Positive Sampling in Wavelet Subspaces

Gilbert G. Walter
Department of Mathematics, University of Wisconsin, Milwaukee, Wisconsin 53201

and

Xiaoping A. Shen
Department of Mathematics, Ohio University, Athens, Ohio 45701

Communicated by Dominique Picard

Received February 2, 2001; revised July 19, 2001

This paper is devoted to the discussion of a “hybrid” sampling series, a series of translates of a nonnegative summability function used in place of an orthogonal scaling function. The coefficients in the series are taken to be sampled values of the function to be approximated. This enables one to avoid the integration which arises in the other series. The approximations based on this hybrid series have certain desirable convergence properties: they are locally uniformly convergent for locally continuous functions, they have quadratic uniform convergence rate for functions in certain Sobolev spaces, they are locally bounded when the function is locally bounded and therefore, in particular, Gibbs’ phenomenon is avoided. Numerical experiments are given to illustrate the theoretical results and to compare these approximations with the scaling function approximations.

1. INTRODUCTION

The approximation of a continuous function by its expansion in a scaling function series has a number of advantages over other series expansions. For example, these scaling function series always converge uniformly to the function (Kelly, Kon, and Raphael [3], Walter [5]). However, these approximations

(a) may not give nonnegative approximations to nonnegative functions,
(b) may oscillate excessively, and
(c) have coefficients that at the finest scales involve computation of an integral.

The failure of nonnegativity as well as the excessive oscillation can be avoided by using a nonnegative summability function in place of an orthogonal scaling function as in Walter and Shen [7]. In this work, we shall try to avoid the integration as well, while preserving the other two properties. We use the same summability function in a “hybrid” sampling
series in which the coefficients are sampled values of the function rather than the usual ones. This avoids integration while maintaining certain desirable convergence properties and avoiding excessive oscillations. It will also enable us to find the coefficients at coarser scales by using a modified decomposition algorithm.

2. BACKGROUND

We present here a few elements of orthogonal wavelet theory, in which an orthonormal basis \( \{ \psi_{mn} \} \) of \( L^2(\mathbb{R}) \) is constructed having the form

\[
\psi_{mn}(t) = 2^{m/2} \psi_{mn}(2^m t - n), \quad n, m \in \mathbb{Z},
\]

where \( \psi(t) \) is the “mother wavelet.” Usually it is not constructed directly but rather from another function called the “scaling function” \( \phi(t) \in L^2(\mathbb{R}) \cap L^1(\mathbb{R}) \). The scaling function \( \phi \) is chosen in such a way that

(i) \[
\int_{-\infty}^{\infty} \phi(t) \phi(t - n) \, dt = \delta_{0,n}, \quad n \in \mathbb{Z},
\]

(ii) \[
\phi(t) = \sum_{k=-\infty}^{\infty} \sqrt{2} c_k \phi(2t - k), \quad \{c_k\}_{k \in \mathbb{Z}} \in l^2,
\]

(iii) For each \( f \in L^2(\mathbb{R}) \), \( \epsilon > 0 \), there is a function \( f_m(t) \) such that \( \| f_m - f \| < \epsilon \).

These conditions lead to a “multiresolution approximation” \( \{ V_m \}_{m \in \mathbb{Z}} \), consisting of closed subspaces of \( L^2(\mathbb{R}) \). The space \( V_0 \) is taken to be the closed linear span of \( \{ \phi(2^m t - n) \}_{n \in \mathbb{Z}} \). Because of (2.1) (ii), the \( V_m \) are nested, i.e., \( V_m \subseteq V_{m+1} \), and because of (2.1) (iii), \( \bigcup_m V_m \) is dense in \( L^2(\mathbb{R}) \).

Frequently, conditions (2.1) are expressed in terms of their Fourier transforms. We give a sufficient condition for (2.1) as

(i) \[
\sum_{k=-\infty}^{\infty} |\hat{\phi}(\omega + k)|^2 = 1,
\]

(ii) \[
\hat{\psi} (\omega) = \left( \frac{1}{\sqrt{2}} \sum_{k=-\infty}^{\infty} c_k e^{i k \omega / 2} \right) \hat{\phi} \left( \frac{\omega}{2} \right) = m_0 \left( \frac{\omega}{2} \right) \hat{\phi} \left( \frac{\omega}{2} \right),
\]

where \( m_0 \left( \frac{\omega}{2} \right) = \frac{1}{\sqrt{2}} \sum_{k=-\infty}^{\infty} c_k e^{i k \omega / 2} \in L^2(-2\pi, 2\pi) \),

(iii) \( \hat{\phi} (\omega) \) is continuous at \( \omega = 0 \) and \( \hat{\phi}(0) = 1 \).

The mother wavelet comes from \( \phi(t) \) via (2.1) (ii) or (2.2) (ii),

(i) \[
\psi(t) = \sum_{k=-\infty}^{\infty} (-1)^k \sqrt{2} c_{1-k} \phi(2t - k),
\]

(ii) \[
\hat{\psi} (\omega) = e^{-i \omega / 2} m_0 \left( \frac{\omega}{2} + \pi \right) \hat{\phi} \left( \frac{\omega}{2} \right).
\]
The Haar orthogonal system is based on a scaling function which has compact support and is nonnegative. Many other wavelets with compact support, beginning with those of Daubechies [1], have been constructed. These are continuous, which the Haar system is not, but are never nonnegative (Janssen [2]). To find nonnegative estimation in \( V_0 \), we first ask if there are any nonnegative functions in \( V_0 \), and then if there is a basis of such functions.

Let \( \phi(t) \) be any continuous scaling function on \( \mathbb{R} \) satisfying (2.2) with support in a compact interval. We know that all such \( \phi(t) \) generate a partition of unity by (2.2) (iii):

\[
\sum_{n \in \mathbb{Z}} \phi(t - n) \equiv 1, \quad t \in \mathbb{R}. \tag{2.3}
\]

This is a property shared by the scaling functions due to Daubechies as well as the “Coiflets” of Coifman (Daubechies [1]). We define the Abel means of their series (2.3) to be

\[
\rho^r(t) := \sum_{n = -\infty}^{\infty} r^{\left| n \right|} \phi(t - n), \quad 0 < r \leq 1, \quad t \in \mathbb{R}. \tag{2.4}
\]

This series converges uniformly on \([0, 1]\) since it is locally finite and \( \rho^r(t) \to 1 \) for \( t \in [0, 1] \) as \( r \to 1 \). This result also holds uniformly on any finite interval \([M, N]\). Consequently, there exists a real number \( r_0 \), \( 0 < r_0 < 1 \), such that \( \rho^r(t) \geq 1/2 \) for \( 1 > r \geq r_0, t \in [M, N] \).

The properties of \( \rho^r(t) \) are given by the next lemma.

**Lemma 2.1.** Let \( \phi(t) \) be a continuous scaling function on \( \mathbb{R} \) with compact support satisfying (2.2); let \( V_0 = \text{CLS}\{\phi(t - n), n \in \mathbb{Z}\} \); then there is an \( 0 < r_0 < 1 \) such that \( \rho^r \) given by (2.4), for \( r_0 \leq r < 1 \), satisfies

(i) \( \rho^r(t) \geq 0, \quad t \in \mathbb{R}, \)

(ii) \( \rho^r \in V_0. \)

We can say much more than this. In fact, we have

**Theorem 2.1.** Let \( \rho^r(t) = \sum_n r^{\left| n \right|} \phi(t - n) \), where \( r \) is chosen so large that \( \rho^r(t) \geq 0 \), for \( t \in \mathbb{R} \); then \( \{2^{4n} \rho^r(2^n t - n)\}_{n \in \mathbb{Z}} \) is a Riesz basis of \( V_m \); its dual basis is generated by \( \tilde{\rho}_r(t) \), where \( \tilde{\rho}_r(t) = 1/2\pi(1 - r^2)(1 + r^2)\phi(t) - r[\phi(t + 1) + \phi(t - 1)] \).

The proofs of Lemma 2.1 and Theorem 2.2 can be found in [7]. However, this biorthogonal system does not give us the positive kernel we need. One of the modifications, used in [7],

\[
K_r(t, s) = \left( \frac{1 - r}{1 + r} \right)^2 \sum_n \rho^r(t - n) \rho^{r^*}(s - n),
\]

does give a positive kernel with the desired properties. In this case, the approximation in \( V_0 \) is given by

\[
f_{0, r}(t) = \int_{-\infty}^{\infty} K_r(t, s) f(s) \, ds.
\]
In this paper, we introduce another positive kernel based on the same summability function. It is defined by

$$G^r(t, s) = \frac{1 - r}{1 + r} \sum_n \delta(s - n) \rho^r(t - n), \quad (2.5)$$

where $\delta(t)$ is the delta distribution. It gives us the approximation in $V_0$ to a continuous $f \in L^2$ as

$$f_0^r(t) = \int_{-\infty}^{\infty} G^r(t, s) f(s) \, ds.$$ 

The summability function $\rho^r(t)$, as defined in Theorem 2.2, is no longer compactly supported. A dilation equation relating $\rho^r(t)$ and a series of $\{\rho^r(2t - n)\}$ may be found but is now an infinite sum. However, the series is rapidly convergent [8] and its Fourier transform is given by

$$\hat{\rho}^r(\omega) = Q_r(\omega) \hat{\phi}(\omega) = \frac{1 - r^2}{1 - 2r \cos \omega + r^2} \left( \frac{\omega}{2} \right)^r \hat{\phi} \left( \frac{\omega}{2} \right).$$ \quad (2.6)

Here, we use the identity $\hat{\phi}(\omega) = m_0(\frac{\omega}{2}) \hat{\phi}(\frac{\omega}{2})$.

This is the Fourier transformed version of the dilation equation for $\rho^r(t)$. By expanding the periodic factor in a Fourier series and taking the inverse Fourier transform, we can find the coefficients of the dilation equation in the time domain as follows:

$$a_k = \sum_{n=-\infty}^{\infty} \left( c_{k-2n} r^{|n|} \frac{1 + r^2}{1 - r^2} - c_{k-1-2n} r^{|n|+1} \frac{r^{n+1}}{1 - r^2} - c_{k+1-2n} r^{|n|+1} \frac{r^{n+1}}{1 - r^2} \right).$$ \quad (2.7)

Notice that, for every fixed $k$, (2.7) only contains a finite number of terms. We then have the dilation equation:

$$\rho^r(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} a_k \rho^r(2t - k).$$

Similarly, the biorthogonal function $\tilde{\rho}^r$ satisfies

$$\hat{\tilde{\rho}}^r(\omega) = \frac{1 - 2r \cos \omega + r^2}{1 - 2r \cos(\omega/2) + r^2} m_0(\frac{\omega}{2}) \hat{\rho}^r \left( \frac{\omega}{2} \right),$$ \quad (2.8)

and hence

$$\tilde{\rho}^r(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} \tilde{a}_k \tilde{\rho}^r(2t - k),$$

with coefficients given by

$$\tilde{a}_k = \sum_{n=-\infty}^{\infty} \left( c_{k-n} r^{|n|} \frac{1 + r^2}{1 - r^2} - c_{k-2-n} r^{|n|+1} \frac{r^{n+1}}{1 - r^2} - c_{k+2-n} r^{|n|+1} \frac{r^{n+1}}{1 - r^2} \right).$$ \quad (2.9)
The pair of biorthogonal mother wavelets can be defined by
\[
\xi_r(t) = \sqrt{2} \sum_{k} (-1)^k \tilde{a}_{1-k} \rho^r(2t-k), \tag{2.10}
\]
with its dual
\[
\xi_r(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} (-1)^k a_{1-k} \tilde{\rho}_r(2t-k), \tag{2.11}
\]
as given in [1, p. 263]. Then the Fourier transformed versions of (2.10) and (2.11) are
\[
\hat{\xi}_r(\omega) = e^{-i(\omega/2)Q_r(\omega)} \sqrt{(1 - 2r^2)/(1 - 2r \cos \omega + r^2)} \frac{Q_r(\omega/(2 + \pi))}{m_0(\omega/(2 + \pi))} \hat{\phi}(\omega/2),
\]
and
\[
\hat{\xi}_r(\omega) = \frac{(1 + r^2 - 2r^2 \cos \omega)}{(1 - 2r \cos \omega + r^2)} \hat{\psi}(\omega),
\]
respectively. Hence, by taking the inverse Fourier transform and using the Fourier series of the periodic factor, we have in the first case that
\[
\xi_r(t) = \sum_{n} r^{2|n|} \psi(t - n) - \frac{r}{1 + r^2} \sum_{n} r^{2|n|} [\psi(t - n - 1) + \psi(t - n + 1)]. \tag{2.12}
\]
Notice that both \(\xi_r(t)\) and \(\tilde{\xi}_r(t)\) belong to the wavelet subspace \(W_0\), the space with basis \(\{\psi(t - n)\}\) and each has as many vanishing moments as \(\psi\) has. Furthermore, it can be shown that each constitutes a Riesz basis of \(W_0\).

3. THE CONVERGENCE OF THE POSITIVE HYBRID SERIES

We assume that \(f\) is a bounded uniformly continuous function in \(L^1(\mathbb{R})\). We shall approximate it by a hybrid series in \(V_m\) given by
\[
f_m^r(t) = \sum_{n} f \left( n \left( \frac{2^m}{n} \right) \right) \left( \frac{2^m t - n}{1 + r} \right) \frac{1 - r}{1 + r^2}, \tag{3.1}
\]
This series can be expressed as
\[
f_m^r(t) = \int_{-\infty}^{\infty} G_m^r(t, s) f(s) \, ds,
\]
where $G_m^r$, the $2^m$ dilation of the kernel of (2.5), is given by

$$G_m^r(t, s) = 2^m \sum_n \delta \left(s - \frac{n}{2^m}\right) \rho^r \left(2^m t - n\right) \frac{1 - r}{1 + r}.$$  

It will be shown to constitute a positive delta sequence and therefore (3.1) will converge to $f(t)$ uniformly and will avoid Gibbs’ phenomenon (excessive oscillations). We shall use the coefficients in (3.1) together with the decomposition algorithm to construct the coefficients at the coarser scales.

The properties of a positive delta sequence needed for convergence are

(i) $G_m^r(t, s) \geq 0,$

(ii) $\int_{-\infty}^{\infty} G_m^r(t, s) \, ds = 1,$

(iii) $\forall \gamma > 0, \int_{|t-s| \geq \gamma} G_m^r(t, s) \, ds \rightarrow 0$ as $m \rightarrow \infty.$

The first inequality of (3.2) is obvious (but must be taken in the sense of distributions). The second inequality follows immediately from the fact that

$$\sum_n \rho^r(t - n) \frac{1 - r}{1 + r} \equiv 1.$$  

In the case of (iii), we use the fact

$$\int_{|t-s| \geq \gamma} G_m^r(t, s) \, ds = \sum_{|n-2^m t| \geq 2^m \gamma} \rho^r \left(2^m t - n\right) \frac{1 - r}{1 + r}$$

$$= \sum_{|n-x| \geq 2^m \gamma} \rho^r(x - n) \frac{1 - r}{1 + r}.$$  

But $\rho^r(x) = O(r^{-|x|})$ (Walter and Shen [7]) and hence this series converges to 0 uniformly as $m \rightarrow \infty$. Thus, the three inequalities of (3.2) are satisfied and give us the following result.

**Proposition 3.1.** Let $f$ be a function such that both $f$ and $\hat{f} \in L^1(\mathbb{R})$, and let $f_m^r \in V_m$ be given by (3.1), then

(a) $f_m^r \rightarrow f$ uniformly as $m \rightarrow \infty$;

(b) if $M_1 \leq f(t) \leq M_2$, then $M_1 \leq f_m^r(t) \leq M_2$.

Let $e_m(t) = f(t) - f_m^r(t)$. Then by (3.2) (ii), we have

$$|e_m(t)| = \left|\int_{-\infty}^{\infty} G_m^r(t, s)(f(s) - f(t)) \, ds\right|$$

$$\leq \int_{|t-s| \geq \gamma} G_m^r(t, s)(f(s) - f(t)) \, ds$$

$$+ \int_{|t-s| < \gamma} G_m^r(t, s)(f(s) - f(t)) \, ds$$

$$= l_1 + l_2.$$
We now choose $\gamma$ such that $|f(s) - f(t)| < \epsilon/2$ for $|t - s| < \gamma$, which can be done since $f$ is uniformly continuous. Then the second integral satisfies

$$I_2 \leq \int_{|t-s|<\gamma} G_m'(t,s) \frac{\epsilon}{2} ds \leq \frac{\epsilon}{2} \int_{-\infty}^{\infty} G_m'(t,s) ds = \frac{\epsilon}{2},$$

and the first integral satisfies

$$I_1 \leq \left| \int_{|t-s|\geq\gamma} G_m'(t,s)(|f(s)| + |f(t)|) ds \right| \leq 2\|f\| \int_{|t-s|<\gamma} G_m'(t,s) ds,$$

since $\hat{f} \in L^1(\mathbb{R})$ implies that $f$ is bounded. By (iii), we see that this converges to 0 and hence may be chosen less than $\epsilon/2$ for $m \geq m_0$. Therefore, we have

$$|e_m(t)| \leq \epsilon, \quad \text{for } m \geq m_0,$$

which gives conclusion (i). Conclusion (ii) is even easier,

$$f_m^*(t) = \int_{-\infty}^{\infty} G_m'(t,s) f(s) ds \leq \int_{-\infty}^{\infty} G_m'(t,s) M_2 ds,$$

and similarly for the other inequality.

We also have local convergence results. They are based on the same inequalities except we have to stay away from the end points of the interval of continuity.

**Proposition 3.2.** Let $f \in L^1(\mathbb{R})$ be piecewise continuous and be continuous in $[a, b]$; and let $[\alpha, \beta] \subseteq (a, b)$; then

(i) $f_m^*(t) \rightarrow f(t)$ uniformly in $[\alpha, \beta]$;

(ii) if $M_1 \leq f(t) \leq M_2$ in $(a, b)$, then $\forall \epsilon > 0$, $\exists m_0$ such that

$$M_1 - \epsilon \leq f_m^*(t) \leq M_2 + \epsilon, \quad \text{for } t \in [\alpha, \beta], \quad m > m_0.$$

The proofs of these assertions are similar, though a little harder, and will be omitted.

In order to obtain the rate of convergence, we will compare $f_m^*(t)$ to the projection $f_m(t)$ of $f(t)$ onto $V_m$. We know its rate of convergence ([3, 4, 5 p. 128]) provided that the original scaling function $\phi(t)$ has vanishing moments. This is given by

**Proposition 3.3.** Let $\int_{-\infty}^{\infty} t^k \phi(t) dt = 0$, $k = 1, 2, \ldots, \lfloor \lambda \rfloor$, and let $f \in H^\alpha$ for $\alpha > \lambda + \frac{1}{2}$; then

$$|f_m(t) - f(t)| = O(2^{-m\lambda}), \quad \text{uniformly in } \mathbb{R}.$$
We use this result to get our convergence rate for the series (3.1). We have the following

**THEOREM 3.1.** Let \( \phi(t) \) be a continuous orthogonal scaling function with compact support; if

(i) \( f \in H^\alpha \) for \( \alpha > \frac{3}{2} \), then \( \| f_m f - f \|_\infty = O(2^{-m}) \).

(ii) \( f \in H^\alpha \) for \( \alpha > \frac{5}{2} \), and, in addition, \( \int_\infty^{-\infty} t \phi(t) \, dt = 0 \), then \( \| f_m f - f \|_\infty = O(2^{-2m}) \).

**Proof.** Since by Proposition 3.3 these conclusions hold for \( f_m f \) replaced by \( f_m \) (no vanishing moment and one vanishing moment, respectively), we can reach our conclusion by examining the difference

\[ E_m(t) = f_m'(t) - f_m(t). \]

We adopt the notation

\[ \rho(t) = \rho'(t) \frac{1 - r}{1 + r}, \quad \tilde{\rho}(t) = \tilde{\rho}_r(t) \frac{1 + r}{1 - r}, \]

for simplicity. Thus \( E_m \) becomes

\[
E_m(t) = \sum_{n=\infty}^{\infty} \left[ f \left( \frac{n}{2^m} \right) - \int_\infty^{\infty} f(s) \tilde{\rho}(2^m s - n) 2^m \, ds \right] \rho(2^m t - n)
\]

\[
= \sum_{n=\infty}^{\infty} \int_\infty^{\infty} f(s) \left[ \delta(s - \frac{n}{2^m}) - \tilde{\rho}(2^m s - n) 2^m \right] ds \rho(2^m t - n)
\]

\[
= \frac{1}{2\pi} \int_\infty^{\infty} \tilde{f}(\omega) \left[ 1 - \tilde{\rho}(2^{-m} \omega) \right] \sum_{n=\infty}^{\infty} e^{-in\omega - m} \rho(2^m t - n) \, d\omega.
\]

Then we have

\[
|E_m(t)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \tilde{f}(\omega) \right| \left| (1 - \tilde{\rho}(2^{-m} \omega)) \right| d\omega \sum_{n=\infty}^{\infty} \rho(2^m t - n);
\]

but

\[
\sum_{n=\infty}^{\infty} \rho(2^m t - n) = \sum_{j=\infty}^{\infty} \phi(2^m t - j) \sum_{k=\infty}^{\infty} r^{|k|} \frac{1 - r}{1 + r} = 1,
\]

and hence, if \( f \in H^\alpha \), we have

\[
|E_m(t)|^2 \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \tilde{f}(\omega) \right|^2 (\omega^2 + 1)^{\alpha} d\omega \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|1 - \tilde{\rho}(2^{-m} \omega)|^2}{(\omega^2 + 1)^{\alpha}} d\omega,
\]

or

\[
|E_m(t)|^2 \leq \left( \frac{1}{2\pi} \| f \|_\alpha \right)^2 \int_{-\infty}^{\infty} \frac{|1 - \tilde{\rho}(2^{-m} \omega)|^2}{(\omega^2 + 1)^{\alpha}} d\omega. \tag{3.3}
\]

Now

\[
\tilde{\rho}(\omega) = \frac{1}{(1 - r)^2} \left[ (1 + r^2) \phi(\omega) - r \phi(\omega) (e^{i\omega} + e^{-i\omega}) \right]
\]
from Theorem 2.2. Hence \( \hat{\rho}(0) = (1 - r)^2 / (1 - r)^2 = 1 \), and
\[
\hat{\rho}'(0) = \frac{(1 + r^2 - 2r)}{(1 - r)^2} \hat{\phi}'(0) - \frac{r}{(1 - r)^2} \hat{\phi}(0)(1 - 1).
\]
Thus \( \hat{\rho}'(0) = 0 \) if \( \hat{\phi}'(0) = 0 \), i.e., \( \hat{\rho}(t) \) has a vanishing first moment if \( \phi(t) \) does. (This does not extend to higher moment, however.) From this, it follows that
\[
1 - \hat{\rho}(\omega) = O(|\omega|), \quad \text{if } \hat{\phi}'(0) \neq 0,
\]
and
\[
1 - \hat{\rho}(\omega) = O(|\omega|^2), \quad \text{if } \hat{\phi}'(0) = 0.
\]
By using this result in (3.3), we find that \( |E_m(t)| \leq C2^{-m} \) in the first case and \( |E_m(t)| \leq C2^{-2m} \) in the second case.

4. DECOMPOSITION ALGORITHM

The mother wavelet and its dual, which we denote by \( \xi_r \) and \( \tilde{\xi}_r \), could have been defined in a number of different ways, since both \( \{ \rho_r(t - n) \} \) and \( \{ \tilde{\rho}_r(t - n) \} \) span the same space \( V_0 \). They need merely to be biorthogonal and span the orthogonal complement of \( V_0 \) in \( V_1 \). However, in order to get the exact reconstruction property, we need a definition similar to (2.6) and (2.7) [1, p. 263], which is the one we shall use.

The decomposition algorithm involves the coefficients of the dilation equation (2.1) for estimating the coefficients of both the summability function series and the corresponding mother wavelet series at coarser scales. We denote by \( \{ a^m_k \} \) the former at scale \( m \), and by \( \{ \beta^m_k \} \) the latter at the same scale. At the finest scale of interest, the coefficients are given by \( a^m_k = f(k2^{-m})2^{-m/2} \), which gives us our approximation (3.1). In terms of the normalized functions of (3.5), this is
\[
f^m_m(t) = \sum_n a^m_n 2^{m/2} \rho(2^m t - n). \tag{4.1}
\]
At the other scales, we use the same formulae as in (4.1), but the coefficients are now obtained by using the dilation equation
\[
a^{m-1}_n = \sum_k a^m_{2n+k} \tilde{a}_k.
\]
These are no longer guaranteed to be nonnegative for \( f \geq 0 \), but often will remain so, as we shall see in the next section. The corresponding wavelet series, which gives the error in going to coarser scales, is given similarly by
\[
g^m_m(t) = \sum_n \beta^m_n 2^{m/2} \tilde{\xi}(2^m t - n),
\]
where the coefficients are obtained from the other dilation equation (2.6). The \( \beta^m_n \) satisfy
\[
\beta^{m-1}_n = \sum_k a^m_{2n+k} \tilde{a}_{1-k}(-1)^k.
\]
We also have an approximation to the original function in terms of the dual (biorthogonal) series
\[ \tilde{f}_m(t) = \sum_n \tilde{\alpha}_n^m 2^{m/2} \tilde{\rho}(2^m t - n). \]
Again at the finest scale, the coefficients are given by \( \tilde{\alpha}_n^m = f(k 2^{-m}) 2^{-m/2} \), but the corresponding approximation is no longer nonnegative even though the \( \tilde{\alpha}_n^m \) are. Now, however, the coefficients at the other (coarser) scales given by
\[ \tilde{\alpha}_{n-1}^m = \sum_k \tilde{\alpha}_n^m a_k \]
are nonnegative as well. This follows from the following lemma.

**Lemma 4.1.** Let \( \{a_k\} \) be the filter coefficients of the dilation equation for \( \rho'(t) \) given by (2.7), where \( \{c_k\} \) is the sequence of filter coefficients for any scaling function of compact support; and let \( f \in L^2(\mathbb{R}) \) be uniformly continuous; then, for some \( 0 < r_0 < 1 \) and \( r_0 < r < 1 \), we have \( a_k \geq 0 \) for all \( k \in \mathbb{Z} \).

**Proof.** By the orthogonality of the translates of the scaling function, it follows that the filter coefficients satisfy [1, p. 132],
\[ \sum_n c_{2n} = \sum_n c_{2n+1} = 1/\sqrt{2}, \]
and hence \( \sum_n c_{k-2n} = 1/\sqrt{2} \). Thus, by the Abel summability theorem, the term \( \sum_n r^{\text{in}} c_{k-2n} \) of (2.7) can be made larger than, say, \( 1/2 \) for \( r \) sufficiently close to 1. We denote this term by \( \theta_k \), and observe that
\[ \sum_n c_{k-2n} - c_{k-2n-1} = 0. \]
It follows that the Abel means of this sum can be made as small as wanted, i.e.,
\[ \left| \sum_n r^{\text{in}} (c_{k-2n} - c_{k-2n-1}) \right| = |\theta_k - \theta_{k-1}| < \varepsilon, \]
for \( r \) sufficiently large. Now we write (2.7) as
\[ a_k = \frac{1 + r^2}{1 - r^2} \theta_k - \frac{r}{1 - r^2} [\theta_{k-1} + \theta_{k+1}], \]
\[ = \left[ \frac{1 + r^2}{1 - r^2} - \frac{2r}{1 - r^2} \right] \theta_k + \frac{r}{1 - r^2} [\theta_{k-1} + \theta_{k+1}]. \]
Then we divide both sides of this equation by \( 1 - r \) to obtain
\[ \frac{a_k}{1 - r} = \frac{1}{1 + r} \theta_k - \frac{r}{(1 + r)(1 - r)^2} [\Delta \theta_k - \Delta \theta_{k+1}] \]
\[ \geq \frac{1}{1 + r} (1/2) - \frac{r}{(1 + r)(1 - r)^2} [2\varepsilon], \]
for $r$ sufficiently close to 1, and say, $r > 1/2$. Then we have

$$\frac{a_k}{1-r} \geq \frac{1}{1+r}(1/2 - 8\varepsilon),$$

and since $\varepsilon$ can be made as small as we want, the conclusion follows. 

Now it is straightforward to show that the coefficients $\tilde{\alpha}_m^m$ of (4.3) are nonnegative for $r$ sufficiently close to 1. Thus we have two choices for our approximation. If we are interested primarily in the continuous function approximation, then we use (3.1), which is nonnegative for all $t \in \mathbb{R}$. But the approximations at coarser scales may be negative for some values of $t$, and the coefficients may also be negative. If we are interested mainly in the coefficients at the various scales, we can use the approximation by the dual (biorthogonal) series (4.2), whose coefficients are nonnegative at each scale. The approximating function may now, however, be negative for some values of $t$.

### 5. NUMERICAL EXAMPLES

We now consider a few simple numerical examples in which we compare our positive hybrid approximation to the standard approximation. The Coifman wavelet system will be used in these numerical examples, and we will start with some elements of this system.

#### 5.1. About the Orthogonal Coifman Wavelet System

We denote the $l$th moment of a function $f$ by

$$M_l(f) := \int_{\mathbb{R}} t^l f(t) \, dt,$$

where $l$ is a nonnegative integer.

**Definition 5.1 (Coifman wavelet system).** An orthonormal wavelet system with compact support is called an orthogonal Coifman wavelet system of degree $L$ if the vanishing moments of $\phi$, $\psi$ are both of degree $L$, that is,

$$M_l(\phi) = 0, \quad \text{for } l = 1, 2, \ldots, L,$$

$$M_l(\psi) = 0, \quad \text{for } l = 0, 1, \ldots, L.$$

**Table 1**

Filter Coefficients of Coifman Wavelet Systems

<table>
<thead>
<tr>
<th>$n$</th>
<th>$c_n/\sqrt{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>-0.051429728471</td>
</tr>
<tr>
<td>-1</td>
<td>0.238929728471</td>
</tr>
<tr>
<td>0</td>
<td>0.602859456942</td>
</tr>
<tr>
<td>1</td>
<td>0.272140543058</td>
</tr>
<tr>
<td>2</td>
<td>-0.051429972847</td>
</tr>
<tr>
<td>3</td>
<td>-0.011070271529</td>
</tr>
</tbody>
</table>
This system was suggested by R. Coifman in 1989 and was dubbed “Coiflets” by I. Daubechies [1, p. 258]. The Coiflets possess certain distinguishing properties. They are much more symmetric than other scaling functions because of these vanishing moments. However, the drawback is that they have larger support. In general, the Coiflets with \( L = 2K \) vanishing moments typically have support width \( 3L - 1 \). This is in contrast with the original Daubechies wavelets, which have support width \( 2L - 1 \) with the same number of vanishing moments. The construction of orthogonal Coiflets and their interpolation properties have been studied by several authors (see, for example, Tian and Wells [4]).

To illustrate the properties of positive hybrid sampling series, we use the Coifman wavelet system of degree 2, which is constructed in Daubechies [1, p. 260]. The support for this system is an interval of length \( N = 5 \). The lowpass filter coefficients are given in Table 1. The graph of the scaling and the wavelet functions are shown in Figs. 1 and 2.

### 5.2. Choices of Positivity Parameter \( r \)

To maximize the rate of decay of the summability function \( \rho^r(t) \), the parameter \( r \) should be chosen as small as possible, while keeping \( \rho^r(t) \geq 0 \). The optimum value satisfying this requirement is dependent on the length of the support interval of the associated Coiflet. Several numerical experiments have been performed to find an appropriate \( r \). We illustrate
one of the resulting summability functions in Fig. 3. The associated dual function is shown in Fig. 4. Compared with the value of $r$ needed for the Daubechies wavelets, it appears that the value of $r$ is much smaller for the Coiflets (see Walter and Shen [7] for details). Consequently, the sampling series converges faster.

5.3. Comparison of the Positive Hybrid and Conventional Series

We test the behavior of the approximation given by our positive hybrid sampling function against the conventional sampling series by means of the following two examples.

**Example 5.1.** Sawtooth wave function (Fig. 5):

$$f(t) = \text{sgn}(t) - 2t, \quad |t| < \frac{1}{2}.$$ 

The function is approximated by the positive hybrid sampling function $f_{m}^{r}(t)$, given in (3.1). The smallest value of $r$ which gives positivity is around 0.2; $r = 0.22$ is used at the scale $m = 8$. The result is illustrated in Fig. 6. This result is compared to the approximation obtained from the conventional sampling approximation $f_{m}^{I}$ at the same
FIG. 5. The sawtooth wave function.

FIG. 6. The approximation of the sawtooth wave by positive hybrid sampling series.

FIG. 7. The approximation of the sawtooth wave function by Coiflet sampling series.
FIG. 8. The square wave function.

FIG. 9. The approximation of the square wave by positive hybrid sampling series.

FIG. 10. The approximation of the square wave function by Coiflet sampling series.
level \( m \), shown in Fig. 7, where \( f^I_m \) is defined by

\[
f^I_m(t) = \frac{1}{2^m} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{2^m}\right) \phi_{m,n}(t).
\]

**Example 5.2.** The square wave function (Fig. 8):

\[
g(t) = \chi_{[-1/2, 1/2]}(t).
\]

Similar approximations are made at the same scale, and the results are shown in Figs. 9 and 10.

**6. CONCLUDING REMARK**

We conclude our discussion with the observation that these examples behave exactly as we would expect from the theory. In both examples, the conventional approximation shows considerable overshoot at the discontinuity. Since we know that these series exhibit Gibbs' phenomenon, this overshoot will never go away with finer scales. In contrast, the positive hybrid series shows no such overshoot and the approximation is always nonnegative. In fact, the outcomes are much better than expected. The reason for this is unclear at present.

**REFERENCES**