Optimal Wavelet Expansion via Sampled-Data Control Theory

Kenji Kashima[1] Yutaka Yamamoto [2] Masaaki Nagahara [3]

Abstract

Wavelet theory provides a new type of function expansion and and has found many applications in signal processing. Discrete wavelet transform of a signal $x(t)$ in $\mathcal{L}^2(\mathbf{R})$ is usually computed by the so-called pyramid algorithm. It however requires a proper initialization, i.e., expansion coefficients with respect to the basis of one of the desirable approximation subspaces. An interesting question is how we can obtain such coefficients when only sampled values of $x(t)$ are available. This paper provides a design method for a digital filter that optimally gives such coefficients assuming certain a priori knowledge on the frequency characteristic of the target functions. We then extend the result to the case of non-orthogonal wavelets. Examples show the effectiveness of the proposed method.

1 Introduction

Wavelet theory provides a new type of signal representation: it expands functions in terms of small waves, localized both in time and frequency, in contrast to Fourier analysis where basis functions are sinusoids. It hence applies to many cases where local information is important, for example, image compression. In particular, discrete wavelet transform fits naturally with the digital computer with its basis functions defined by summations but not integrals or derivatives.

Wavelet analysis of a signal begins by approximation by projecting it onto one of the approximation subspaces, which is spanned by shifted scaling function, constituting multiresolution analysis. This determines the finest resolution, and one is led to find expansion coefficients of coaser orders. The pyramid algorithm [8] can then be invoked to obtain such lower scale approximation coefficients. Furthermore, this can be implemented by a multirate filter bank [6, 8, 9]. In this procedure, one has to first find expansion coefficients with respect to the basis formed by the shifted scaling functions. The subsequent approximation accuracy in the pyramid algorithm depends on these coefficients.

In digital signal processing, however, it is more common practice that only sampled values of signals are available. Then it is not possible to accurately compute such expansion coefficients in the first place. One may then directly employ the sampled values as such coefficients, with the idea that the intersampling changes are not large. This is convenient, but obviously not satisfactory from thea theoretical point of view, and hence referred to as "wavelet crime" by Strang and Nguyen [8].

Pu and Francis [7] gave an optimal filter to initialize the expansion from sampled values, and also gave an error analysis. They have shown that the wavelet crime can lead to a very large error. On the other hand, their filter is not causal and infinite-dimensional, and hence not easily realizable.

We instead propose to use a finite-dimensional and causal (with finite-time delay) digital filter; for its design, we assume a certain frequency domain characteristic, and employ the sampled-data control theory as per [5, 11, 12, 13] etc. While the performance may become less optimal compared to that of [7], it is easily realizable and, moreover, as we see later, it can be extended to non-orthogonal wavelets.

The paper is organized as follows: We first introduce discrete wavelet transform and formulate our problem. We reduce this to an equivalent discrete-time $\mathcal{H}^\infty$ problem. Then these results are extended to the case of oversampled signals and non-orthogonal scaling functions. The optimal filter is compared with the conventional wavelet crime to exhibit the difference.

2 Problem Formulation

2.1 Discrete Wavelet Transform

We begin with the basic setup in wavelet theory [6, 8]. For the signal space $\mathcal{L}^2(\mathbf{R})$, a *multiresolution analysis* (MRA) is a series $\{\mathcal{V}_j\}_{j\in\mathbf{Z}}$ of closed subspaces having following properties:

1. $\mathcal{V}_j \subset \mathcal{V}_{j+1}$ for all $j \in \mathbf{Z}$

[1] kashima@acs.i.kyoto-u.ac.jp, Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, JAPAN

[2] yy@i.kyoto-u.ac.jp, the same address as above; author to whom all correspondence should be addressed.

[3] nagahara@acs.i.kyoto-u.ac.jp, the same address as above.

2. $\lim_{j \to -\infty} \mathcal{V}_j = \{0\}$ and $\lim_{j \to \infty} \mathcal{V}_j = \mathcal{L}^2(\mathbf{R})$

3. $f(t) \in \mathcal{V}_j \Leftrightarrow f(2t) \in \mathcal{V}_{j+1}$

4. $f(t) \in \mathcal{V}_0 \Leftrightarrow f(t - kh) \in \mathcal{V}_0$

5. $\mathcal{V}_0$ is spanned by $\{\phi(\cdot - kh)\}_{k \in \mathbf{Z}}$

where $h > 0$ is fixed. The function ϕ in condition 5 is called a *scaling function* and $\mathcal{V}_j$ is spanned by $\{\phi_{j,k}\}_{k \in \mathbf{Z}}$ where

$$\phi_{j,k}(t) := 2^{j/2}\phi(2^j t - kh)$$

for all $j \in \mathbf{Z}$ according to condition 3. Thus for $x(t) \in \mathcal{L}^2(\mathbf{R})$, its orthogonal projection $v_j(t)$ onto $\mathcal{V}_j$ is viewed as the approximation of $x(t)$ at scale j. The scale of the resolution increases as we move up in the nest.

The wavelet subspace $\mathcal{W}_0$ is defined to be the orthogonal complement of $\mathcal{V}_0$ in V_1.

$$\mathcal{V}_1 = \mathcal{V}_0 \oplus \mathcal{W}_0$$

which extends to

$$\mathcal{L}^2(\mathbf{R}) = \mathcal{V}_{j_0} \oplus \mathcal{W}_{j_0} \oplus \mathcal{W}_{j_0+1} \oplus \cdots. \tag{1}$$

Since these wavelets reside in the space spanned by the subsequent space spanned by the shifted and dilated scaling functions $\mathcal{W}_j \subset \mathcal{V}_{j+1}$ follows; i.e.,

$$\psi(t) = \sum_k h(k)\phi(2t - k), k \in \mathbf{Z} \tag{2}$$

holds. For an appropriate series $\{h(k)\}$, the function generated by (2) gives the mother wavelet $\psi(t)$ such that W_j is spanned by $\{\psi_{j,k}\}_{k \in \mathbf{Z}}$ where

$$\psi_{j,k}(t) := 2^{j/2}\psi(2^j t - kh). \tag{3}$$

We have now constructed a set of functions that could span all of $\mathcal{L}^2(\mathbf{R})$. According to (1), any function $x(t) \in \mathcal{L}^2(\mathbf{R})$ could be written as a series expansion in terms of the scaling function and wavelets:

$$x(t) = \sum_{k=-\infty}^{\infty} c_{j_0}(k)\phi_{j_0,k}(t) + \sum_{k=-\infty}^{\infty} \sum_{j=j_0}^{\infty} d_j(k)\psi_{j,k}(t). \tag{4}$$

This is called the wavelet expansion and *Discrete Wavelet Transform* (DWT) is the map from $x(t) \in \mathcal{L}^2(\mathbf{R})$ to $\{(c_{j_0}(k),\ d_{j_0}(k),\ d_{j_0+1}(k), \ldots)\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$.

2.2 Computation of DWT

Firstly it is known that we can apply the filter bank shown in Figure 1 to compute $\{d_{j-1}(k)\}_{k \in \mathbf{Z}}$ and $\{c_{j-1}(k)\}_{k \in \mathbf{Z}}$ along with $\{c_j(k)\}_{k \in \mathbf{Z}}$ [8]. Here HPF and LPF denote, respectively, the ideal highpass and lowpass filters with cutoff frequency π/h[rad/sec], and $\downarrow 2$ is a downsampler defined by

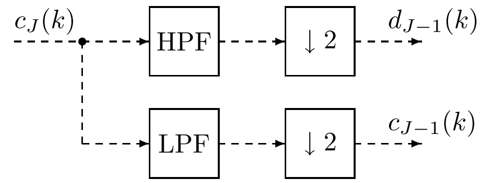

Figure 1: First stage of hte pyramid algorithm.

$$\downarrow 2 : \ell^2(\mathbf{Z}) \longrightarrow \ell^2(\mathbf{Z}) : \{x(k)\}_{k \in Z} \mapsto \{x(2k)\}_{k \in Z}.$$

In practice, computation of the DWT of $x(t) \in \mathcal{L}^2(\mathbf{R})$ goes as follows:

1. Select some level J where $v_J(t) \in \mathcal{V}_J$ represents $x(t)$ to a desired degree of resolution.

2. If the approximation coefficients at this level $\{c_J(k)\}_{k \in \mathbf{Z}}$ are available, we can obtain $\{d_{J-1}(k)\}_{k \in \mathbf{Z}}$ and $\{c_{J-1}(k)\}_{k \in \mathbf{Z}}$ by filter bank above.

3. Since we can similarly obtain $\{d_{j-1}(k)\}_{k \in \mathbf{Z}}$ and $\{c_{j-1}(k)\}_{k \in \mathbf{Z}}$ along with $\{c_j(k)\}_{k \in \mathbf{Z}}$, repeat a number of times down to the coarsest desired scale j_0.

This algorithm is called the *pyramid algorithm*.

As shown in step 2, the pyramid algorithm should be initialized by $\{c_J(k)\}_{k \in \mathbf{Z}}$. In what follows, we first consider the case of orthogonal wavelets.

Assumption 1 *Scaling function ϕ is orthonormal, i.e.,*

$$\langle \phi(t), \phi(t - kh) \rangle_{\mathcal{L}^2(\mathbf{R})} = \delta_{k,0}$$

for all $k \in \mathbf{Z}$.

where δ_{ij} denotes the Kronecker delta.

This assumption makes $\{\phi_{J,k}\}_{k \in \mathbf{Z}}$ an orthonormal system of $\mathcal{V}_J$ and the projection coefficients are given by the inner product $\langle x, \phi_{J,k} \rangle_{\mathcal{L}^2(\mathbf{R})}$. However even in this case it will be difficult to compute by means of its definition itself. Since often only sampled data of $x(t)$ are available, common practice is to use these sampled values directly in place of the correct coefficients $\{c_J(k)\}_{k \in \mathbf{Z}}$. Because this is of course for convenience and causes errors, it is called the "wavelet crime" in [8]. Our problem here is to compute $\{c_J(k)\}_{k \in \mathbf{Z}}$ directly from these samples via digital filters. This problem is called the optimal initialization problem of the DWT and is formulated and solved by giving an optimal solution in [7].

2.3 Problem Formulation

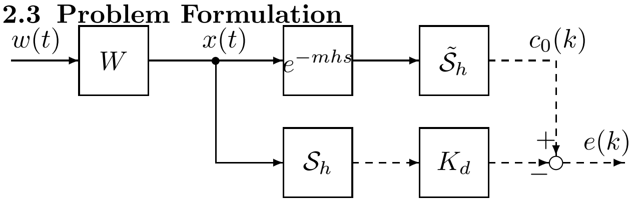

Figure 2: Error system

We formulate the design problem of a digital filter for optimal initialization. Let $J = 0$ without loss of generality. Consider the block diagram Figure 2. We first take $m = 0$. It is appropriate to consider a *class* of functions $x(t)$ for such expansion, and as in $\mathcal{H}^\infty$ control theory, we attempt to minimize the $\mathcal{L}^2$-induced norm from w to the error e. W is an LTI weighting low-pass filter that makes this problem well-posed. This governs the frequency domain characteristic of the target functions. For a given scaling function ϕ, $\tilde{\mathcal{S}}_h$ is the operator defined by

$$\tilde{\mathcal{S}}_h : \mathcal{L}^2(\mathbf{R}) \longrightarrow \ell^2(\mathbf{Z}) : x(t) \mapsto \{c_0(k)\}_{k \in Z}.$$

$\mathcal{S}_h$ denotes the sampler with sampling period h. K_d is the digital filter to be designed. The upper path gives the idealized wavelet expansion coefficients for x whereas the lower path gives an approximant resulting from a sampler and a digital filter. The objective here is to minimize the $\mathcal{L}^2$-induced norm of the operator

$$\mathcal{T}_{ew} : \mathcal{L}^2(\mathbf{R}) \longrightarrow \ell^2(\mathbf{Z}) : x(t) \mapsto \{e(k)\}_{k \in Z}. \quad (5)$$

In other words, the worst case norm of this block diagram against signals in $\mathcal{L}^2(\mathbf{R})$.

Problem 1 *For $m = 0$, given stable $W(s)$ and orthonormal scaling function ϕ, find a digital filter $K_d(z)$ which minimizes $\|\mathcal{T}_{ew}\|$ in (5).*

Pu and Francis [7] have shown the following theorem for this problem.

Theorem 1 *Problem 1 has the (not necessarily unique) solution*

$$\hat{k_d}(e^{j\omega h}) = \frac{\sum_k |\hat{w}(j\omega h + j2\pi k)|^2 \hat{\phi}(-j\omega h - j2\pi k)}{\sum_k |\hat{w}(j\omega h + j2\pi k)|^2}$$

and the corresponding worst case error norm is

$$\left[\sup_\omega \sum_k \left| \left[\hat{k_d}(e^{j\omega h}) - \hat{\phi}(-j\omega h - j2\pi k)\right] \hat{w}(j\omega h + j2\pi k) \right|^2 \right]^{1/2}$$

Here $\hat{k}_d(e^{j\omega})$, $\hat{w}(j\omega)$ and $\hat{\phi}(j\omega)$ denote discrete-time or continuous-time Fourier transform of K_d, W and ϕ respectively.

Unfortunately, this filter in Theorem 1 is neither causal nor finite-dimensional, hence difficult to implement. We thus reformulate the problem as follows: First note that

$$\begin{aligned} (\tilde{\mathcal{S}}_h x)(k) &= \langle x, \phi_{0,k} \rangle_{\mathcal{L}^2(\mathbf{R})} \\ &= \int_{-\infty}^{\infty} x(t)\phi(t - kh)dt \\ &= \int_{-\infty}^{\infty} x(t)\check{\phi}(kh - t)dt \\ &= \mathcal{S}_h(\mathbf{F}x)(k). \end{aligned}$$

Here $\mathbf{F}$ denotes a continuous-time linear time-invariant system with impulse response $\check{\phi}(t) := \phi(-t)$. This means that Problem 1 is an optimal discretization problem of non-causal continuous-time system $\mathbf{F}$. In view of this, we make yet another assumption.

Assumption 2 *There exists an integer $l > 0$ such that the scaling function $\phi(t)$ has compact support in $[0, lh]$, i.e.,*

$$\phi(t) = 0, \; t \notin [0, lh]. \quad (6)$$

Many well-known wavelets satisfy this assumption, for example, Haar scaling function ϕ_H and 2nd order B-spline (triangle) spline ϕ_T defined by

$$\phi_H(t) = \begin{cases} 1, & t \in [0, h] \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\phi_T(t) = \begin{cases} t, & t \in [0, h] \\ 2h - t, & t \in [h, 2h] \\ 0, & \text{otehrwise,} \end{cases} . \quad (7)$$

With this assumption, for $m \geq l$ allowing an m step delay, the error system showed in Figure 2 become causal. Our design problem then becomes as follows:

Problem 2 *Given an integer $m > 0$, stable $W(s)$ and scaling function ϕ satisfying Assumptions 1 and 2, find a digital filter $K_d(z)$ that minimizes $\|\mathcal{T}_{ew}\|$ in (5), where (A, B, C) denotes a minimal realization of W.*

3 Reduction to finite dimensional problem

Suppose $l = 1$ for simplicity; the same method applies to the case of $l > 1$.

First we apply the lifting technique [1, 10] to continuous-time signals in Figure 2, and $\tilde{x}[k](\theta)$ denotes lifted signal, i.e., $\tilde{x}[k](\theta) := x(kh+\theta)$. By lifting x and w, we can obtain a lifted representation of system W

$$\begin{array}{rcccl} X(k+1) & = & e^{Ah}X(k) & + & \int_0^h Be^{A(h-\tau)}\tilde{w}[k](\tau)d\tau \\ \tilde{x}[k](\theta) & = & Ce^{A\theta}X(k) & + & \int_0^\theta Ce^{A(\theta-\tau)}B\tilde{w}[k](\tau)d\tau. \end{array}$$

where $A \in \mathbf{R}^{n\times n}$ and $X(k) \in \mathbf{R}^n$. We then have, from (6), that

$$\begin{aligned} (\tilde{\mathcal{S}}_h x)(k) &= \int_{kh}^{(k+1)h} \phi(t-kh)x(t)dt \\ &= \int_0^h \phi(\tau)x(kh+\tau)d\tau \\ &= \int_0^h \phi(\tau)\tilde{x}[k](\tau)d\tau. \end{aligned}$$

The error system in Figure 2 for $m = l-1$ can be rewritten as that in Figure 3 where

$$A_d := e^{Ah} \in \mathbf{R}^{n\times n}$$

$$\mathcal{B}_1 : \mathcal{L}^2[0,h] \longrightarrow \mathbf{R}^n : \tilde{x}(\tau) \mapsto \int_0^h e^{A(h-\tau)}B\tilde{x}(\tau)d\tau$$

$$C_{d1} := \int_0^h Ce^{A\theta}\phi(\theta)d\theta \in \mathbf{R}^{1\times n}$$

$$\begin{array}{rcl} \mathcal{D}_{11} & : & \mathcal{L}^2[0,h] \longrightarrow \mathbf{R} \\ & : & \tilde{w}(\tau) \mapsto \int_0^h \int_0^\theta Ce^{A(\theta-\tau)}B\phi(\theta)\tilde{w}(\tau)d\tau d\theta. \end{array}$$

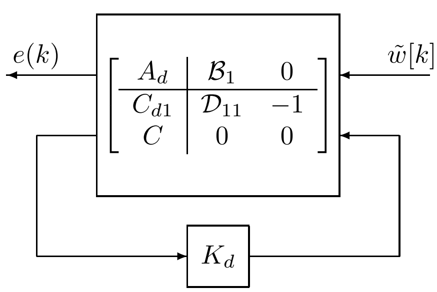

Figure 3: Error system

This can be converted ([2, 3]) to an equivalent discrete-time system by defining B_{d1}, D_{d11} as

$$LL^T := \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{D}_{11} \end{bmatrix} \begin{bmatrix} \mathcal{B}_1^* & \mathcal{D}_{11}^* \end{bmatrix}, \quad \begin{bmatrix} B_{d1} \\ D_{d11} \end{bmatrix} := L.$$

Therefore we can obtain the following theorem.

Theorem 2 *Problem 2 is equivalent to an ℓ^2 induced norm minimization problem of discrete-time system in Figure 4. Here G_{dt} is given by (8) and $A_{dt} \in \mathbf{R}^{(n+m)\times(n+m)}$.*

$$\begin{aligned} G_{dt} &= \left[\begin{array}{c|c} A_{dt} & B_{dt} \\ \hline C_{dt} & D_{dt} \end{array}\right] \\ &:= \left[\begin{array}{ccccc|cc} A_d & 0 & \cdots & \cdots & 0 & B_{d1} & 0 \\ C_{d1} & 0 & \cdots & \cdots & 0 & D_{d11} & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & 1 & 0 & -1 \\ C & 0 & \cdots & \cdots & 0 & 0 & 0 \end{array}\right] \end{aligned} \quad (8)$$

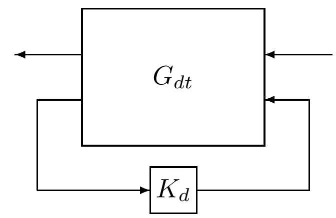

Figure 4: Equivalent discrete time system

For $l > 1$, we can apply the same method by lifting the scaling function. This problem is solvable (sub)optimally and the dimension of the obtained filter does not exceed $n+m$.

4 Extensions

We now give some generalizations, including the case of non-orthogonal wavelets.

4.1 Projection onto higher resolution subspace

We have given a procedure to optimally initialize the wavelet expansion, i.e., optimally projecting onto $\mathcal{V}_0$ which has the same scale of resolution as the sampling rate. It is however possible to project onto a subspace with a *higher scale* resolution than $\mathcal{V}_0$, by making use of a given frequency distribution (i.e., weighting W) of the original signals. This leads us to the design problem depicted in Figure 5.

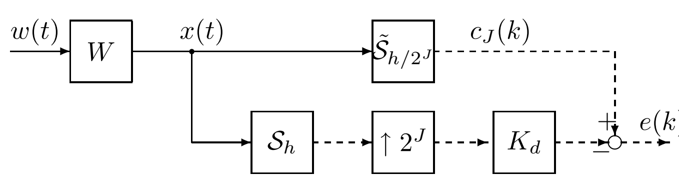

Figure 5: Error system

The differences with Problem 2 are $\tilde{\mathcal{S}}_{h/2^J}$ and $\uparrow 2^J$. First define $\tilde{\mathcal{S}}_{h/2^J}$:

$$\tilde{\mathcal{S}}_{h/2^J} : \mathcal{L}^2(\mathbf{R}) \longrightarrow \ell^2(\mathbf{Z}) : x(t) \mapsto \{c_J(k)\}_{k\in Z},$$

and the output of upper path is $c_J(k)$. In accordance with this, the sampled values are also upsampled by

$\uparrow 2^J$ to match with the upper path in the sampling rate:

$$\uparrow 2^J \;:\; \ell^2(\mathbf{Z}) \longrightarrow \ell^2(\mathbf{Z})$$
$$: \; v(k) \mapsto u(k) := \begin{cases} v(k) & , \quad k = 2^J l, l = 0, 1, \ldots \\ 0 & , \quad \text{otherwise} \end{cases}$$

and processed by $K_d(z)$. This gives an approximation problem in the space $\mathcal{V}_J$ instead of $\mathcal{V}_0$.

A difficulty here is that the overall error system becomes multirate due to the upsampler, but one can reduce this problem to a single-rate problem via discrete-time lifting [4, 3, 11, 12].

4.2 Non-orthogonal scaling function

When Assumption 1 is not satisfied, taking the inner product with the scaling function does not directly give projection coefficients. In such a case we can employ the dual basis $\{\xi_k(t)\}_{k\in\mathbf{Z}}$ such that

$$\langle \phi_{0,p}, \xi_q \rangle_{\mathcal{L}^2(\mathbf{R})} = \delta_{p,q}. \tag{9}$$

According to MRA, we can take

$$\xi_k(t) = \sum_k g(k)\phi(t-k), k \in \mathbf{Z}. \tag{10}$$

for the dual basis. Here desirable coefficients given by $\langle x, \xi_k \rangle_{\mathcal{L}^2(\mathbf{R})}$ are represented as

$$c_0(k) = \sum_l g(l) c_0'(k+l), k \in \mathbf{Z}$$

where $c_0'(k) = \langle x, \phi_{J,q} \rangle_{\mathcal{L}^2(\mathbf{R})}$. Truncating the sum above suitably to approximate the expansion, we can obtain $c_0(k)$ along with $c_0'(k)$ via FIR digital filter

$$K_L(z) = \sum_{l=l_s}^{l_f} g(l) z^{l-l_f}.$$

For $c_0'(k)$ we can apply the method in the previous section.

5 A Design Example

We present a design example for the case of the Haar and 2nd order B-spline wavelets. Put the sampling period $h = 1$, delay step $m = 2$ and weighting function $W(s)$ be

$$W(s) = \frac{1}{(Ts+1)(10Ts+1)}$$

where $T := 22.05/\pi$. Note here that the sampling rate for audio CD (and MD) is 44.1kHz, and the above time constants correspond to the frequecies 1kHz and 10kHz when normalized to $h = 1$. The weighting W above simulates a fairly wide-range orchestral music frequency energy distribution observed in many commercial CD's.

We first compare the present method with that by the "wavelet crime." Note that the Haar wavelet satisfies Assumption 1. Figures 6, 7 show the time responses of the two methods against the source signal $x(t) = 4\sin t$. Figure 6 shows the projection onto $\mathcal{V}_0$ whereas Figure 7 shows the projection onto $\mathcal{V}_1$, per the method given in Section 4.2. The latter exhibits more fidelity. Figure 8 exhibits the corresponding squared error. The present filter shows much advantage over the wavelet crime.

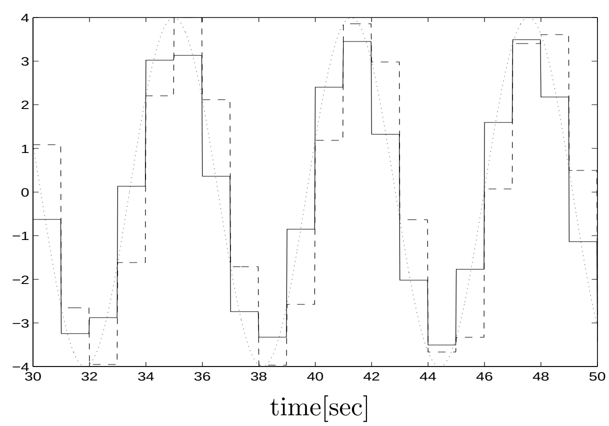

Figure 6: Time response: source signal (dotted), wavelet crime(dash) and proposed (solid)

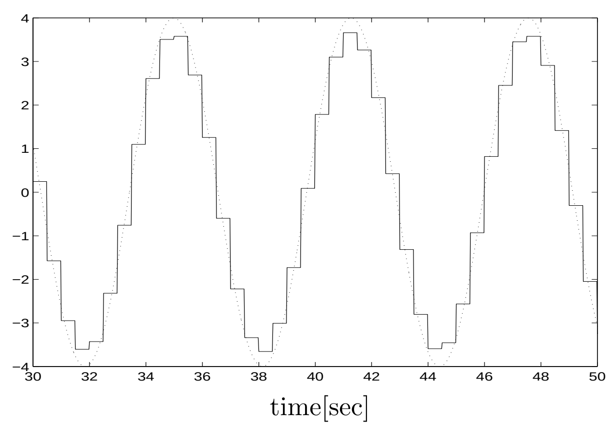

Figure 7: Time response: source signal (dotted) and proposed (projection onto $\mathcal{V}_1$) (solid)

Now consider the 2nd order B-splines. The scaling function is ϕ_T given by (7). Note however that this does not satisfy Assumption 1 and its corresponding subspaces consist of piecewise linear functions. For this scaling function the dual basis is given by (10) where

$$g(l) = \sqrt{3}(\sqrt{3}-2)^{|l|}.$$

Here we take

$$K_L(z) = \sum_{l=-3}^{3} g(k) z^{l-3}$$

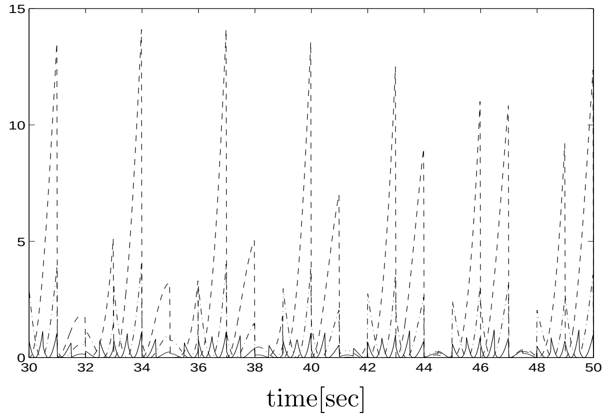

Figure 8: Time responses of squared errors: wavelet crime (dash), projection onto $\mathcal{V}_0$ (dotted) and projection onto $\mathcal{V}_1$ (solid)

and the total delay is taken to be 5 steps. The results for $x(t) = \sin 0.3t + 0.5 \sin t$ are shown in Figure 9, 10.

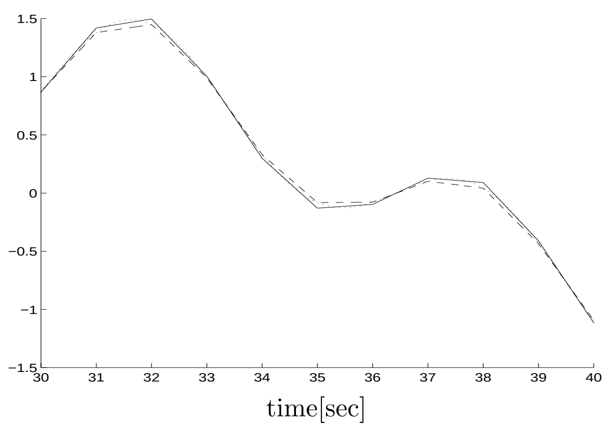

Figure 9: Time response: source signal (solid), wavelet crime(dash) and projection onto $\mathcal{V}_0$(dotted)

Both examples exhibits quite an admirable performance of the proposed method.

6 Concluding Remarks

We have presented a new method, based on sampled-data control theory, for designing a digital filter in the optimal initialization problem of the DWT. The obtained filters are finite-dimensional and show advantages to a naive representation of using sampled values directly. We have also extended the results to the higher order expansion and to the non-orthogonal wavelets.

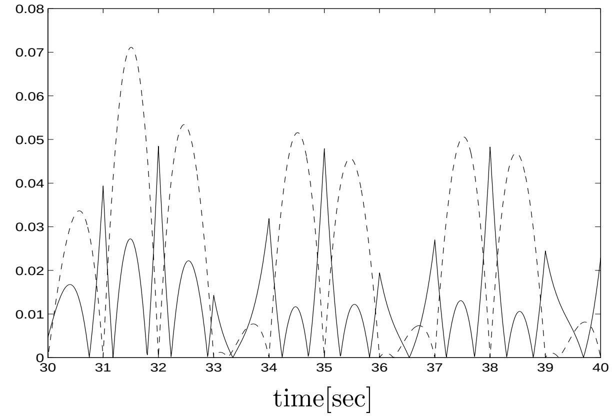

Figure 10: Time responses of squared errors: wavelet crime (dash) and projection onto $\mathcal{V}_0$(solid)

References

[1] B. Bamieh, J. B. Pearson, B. A. Francis and A. Tannenbaum, A lifting technique for linear periodic systems with applications to sampled-data control systems, *Syst. Control Lett.*, **17**: 79–88, 1991.

[2] T. Chen and B. A. Francis, "On the $\mathcal{L}^2$-induced norm of a sampled system," *Systems and Control Letters,* vol. 15, pp. 211–219, 1990.

[3] T. Chen and B. A. Francis, *Optimal Sampled-data Control Systems,* Springer, 1995.

[4] T. Chen and B. A. Francis, "Design of multirate filter banks by H_∞ optimization," *IEEE Trans. Signal Processing,* **SP-43,** pp. 2822–2830, 1995.

[5] P. P. Khargonekar and Y. Yamamoto, "Delayed signal reconstruction using sampled-data control," *Proc. of 35th Conf. on Decision and Control*, pp. 1259–1263, 1996.

[6] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[7] PU Qian and B. A. Francis, "Optimal Initialization of the Discrete Wavelet Transform," submitted for publication May 5, 1999.

[8] G. Strang and T. Nguyen, *Wavelet and Filter Banks*, Wellesley-Cambridge, 1997.

[9] P. P. Vidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.

[10] Y. Yamamoto, "New approach to sampled-data systems: a function space method," *Proc. 29th Conf. on Decision and Control*, pp. 1881–1887, 1990.

[11] Y. Yamamoto, H. Fujioka and P. P. Khargonekar, "Signal reconstruction via sampled-data control with multirate filter banks," *Proc. 36th Conf. on Decision and Control*, pp. 3395–3400, 1997.

[12] Y. Yamamoto and P. P. Khargonekar, "From sampled-data control to signal processing," in *Learning, Control and Hybrid Systems*, Springer Lecture Notes in Control and Information Sciences, vol. 241, pp. 108–126, 1998.

[13] Y. Yamamoto, M. Nagahara and H. Fujioka, "Multirate signal reconstruction and filter design via sampled-data H^∞ control," *MTNS2000*, 2000.

[14] G. Zelniker and F. J. Taylor, *Advanced Digital Signal Processing*, Dekker, 1994.