \frac{\hat{f}_m^2}{2\pi} \int_{-\pi}^{\pi} \omega^{2N} d\omega = \frac{\hat{f}_m^2 \pi^{2N}}{2N+1}$$

with equality holding if and only if $|\hat{f}(\omega)| \equiv \hat{f}_m$. Hence, for this (large) class of signals, the right-hand side of (53) decays to zero as N gets large.

Since $N \leq L/2$, solutions to Problem 3 remain suboptimal for Problem 1 for finite length filters. However, for a given ϵ , one can choose N such that

$$\frac{M_N(f)}{4^{N+M+1}(4^N - 1)} \leq \epsilon \quad (54)$$

and, hence, that the value of $\mathcal{E}_0(f, \varphi)$ achieved by a solution to Problem 3 is sufficiently close to its minimal value. The expression in (54) also provides some guidance as to the signals for which the objective in Problem 1 can be tightly bound. Recall that $M_N(f)$ is the $2N$ th moment of the power spectrum of $f(t)$, and, as such, is a measure of the extent to which $\hat{f}(\omega)$ is concentrated around $\omega = 0$. Functions with spectra that are concentrated around $\omega = 0$ will tend to have smaller moments and, hence, will tend to require only small values of N to satisfy (54). In contrast, functions with spectra that are dispersed across the whole band $[-\pi, \pi]$ will tend to have larger moments and, hence, may require larger N . Recall that all previous design methods also approximate the objective in Problem 1. A feature of our method is that the approximation error can be explicitly controlled.

IV. DESIGN EXAMPLES

In the section, we demonstrate the performance, flexibility, and efficiency of our design method using three examples.

Example 1: We first consider the signal $f_1(t) = \sin \pi t / \pi t$, which has a rectangular spectrum (of unit ‘‘height’’) on $[-\pi, \pi]$.

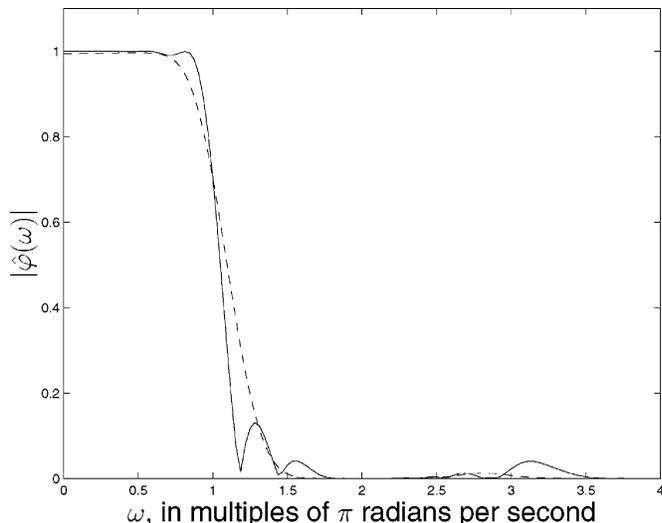


Fig. 2. Magnitude spectra of scaling functions in Fig. 1. The solid line shows the magnitude spectrum of our scaling function for $f_1(t)$, and the dashed line shows that of the scaling function generated by the standard Daubechies length 20 filter.

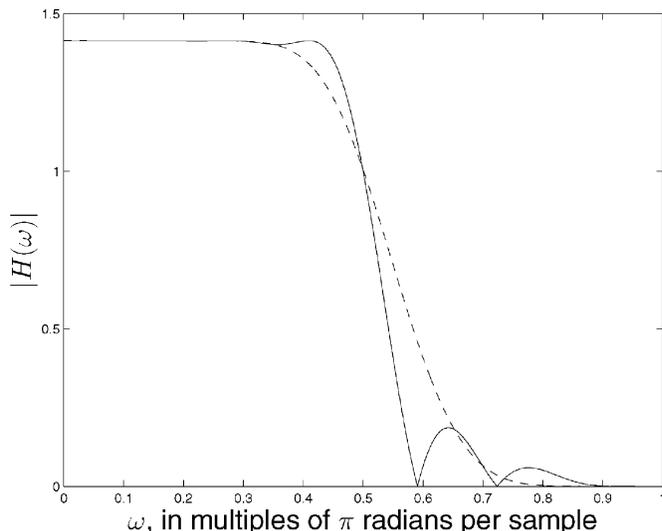


Fig. 3. Frequency responses of the scaling filters in Example 1. The solid line shows the frequency response of our scaling filter for $f_1(t)$, and the dashed line shows that of the standard length 20 Daubechies scaling filter.

As pointed out in Section III-C, this function results in an instance of Problem 1, which is quite difficult to approximate using our method. The performance of a number of length 20 filters designed by our method is compared with that of the standard length 20 Daubechies scaling filter, which has $N = 10$, in Table I. (In the table, $\mathcal{E}_0(f, \varphi)$ was approximated by $\hat{\mathcal{E}}_0^{(K)}(f, h)$ in (15) with $K = 10$.) As can be seen from Table I, our method can provide a substantial reduction in the projection error $\mathcal{E}_0(f, \varphi)$, when compared with the standard Daubechies design. Moreover, our design technique is flexible and efficient. As shown in Table I, it lends itself to the investigation of the tradeoffs between the number of vanishing moments N , the number of derivatives that are guaranteed to be continuous M , and the (achieved) projection error.

To study the characteristics of filters designed by our method in more detail, we select the case where $N = 4$ and $M = 0$,

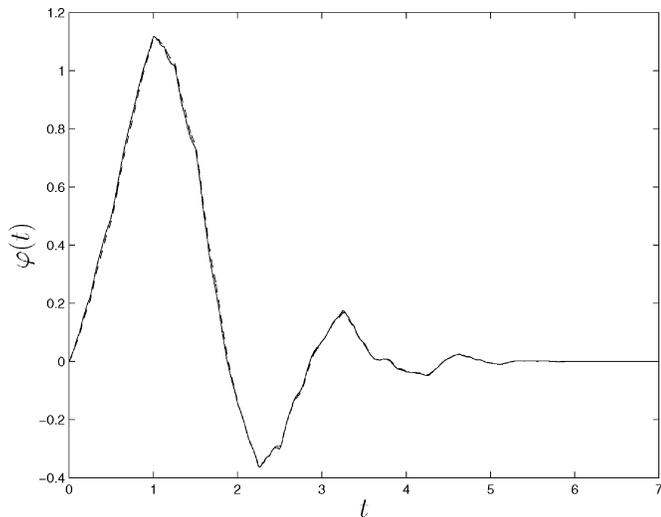


Fig. 4. Scaling functions for $f_2(t)$. The solid line shows our designed scaling function, and the dashed line shows the scaling function designed in [11]. These functions are almost indistinguishable at the scale of the figure.

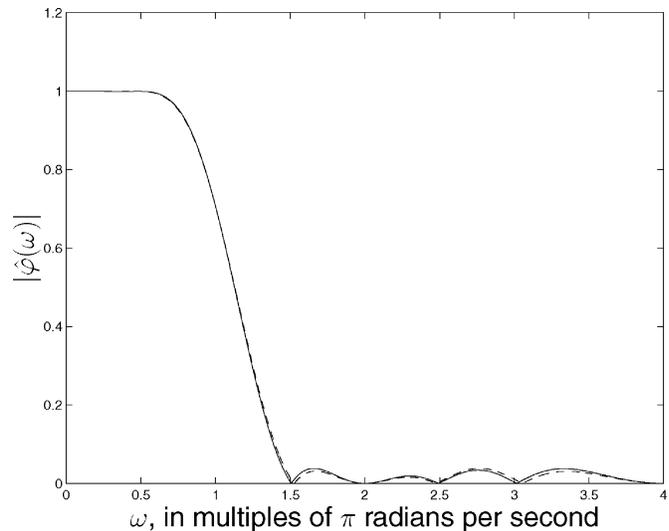


Fig. 5. Magnitude spectra of scaling functions in Fig. 4. The solid line shows the magnitude spectrum of our designed scaling function for $f_2(t)$, and the dashed line shows that of the scaling function designed in [11].

which results in a 26% reduction in the value of $\sqrt{\mathcal{E}_0(f, \varphi)}$ over the Daubechies filter. As can be seen in Fig. 1, the scaling function generated by our filter has a similar shape to that generated by the Daubechies filter, even though the coefficients of the filters are substantially different; see Table II. However, the differences between scaling functions are quite clear in the frequency domain, as shown in Fig. 2. The differences are perhaps most clear when plotting the frequency response of the scaling filters, as we have done in Fig. 3, in which it is clear that our filter has used its extra degrees of freedom (it has $N = 4$, whereas the Daubechies filter has $N = 10$) to provide a narrower transition band than the Daubechies filter. It has done so by placing two extra zeros in its frequency response (at $\omega \approx 0.59\pi$ and $\omega \approx 0.72\pi$) in addition to those at $\omega = \pi$. The tradeoff for this narrower band is a peak in the passband and slightly higher side-lobes in the stopband of the filter. The improved transition band of our filter clearly manifests itself in the improved transition performance of our scaling function (see Fig. 2). \square

Example 2: In this example, we demonstrate the accuracy of the approximation that led to our convex design method (Problem 3) by comparing the performance our method to that of Gopinath’s method [11]. Gopinath’s method involves solving the nonconvex optimization problem in Problem 2 in the absence of (11). We have chosen the signal from Example 2 in [11] as the signal of interest:

$$f_2(t) = \cos\left(\frac{\pi t}{2}\right) + \frac{2}{5} \cos\left(\frac{\pi t}{4}\right) \quad \text{for } 0 \leq t \leq 511$$

and zero otherwise. This signal is almost, rather than absolutely, bandlimited, but it was chosen because a scaling filter designed using Gopinath’s method [11] is available. To compare our method with Gopinath’s, we will consider a length 8 filter with two vanishing moments, i.e., $N = 2$ and $L = 8$. Our scaling filter $h[\ell]$ and Gopinath’s scaling filter [11] are given in Table III, along with the corresponding projection errors. (The projection error for Gopinath’s filter reported in Table III is slightly different from that reported in [11] because we have used a higher sampling rate in the discretization of $f_2(t)$.) The

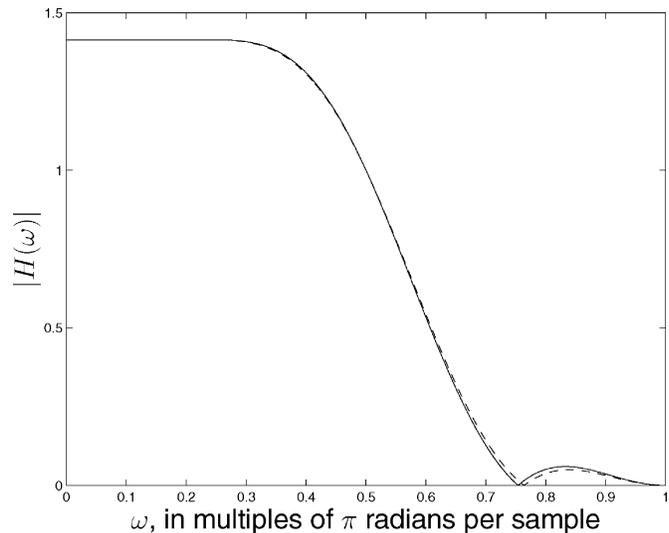


Fig. 6. Frequency responses of the scaling filters. The solid line shows the frequency response of our designed scaling filter for $f_2(t)$, and the dashed line shows that of the scaling filter designed in [11].

corresponding scaling functions and their magnitude spectra are shown in Figs. 4–6, respectively.

The analysis in Section III-C guarantees that the performance of our filter will approach the performance of an optimal filter as N grows. Even though we only have $N = 2$ in this example, it is clear from Table III that our scaling filter is very close to that obtained in [11]. As a result, the scaling functions and magnitude spectra are quite similar (see Figs. 4–6). The filter designed in [11] provides a slightly lower projection error, due to the more accurate approximation of $\mathcal{E}_0(f, \varphi)$ used in [11]. However, the design of that filter requires the solution of a nonconvex optimization problem. This typically involves the rather computationally costly task of generating a sufficiently large number of locally optimal solutions from sufficiently diverse class of “starting points” in order to have a reasonable degree of confidence that the best solution of

TABLE III
SCALING FILTERS FOR $f_2(t)$.

Our design		Design in [11]	
$h[\ell]$	$\sqrt{\mathcal{E}_0(f_2, \varphi)}$	$h[\ell]$	$\sqrt{\mathcal{E}_0(f_2, \varphi)}$
0.3145	0.1916	0.3056	0.1901
0.7481		0.7455	
0.5388		0.5488	
-0.0583		-0.0560	
-0.2048		-0.2032	
0.0419		0.0405	
0.0587		0.0558	
-0.0247		-0.0229	

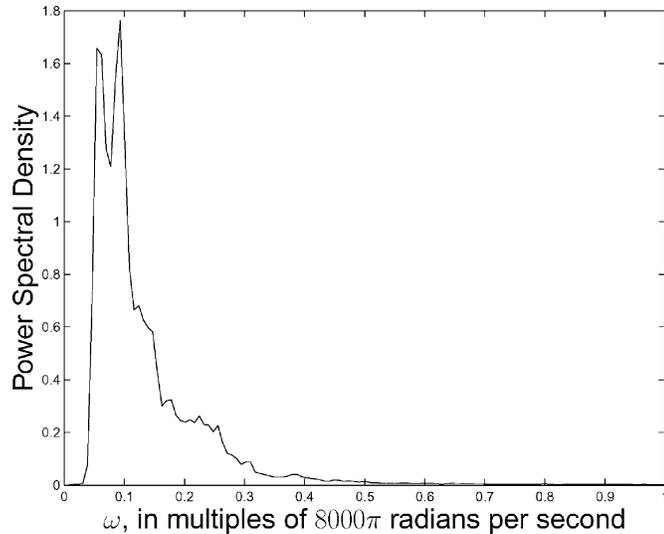


Fig. 7. (Time-averaged) power spectral density of the speech signal used in Example 3.

those obtained is sufficiently close to the global optimum. In contrast, any locally optimal solution to our formulation is globally optimal, and such solutions can be efficiently obtained using, for example, interior point methods. \square

Example 3: In this example, we apply our design method to a signal consisting of several English sentences spoken by a female speaker. The (time-averaged) power spectral density of the speech signal is shown in Fig. 7. By invoking the version of our design method for a stochastic signal [see (12)], we designed a scaling filter for this signal with $L = 12$ and $N = 4$. Our filter and the root mean square (RMS) value of the projection error that it achieves are provided in Table IV, and the resulting scaling function is provided in Fig. 8. In both Table IV and Fig. 8, our filter is compared with the standard Daubechies filter of length 12. Our filter achieves an RMS projection error that is more than 15% smaller than that achieved by the Daubechies filter. It can be seen from Fig. 8 that the shapes of our scaling function and the Daubechies scaling function are quite similar in the time domain, which is an observation we also made in Example 1. However, in the frequency domain, there are significant differences between our scaling function and that of Daubechies (see Fig. 9). As was the case in Example 1, the differences are most obvious in the frequency response of the scaling filter (see Fig. 10). Our scaling filter has a zero in the middle of its stopband, whereas the Daubechies filter has all its zeros at $\omega = \pi$. As a result, our scaling filter provides a faster transition between its

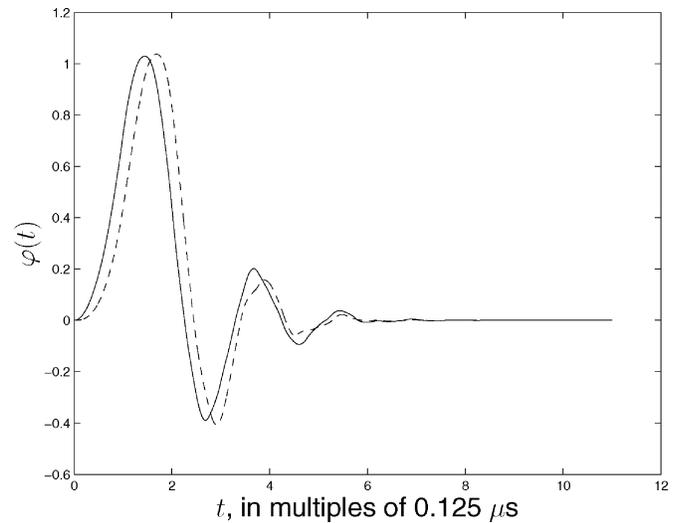


Fig. 8. Our scaling function for the speech signal in Example 3 (solid) and that generated by the standard length-12 Daubechies filter (dashed).

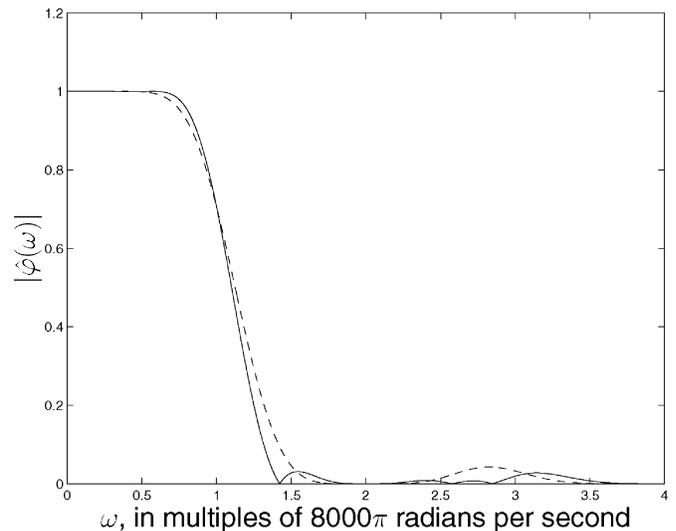


Fig. 9. Magnitude spectra of the scaling functions in Fig. 8. The solid line shows the magnitude spectrum of our scaling function for the speech signal in Example 3, and the dashed line shows that of the function generated by the standard length 12 Daubechies filter.

passband and stopband. As can be seen from Fig. 9 and Table IV, this leads to improved transition behavior in the spectrum of the corresponding scaling function and a reduced mean square projection error. \square

V. CONCLUSIONS

In this paper, we provided a flexible, efficient design technique to find compactly supported wavelet bases that are “matched” to a given bandlimited signal in a least square sense. This design technique complements the efficient design techniques that are currently available for “matching” filterbanks to signals with particular properties. All conventional approaches to this problem involve approximations because the objective function is only an implicit function of the parameters. An important advantage of our approximation over previous approximations is that the resulting design problem

Now, if we recursively apply the above procedure to $\Phi(\omega/2)$ on the right side of (62) K times, we have that

$$\Phi(\omega) = \prod_{k=1}^K \frac{1}{\sqrt{2}} H\left(\frac{\omega}{2^k}\right) \Phi\left(\frac{\omega}{2^K}\right) + \rho_K(\omega) \quad (64)$$

where $\rho_K(\omega) = \sum_{k=0}^{K-1} \epsilon_k(\omega)$ with $\epsilon_k(\omega)$ satisfying

$$|\epsilon_k(\omega)| \leq \frac{Q_m \Phi_m}{\sqrt{2}} \left| \frac{\omega}{4} \right|^N \left(\frac{H_m}{2^N \sqrt{2}} \right)^k. \quad (65)$$

Now, we see that if $H_m < 2^N \sqrt{2}$, then the series $\sum_{k=0}^{\infty} (H_m/2^N \sqrt{2})^k$ converges, and

$$\sum_{k=0}^{\infty} \left(\frac{H_m}{2^N \sqrt{2}} \right)^k = \frac{2^N \sqrt{2}}{2^N \sqrt{2} - H_m}. \quad (66)$$

Therefore, the series $\sum_{k=0}^{\infty} \epsilon_k(\omega)$ is absolutely and uniformly convergent. Let $\rho(\omega) = \lim_{K \rightarrow +\infty} \rho_K(\omega)$. Then, combining (65) with (66) gives

$$|\rho(\omega)| \leq \frac{Q_m \Phi_m}{\sqrt{2} - 2^{-N} H_m} \left| \frac{\omega}{4} \right|^N. \quad (67)$$

Finally, note that the first term in (64) tends to $\hat{\varphi}(\omega)$ as K goes to infinity since $\Phi(0) = 1$, cf. (13) and the sentence preceding (13). This completes the proof of Lemma 2. ■

The proof of Proposition 1 is as follows: Since $h[\ell]$ is orthonormal, we have that $|H(\omega)|^2 + |H(\pi + \omega)|^2 = 2$, [30]. Since $h[\ell]$ satisfies the conditions of Lemma 1, $H(\pi) = 0$, and hence, $H_m = \sqrt{2}$. Equation (18) is then obtained by substituting $H_m = \sqrt{2}$ into (57).

Since $h[\ell]$ satisfies the conditions of Lemma 1, $\varphi(t)$ is orthonormal, and hence

$$\varphi_r(\tau) = \int_{-\infty}^{\infty} \varphi(t) \varphi(t - \tau) dt \quad (68)$$

is an interpolating scaling function [45], i.e., $\varphi_r(k) = \delta[k]$. Therefore, for all real ω , $\Phi_r(\omega) = \sum_k \varphi_r(k) e^{-j\omega k} = 1$. Now, $\hat{\varphi}_r(\omega) = |\hat{\varphi}(\omega)|^2$. Applying Lemma 2 to $\varphi_r(t)$ completes the proof of Proposition 1.

APPENDIX C DERIVATION OF (30)

By taking the continuous-time Fourier transform of the right-hand side of (27), we have that

$$\hat{f}(\omega) = \begin{cases} \frac{1}{2} \sum_k f\left(\frac{k}{2}\right) e^{-j\omega k/2}, & |\omega| \leq 2\pi \\ 0, & \text{otherwise} \end{cases} \quad (69)$$

and, hence, that

$$|\hat{f}(\omega)|^2 = \begin{cases} \frac{1}{2} \sum_n b_1[n] e^{-j\omega n/2}, & |\omega| \leq 2\pi \\ 0, & \text{otherwise.} \end{cases} \quad (70)$$

The above formulae remain valid even though we know that $\hat{f}(\omega) = 0$ for $|\omega| > \pi$.

We also have that

$$H\left(\frac{\omega}{2} + \pi\right) = \sum_{\ell} (-1)^{\ell} h[\ell] e^{-j\omega \ell/2} \quad (71)$$

and, hence, that

$$\left| H\left(\frac{\omega}{2} + \pi\right) \right|^2 = \sum_m (-1)^m r_h[m] e^{-j\omega m/2}. \quad (72)$$

Since we know that $\hat{f}(\omega) = 0$ for $|\omega| > \pi$

$$\int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \left| H\left(\frac{\omega}{2} + \pi\right) \right|^2 d\omega = \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \left| H\left(\frac{\omega}{2} + \pi\right) \right|^2 d\omega = 2\pi \sum_m (-1)^m r_h[m] b_1[m] \quad (73)$$

where, in the last step, we have used (70), (72), and the fact that $\int_{-\pi}^{\pi} e^{-j\omega(m-n)/2} d\omega = 4\pi \delta[n-m]$. Equation (30) then follows by substituting (73) into (23).

APPENDIX D CALCULATION OF $M_N(f)$

Using (25), we have that

$$M_N(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega^{2N} |\hat{f}(\omega)|^2 d\omega. \quad (74)$$

Since $f(t)$ is bandlimited to $[-\pi, \pi]$, we have that $f(t) = \sum_{k \in \mathbb{Z}} f(k) (\sin(\pi(t-k)) / \pi(t-k))$. Hence

$$|\hat{f}(\omega)|^2 = b_0[0] + 2 \sum_{k \geq 1} b_0[k] \cos(k\omega)$$

where $b_j[k]$ was defined in (29). In other words, the power spectrum of $f(t)$ can be written in terms of the autocorrelation of its ‘‘Nyquist-rate’’ samples. Hence

$$\begin{aligned} M_N(f) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega^{2N} |\hat{f}(\omega)|^2 d\omega \\ &= \frac{b_0[0]}{2\pi} \int_{-\pi}^{\pi} \omega^{2N} d\omega \\ &\quad + \frac{1}{\pi} \sum_{k \geq 1} b_0[k] \int_{-\pi}^{\pi} \omega^{2N} \cos k\omega d\omega \\ &= \frac{\pi^{2N} b_0[0]}{2N+1} \\ &\quad + 2 \sum_{\ell=1}^N (-1)^{\ell-1} \pi^{2(N-\ell)} P_{2N-1}^{2\ell-1} \sum_{k=1}^{+\infty} (-1)^k \frac{b_0[k]}{k^{2\ell}}. \end{aligned}$$

APPENDIX E PROOF OF PROPOSITION 2

To simplify the notation, we define

$$\mathcal{L}(f, h) = \frac{b_1[0]}{2} + \sum_{k=1}^{L-1} (-1)^k b_1[k] r_h[k]$$

where $b_j[k]$ was defined in (29), and $r_h[k]$ is the autocorrelation of $h[\ell]$, cf. (28). Using (24), we have that

$$\mathcal{E}_0(f, \varphi) = \mathcal{L}(f, h) + \Delta(f, \varphi) \leq \mathcal{L}(f, h) + \beta_N(f) Q_m^2 \quad (75)$$

where $Q_m = \max_{\omega \in [-\pi, \pi]} |Q(\omega)|$, and $Q(\omega) = 2^N (1 + e^{-j\omega})^{-N} H(\omega)$. If we let $h_1^*[\ell]$ denote an optimal scaling filter

for Problem 1 and let $h_3^*[\ell]$ and λ_3^* denote an optimal solution to Problem 3, then

$$\begin{aligned} \mathcal{E}_0(f, \varphi_3^*) - \mathcal{E}_0(f, \varphi_1^*) \\ \leq \mathcal{L}(f, h_3^*) - \mathcal{L}(f, h_1^*) + \beta_N(f)\lambda_3^* - \Delta(f, \varphi_1^*). \end{aligned} \quad (76)$$

Here

$$\Delta(f, \varphi) = \frac{1}{4\pi} \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \left| H\left(\frac{\omega}{2}\right) \right|^2 \left(1 - \left| \hat{\varphi}\left(\frac{\omega}{2}\right) \right|^2 \right) d\omega. \quad (77)$$

Since $\varphi(t)$ is orthonormal, we have that $\sum_k |\hat{\varphi}(\omega + 2\pi k)|^2 = 1$, [24, p. 132]. Hence, $|\hat{\varphi}(\omega)|^2 \leq 1$, and therefore, $\Delta(f, \varphi) \geq 0$. Using this result, we have that

$$\mathcal{E}_0(f, \varphi_3^*) - \mathcal{E}_0(f, \varphi_1^*) \leq \mathcal{L}(f, h_3^*) - \mathcal{L}(f, h_1^*) + \beta_N(f)\lambda_3^*. \quad (78)$$

Now, since $h_3^*[\ell]$ and λ_3^* are optimal with respect to Problem 3, we have that

$$\mathcal{L}(f, h_3^*) + \beta_N(f)\lambda_3^* \leq \mathcal{L}(f, h_1^*) + \beta_N(f)(Q_{1,m}^*)^2 \quad (79)$$

and hence

$$\mathcal{L}(f, h_3^*) - \mathcal{L}(f, h_1^*) \leq \beta_N(f)(Q_{1,m}^*)^2 - \beta_N(f)\lambda_3^*. \quad (80)$$

Combining (78) and (80), we have that

$$\mathcal{E}_0(f, \varphi_3^*) - \mathcal{E}_0(f, \varphi_1^*) \leq \beta_N(f)(Q_{1,m}^*)^2. \quad (81)$$

The proposition now follows by observing that for any scaling filter satisfying (4) and (11), $|Q(\omega)| \leq 2^{N-M-1/2}$.

ACKNOWLEDGMENT

The first author would like to thank Prof. B. Zheng for his careful guidance at Xidian University, Xi'an, China. This work is an extension of his Ph.D. dissertation.

REFERENCES

[1] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge, 1996.
 [2] A. Rosenfield, *Multiresolution Techniques in Computer Vision*. New York: Springer-Verlag, 1984.
 [3] M. Frisch and H. Messer, "Detection of transient signal of unknown scaling and arrival time using the discrete wavelet transform," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Toronto, ON, Canada, May 1991, pp. 1313–1316.
 [4] S. Mallat, "Zero-crossings of a wavelet transform," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1019–1033, July 1991.
 [5] S. Kadambe and G. F. Boudraux-Bartels, "Application of wavelet transform for pitch detection of speech signals," *IEEE Trans. Inform. Theory*, vol. 38, pp. 917–924, Mar. 1992.
 [6] D. M. Healy and J. B. Weaver, "Two applications of wavelet transforms in magnetic resonance imaging," *IEEE Trans. Inform. Theory*, vol. 38, pp. 840–860, Mar. 1992.
 [7] G. W. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proc. IEEE*, vol. 84, pp. 586–603, Apr. 1996.
 [8] K. M. Wong, J. Wu, T. N. Davidson, and Q. Jin, "Wavelet packet division multiplexing and wavelet packet design under timing error effects," *IEEE Trans. Signal Processing*, vol. 45, pp. 2877–2890, Dec. 1997.
 [9] M. G. Strintzis, "Optimal biorthogonal wavelet bases for signal decomposition," *IEEE Trans. Signal Processing*, vol. 44, pp. 1406–1417, June 1996.
 [10] A. H. Tewfik, D. Sinha, and P. Jorgensen, "On the optimal choice of a wavelet for signal representation," *IEEE Inform. Theory*, vol. 38, pp. 747–765, Feb. 1992.

[11] R. A. Gopinath, J. E. Odegard, and C. S. Burrus, "Optimal wavelet representation of signal and the wavelet sampling theorem," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 262–277, Apr. 1994.
 [12] M. Unser, "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3591–3596, Dec. 1993.
 [13] M. K. Tsatsanis and G. B. Giannakis, "Principal component filter banks for optimal multiresolution analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1766–1777, Aug. 1995.
 [14] S. Ohno and H. Sakai, "Optimization of filter banks using cyclostationary spectral analysis," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Detroit, MI, May 1995, pp. 1292–1295.
 [15] P. Delsarte, B. Macq, and D. Slock, "Signal-adapted multiresolution analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1766–1777, Aug. 1995.
 [16] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Potra, "The role of linear semi-infinite programming in signal-adapted QMF bank design," *IEEE Trans. Signal Processing*, vol. 45, pp. 2160–2174, Sept. 1997.
 [17] P. Moulin and M. K. Mihcak, "Theory and design of signal-adapted FIR paraunitary filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 920–929, Apr. 1998.
 [18] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Processing*, vol. 46, pp. 1528–1543, June 1998.
 [19] —, "Results on optimal biorthogonal filter banks," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 932–947, Aug. 1998.
 [20] B. Xuan and R. H. Bamberger, "FIR principal component filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 930–940, Apr. 1998.
 [21] J. Tuqan and P. P. Vaidyanathan, "A state space approach to the design of globally optimal FIR energy compaction filters," *IEEE Trans. Signal Processing*, vol. 48, pp. 2822–2838, Oct. 2000.
 [22] S. Akkarakaran and P. P. Vaidyanathan, "Filterbank optimization with convex objectives and the optimality of principal component forms," *IEEE Trans. Signal Processing*, vol. 49, pp. 100–114, Jan. 2001.
 [23] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, pp. 569–587, Apr. 2000.
 [24] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
 [25] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
 [26] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998.
 [27] R. A. Gopinath and C. S. Burrus, "Wavelets and filter banks," in *Wavelets: A Tutorial in Theory and Application*, C. K. Chui, Ed. New York: Academic, 1992.
 [28] W. M. Lawton, "Necessary and sufficient conditions for constructing orthonormal wavelet bases," *J. Math. Phys.*, vol. 32, pp. 57–61, 1991.
 [29] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.
 [30] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 906–996, 1988.
 [31] S.-P. Wu, S. Boyd, and L. Vandenbergh, "FIR filter design via spectral factorization and convex optimization," in *Applied and Computational Control, Signal and Circuits*, B. Datta, Ed. Boston, MA: Birkhauser, 1998, vol. 1.
 [32] T. N. Davidson, Z. Q. Luo, and K. M. Wong, "Design of orthogonal pulse shapes for communications via semidefinite programming," *IEEE Trans. Signal Processing*, vol. 48, pp. 1433–1445, May 2000.
 [33] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, "Linear matrix inequality formulation of spectral mask constraints with applications to FIR filter design," *IEEE Trans. Signal Processing*, vol. 50, pp. 2701–2715, Nov. 2002.
 [34] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.
 [35] B. Dumitrescu and C. Popea, "Accurate computation of compaction filters with high regularity," *IEEE Signal Processing Lett.*, vol. 9, pp. 278–281, Sept. 2002.
 [36] L. Hitz and B. D. O. Anderson, "Discrete positive-real functions and their application to system stability," *Proc. Inst. Elect. Eng.*, vol. 116, pp. 153–155, Jan. 1969.
 [37] B. D. O. Anderson, K. Hitz, and N. D. Diem, "Recursive algorithm for spectral factorization," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 742–750, June 1974.
 [38] B. Dumitrescu, I. Tabus, and P. Stoica, "On the parameterization of positive real sequences and MA parameter estimation," *IEEE Trans. Signal Processing*, vol. 49, pp. 2630–2639, Nov. 2001.
 [39] L. Vandenbergh and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 31, pp. 49–95, Mar. 1996.
 [40] Y. Ye, *Interior Point Algorithms: Theory and Analysis*. New York: Wiley, 1997.

- [41] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Meth. Software*, vol. 11–12, pp. 625–653, 1999.
- [42] Y. Genin, Y. Hachez, Y. Nesterov, and P. Van Dooren, "Convex optimization over positive polynomials and filter design," in *Proc. UKACC Int. Conf. Contr.*, Cambridge, U.K., Sept. 2000.
- [43] B. Alkire and L. Vandenberghe, "Interior-point methods for magnitude filter design," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, UT, May 2001.
- [44] T. N. T. Goodman, C. A. Micchelli, G. Rodriguez, and S. Seatzu, "Spectral factorization of Laurent polynomials," *Adv. Comput. Math.*, vol. 7, pp. 429–454, Apr. 1997.
- [45] N. Saito and G. Beylkin, "Multiresolution representations using the auto-correlation functions of compactly supported wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 3584–3590, Dec. 1993.



Jian-Kang Zhang received the B.S. degree in mathematics from Shaanxi Normal University, Xi'an, China, the M.S. degree in mathematics from Northwest University, Xi'an, and the Ph.D. degree in electrical engineering from Xidian University, Xi'an.

He is now with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests include multirate filterbanks, wavelet and multiwavelet transforms and their applications, number theory transforms, and fast algorithms. His current research

focuses on multiscale wavelet transforms, random matrices, precoding and space-time coding for block transmissions, and multicarrier and wideband wireless communication systems.



Timothy N. Davidson (M'96) received the B.Eng. (Hons. I) degree in electronic engineering from The University of Western Australia (UWA), Perth, Australia, in 1991 and the D.Phil. degree in engineering science from the The University of Oxford, Oxford, U.K., in 1995.

He is currently an assistant professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests are in signal processing, communications, and control, with current activity focused

on signal processing for digital communication systems. He has held research positions at the Communications Research Laboratory at McMaster University, the Adaptive Signal Processing Laboratory at UWA, and the Australian Telecommunications Research Institute, Curtin University of Technology, Perth.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize (for "the most outstanding [UWA] graduand" in the pure and applied sciences) and the 1991 Rhodes Scholarship for Western Australia.



Kon Max Wong (F'02) was born in Macau. He received the B.Sc.(Eng), D.I.C., Ph.D., and D.Sc.(Eng) degrees, all in electrical engineering, from the University of London, London, U.K., in 1969, 1972, 1974, and 1995, respectively.

He was with the Transmission Division of Plessey Telecommunications Research Ltd., London, in 1969. In October 1970, he was on leave from Plessey, pursuing postgraduate studies and research at Imperial College of Science and Technology, London. In 1972, he rejoined Plessey as a research engineer and worked on digital signal processing and signal transmission. In 1976, he joined the Department of Electrical Engineering, Technical University of Nova Scotia, Halifax, NS, Canada, and in 1981, he moved to McMaster University, Hamilton, ON, Canada, where he has been a Professor since 1985 and served as Chairman of the Department of Electrical and Computer Engineering from 1986 to 1987 and again from 1988 to 1994. He was on leave as a Visiting Professor at the Department of Electronic Engineering, the Chinese University of Hong Kong, from 1997 to 1999. At present, he holds the title of NSERC-Mitel Professor of Signal Processing and is the Director of the Communication Technology Research Centre at McMaster University. His research interest is in signal processing and communication theory, and he has published over 170 papers in the area.

Prof. Wong was the recipient of the IEE Overseas Premium for the best paper in 1989, is a Fellow of the Institution of Electrical Engineers, a Fellow of the Royal Statistical Society, and a Fellow of the Institute of Physics. He also served as an Associate Editor of the IEEE TRANSACTION ON SIGNAL PROCESSING from 1996 to 1998 and has been the chairman of the Sensor Array and Multichannel Signal Processing Technical Committee of the Signal Processing Society since 1998. He received a medal presented by the International Biographical Centre, Cambridge, U.K., for his "outstanding contributions to the research and education in signal processing" in May 2000 and was honored with the inclusion of his biography in the books *Outstanding People of the 20th Century* and *2000 Outstanding Intellectuals of the 20th Century*, which were published by IBC to celebrate the arrival of the new millennium.