USC–SIPI REPORT #286

On Theory and Regularization of Scale-Limited Extrapolation

by

Li-Chien Lin and C.-C. Jay Kuo

September 1995

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

On Theory and Regularization of Scale-Limited Extrapolation *

Li-Chien Lin† and C.-C. Jay Kuo‡

September 1, 1995

Abstract

We propose a scale-limited signal model based on wavelet representation and study the reconstructability of scale-limited signals via extrapolation in this research. In analogy with the band-limited case, we define a scale-limited time-concentrated operator, and examine various vector spaces associated with such an operator. It is proved that the scale-limited signal space can be decomposed into the direct sum of two subspaces and only the component in one subspace can be exactly reconstructed, where the reconstructable subspace can be interpreted as a space consisting of scale/time-limited signals. Due to the ill-posedness of scale-limited extrapolation, a regularization process is introduced for noisy data extrapolation.

1 Introduction

The band-limited signal model has been widely used in the past three decades [12], [13], and band-limited extrapolation has been extensively studied and applied in signal reconstruction [4], [11]. Possible applications include spectrum estimation, synthetic aperture radar (SAR) imaging, limited-angle tomography, beamforming and high resolution image restoration. The performance of an extrapolation algorithm is highly dependent on a proper modeling of the underlying signal. There are however signals which are not band-limited such as time-limited signals. Wavelet theory has recently attracted a lot of attention as a useful tool for signal modeling, and the multiresolution wavelet representation leads naturally to a scale-limited signal model.

The scale-limited model includes the band-limited model as a special case, since by choosing the wavelet basis to be the sinc functions, the scale-limited model is reduced to the band-limited one. To illustrate the additional modeling power of the scale-limited model, we may consider the following two examples. First, the cubic cardinal B-spline wavelet basis [2] spans a function

*This work was supported by the National Science Foundation Presidential Faculty Fellow (PFF) Award ASC-9350309.

†The author is with the Department of Electrical Engineering, Feng Chia University, Taichung, Taiwan.

‡The author is with the Signal and Image Processing Institute and the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564.

space whose elements are second-order polynomials between knots and with continuous first-order derivate at knots. Many practical signals can be well approximated with such a function space. Second, time-localized wavelet bases such as the Haar and Daubeches wavelets are more suitable than the conventional Fourier basis in modeling signals with interesting transient information such as those arising from the electrocardiogram and radar applications.

There exist two fundamental questions in signal extrapolation, i.e. the reconstructability of a signal via extrapolation, and the sensitivity of the extrapolation process to noise. With respect to the band-limited case, these two questions have been examined thoroughly. The answer to the first question is positive. That is, a band-limited signal can be exactly reconstructed from its any segment when no noise exists. As to the second question, it is well known that the band-limited extrapolation process is an ill-posed problem. By adding a small amount of noise in observed data, the extrapolated solution may change dramatically. To overcome the ill-posedness of the extrapolation process, it is often to introduce a regularization technique.

Theory on band-limited signal modeling and extraplation has been well developed [13]. It can be dated back to the work [12] of Slepian in early 60's. In order to provide a meaningful explanation for band-limited signals, Slepian [12] constructed a complete set of band-limited functions by using the eigenfunctions of a time-concentrated band-limited operator, known as the prolate spheroidal wave functions (PSWFs). Papoulis [11] and Gerchberg [4] develpoed an iterative algorithm for band-limited signal extrapolation, and proved the convergence of the algorithm by using the PSWFs in the 70's. Many interesting problems can also be conveniently solved based on PSWFs. They include band-limite extrapolation for noisy data [19] and with unevenly sampled observations [1]. The discrete prolate spheroidal sequence (DPSS) has also been studied by researchers [6], [15].

In this research we attempt to answer the above two fundamental questions in the scale-limited context. In analogy with the band-limited case, we define a scale-limited time-concentrated operator, and examine various vector spaces associated with such an operator. It is proved that the scale-limited signal space can be decomposed into the direct sum of two subspaces and only the component in one subspace can be exactly reconstructed. We show that the reconstructable subspace can be well interpreted as a vector space consisting of scale/time-limited signals with a significant amount of energy in the observed interval. In contrast with the band-limited case where every band-limited signal can be exactly reconstructed via extrapolation regardless of the length and position of the observation interval, this result appears to be more consistent with our intuition. It is important to point out that the reconstructability of band-limited signals is fundamentally linked to the fact that every band-limited function is analytic. However, this assumption is too strong to hold in many practical situations. We feel that the scale-limited model provides not only a more general and but also more natural choice than the conventional band-limited one, and should receive more attention. We also prove the ill-posedness of scale-limited extrapolation, and introduce a regularization process in handling noisy obervations.

This paper is organized as follow. We define a scale-limited time-concentrated operator, examine various vector spaces associated with such an operator, and study the reconstructability of scale-limited signals for the continuous-time case in Section 2. The discrete-time case is developed in parallel in Section 3. We describe a scale/time-limited extrapolation algorithm proposed by Xia, Kuo and Zhang [18] and prove its convergence in Section 4. The ill-posedness of scale-limited extrapolation is shown, and a regularization process for noisy data extrapolation is introduced in Section 5. Numerical experiments are provided in Section 6 to illustrate the performance of the regularized extrapolation algorithm. Some concluding remarks are given in Section 7.

2 Reconstructability of Continuous-time Scale-limited Signals

The scale-limited signal model is based on multiresolution analysis and wavelet theory. Consider a sequence of successive approximation space $\mathcal{P}_j$ of $L^2(\mathbf{R})$ satisfying,

$$\cdots \subset \mathcal{P}_{-2} \subset \mathcal{P}_{-1} \subset \mathcal{P}_0 \subset \mathcal{P}_1 \subset \mathcal{P}_2 \cdots, \quad \text{with} \quad \overline{\bigcup_j \mathcal{P}_j} = L^2(\mathbf{R}), \quad \bigcap_j \mathcal{P}_j = \{0\}.$$

Let $\phi(t)$ be the associated scaling function so that $\{\phi_{jk}(t)\}_{k\in\mathbf{Z}}$, where $\phi_{jk}(t) = 2^{j/2}\phi(2^j t - k)$, is an orthonormal basis of the wavelet subspace $\mathcal{P}_j$. The mother wavelet function corresponding to $\phi(t)$ is denoted by $\psi(t)$. Then, $\{\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k),\ j,k \in \mathbf{Z}\}$ forms an orthonormal basis in $L^2(\mathbf{R})$. For any $f(t) \in L^2(\mathbf{R})$, we have

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{jk}(t). \tag{2.1}$$

The projection $f_J(t)$ of $f(t)$ in $\mathcal{P}_J$ can be written as

$$f_J(t) = \sum_{k=-\infty}^{\infty} c_{J,k}\phi_{Jk}(t) = \sum_{j<J} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{jk}(t). \tag{2.2}$$

We call $f_J(t)$ a scale-limited signal, since its wavelet coefficients are zero for $j \geq J$. The wavelet coefficients $b_{j,k}$, $j < J$, can be computed from coefficients $c_{J,k}$ by a fast recursive formula, and vice versa [8].

In analogy with the band-limited time-concentrated operator, we define the scale-limited time-concentrated operator H as an integral operator which maps $f(t) \in L^2(\mathbf{R})$ to $g(t) \in L^2[-T,T]$ via

$$\sum_{k=-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(s)\phi_{Jk}(s)\, ds \right) \phi_{Jk}(t) = g(t), \quad t \in [-T,T]. \tag{2.3}$$

In words, this operator projects the function $f(t)$ into the wavelet subspace $\mathcal{P}_J$ and then truncates the projected function in the time domain. The signal extrapolation problem with a scale-limited model $\mathcal{P}_J$ can be formulated as the solution of $Hf = g$ for the projected $f_J(t)$ of $f(t)$ in $\mathcal{P}_J$ based on the observation $g(t)$. By reconstructability, we mean that $f_J(t)$ can be solved uniquely for any $g(t) \in L^2[-T,T]$. The operator H is clearly linear and bounded. To get more insight into this problem, it is important to examine various vector spaces associated with H and its adjoint H^*. It is easy to derive that HH^* defines an integral operator from $L^2[-T,T]$ to itself, i.e.

$$HH^* g(t) = \int_{-T}^{T} g(s) Q_J(s,t) ds, \qquad t \in [-T,T], \tag{2.4}$$

where

$$Q_J(s,t) \triangleq \sum_{k=-\infty}^{\infty} \phi_{Jk}(s)\phi_{Jk}(t), \qquad (s,t) \in \mathbf{R}^2, \tag{2.5}$$

is the reproducing kernel for the reproducing kernel Hilbert space $\mathcal{P}_J$ [17], [18].

In what follows, we assume that $Q_J(s,t)$ is continuous in $[-T,T]^2$ and finite in $\mathbf{R}^2$. The eigenfunctions of HH^* can be examined by using the adjoint operator theory. For a real scaling function $\phi_{Jm}(t)$, the operator HH^* is self-adjoint and positive semidefinite so that all eigenvalues λ_k of HH^* are real and nonnegative. We can arrange them in a descending order in terms of magnitude,

$$\infty > |\lambda_0| \geq |\lambda_1| \geq \cdots \geq 0,$$

and use $r_0(t)$, $r_1(t)$, $\cdots$ to denote their corresponding eigenfuctions, i.e.

$$HH^* r_k(t) = \lambda_k r_k(t), \qquad t \in [-T,T]. \tag{2.6}$$

The compactness of HH^* can be proved in the following lemma.

Lemma 1 *The operator HH^* is compact.*

Proof: Let us define the kernel

$$Q_{J,M} = \sum_{|m| \leq M} \phi_{Jm}(s)\phi_{Jm}(t).$$

We know from wavelet theory that the scaling function is well concentrated in the time domain so that

$$\|Q_J - Q_{J,M}\| = \sum_{|m|>M} \int_{-T}^{T}\int_{-T}^{T} \phi_{Jm}(s)\phi_{Jm}(t)dt\,ds \to 0, \qquad \text{as } M \to \infty.$$

Let $H_M H_M^*$ denote the integral operator with kernel $Q_{J,M}$. Then, we have

$$\|HH^* - H_M H_M^*\| \leq \|Q_J - Q_{J,M}\| \to 0, \qquad \text{as } M \to \infty.$$

Since the rank of $H_M H_M^*$ is $(2M+1)^2$, it is compact. With the result in [10, pages 384, theorem 5.24.8], we conclude that HH^* is compact. □

By using the above lemma and spectrum theory [10], we claim that the set of functions $\{r_k(t)\}_{k\geq 0}$ is complete and forms an orthogonal basis in $L^2[-T,T]$.

Now, let us focus on the set of eigenfunctions with nonzero eigenvalues, i.e. $r_k(t)$ with $k \in \mathbf{K}$, where

$$\mathbf{K} = \{0, 1, \cdots K : \lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_K > 0 \text{ and } \lambda_{K+1} = \lambda_{K+2} = \cdots = 0\}.$$

By using (2.4) and (2.5), we can extend these eigenfunctions from $[-T,T]$ to $\mathbf{R}$ to define a new set of functions

$$\hat{r}_k(t) = \frac{1}{\lambda_k}\int_{-T}^{T} r_k(s) Q_J(s,t) ds = \frac{1}{\lambda_k}\sum_n \left(\int_{-T}^{T} r_k(s)\phi_{Jn}(s) ds\right)\phi_{Jn}(t), \quad t \in \mathbf{R}. \tag{2.7}$$

Some properties of $\hat{r}_k \in L^2(\mathbf{R})$ were derived in [18] and summarized in Lemma 2.

Lemma 2 *The eigenfuction functions $\hat{r}_k(t)$, $k \in \mathbf{K}$, satisfy the following three properties.*

1. *For $k \in \mathbf{K}$, $\hat{r}_k(t) \in \mathcal{P}_J$, i.e. they are scale-limited.*

2. *The functions $\hat{r}_k(t)$, $k \in \mathbf{K}$, are orthonormal in $(-\infty,\infty)$ and orthogonal in $[-T,T]$, that is,*

$$\int_{-\infty}^{\infty} \hat{r}_k(t)\hat{r}_l(t) dt = \delta_{kl}, \quad \text{and} \quad \int_{-T}^{T} \hat{r}_k(t)\hat{r}_l(t) dt = \lambda_k \delta_{kl}.$$

3. *For any $f(t) \in \mathcal{P}_J$ and $k \in \mathbf{K}$,*

$$\int_{-T}^{T} f(s) r_k(s) ds = \lambda_k \int_{-\infty}^{\infty} f(s)\hat{r}_k(s) ds.$$

4. *The corresponding eigenvalues λ_k with $k \in \mathbf{K}$ are real and $0 < \lambda_k \leq 1$.*

Let us denote the space generated by the orthonormal basis $\hat{r}_k$, $k \in \mathbf{K}$, in $L^2(\mathbf{R})$ by

$$\mathcal{U}_J \triangleq \left\{ \text{closed linear span of } \{\hat{r}_k\}_{k\in\mathbf{K}} \text{ in } L^2(\mathbf{R}) \right\}. \tag{2.8}$$

It is clear from Property 1 of Lemma 2 that $\mathcal{U}_J$ is a linear subspace in $\mathcal{P}_J$. Also, by Properties 2 and 3, the eigenvalue λ_k can be interpreted as the energy contribution of $\hat{r}_k(t)$ in the time interval $[-T,T]$. However, unlike the band-limited case, the eigenfunctions $\hat{r}_k(t)$ are in general not complete in the wavelet subspace $\mathcal{P}_J$ so that $\mathcal{U}_J \neq \mathcal{P}_J$. This will be proved in the following main theorem. For convenience, we use $\mathcal{P}_T$ to denote a space consisting of all functions $f(t) \in L^2(\mathbf{R})$ with $f(t) = 0$ for $t \notin [-T,T]$. The orthogonal complements of $\mathcal{P}_J$ and $\mathcal{P}_T$ are $\mathcal{P}_J^{\perp}$ and $\mathcal{P}_T^{\perp}$. Clearly, $\mathcal{P}_T^{\perp}$ contains all functions $f(t) \in L^2(\mathbf{R})$ with $f(t) = 0$ for $t \in [-T,T]$.

Theorem 1 *For the $\mathcal{U}_J$ defined in (2.8), we have $\mathcal{P}_J = \mathcal{U}_J \oplus \mathcal{U}_J^{\perp}$ where $\mathcal{U}_J^{\perp}$ is the orthogonal complement of $\mathcal{U}_J$ in $\mathcal{P}_J$ and*

$$\mathcal{U}_J^{\perp} = \mathcal{P}_J \cap \mathcal{P}_T^{\perp}.$$

Proof: Given $J > 0$, any $f(t) \in \mathcal{P}_J$ can be decomposed as $f(t) = f_1(t) + f_2(t)$, where $f_1(t)$ is the projection of $f(t)$ onto $\mathcal{U}_J$, and $f_2(t)$ is the projection of $f(t)$ onto $\mathcal{U}_J^{\perp}$. Therefore, we have

$$f_1(t) = \sum_{n\in K} a_n \hat{r}_n(t), \quad \text{where} \quad a_n = \langle f, \hat{r}_n \rangle = \int_{-\infty}^{\infty} f(t)\hat{r}_n(t)dt. \tag{2.9}$$

We want to prove that $f_2(t) \in \mathcal{P}_J$ and $f_2(t) \in \mathcal{P}_T^{\perp}$. First, since $f(t) \in \mathcal{P}_J$, it follows that $f_2(t) \in \mathcal{P}_J$. Next, we will show that $f_2(t) = 0$ for $t \in [-T, T]$. This is equivalent to proving

$$f(t) = f_1(t) = \sum_{n\in K} a_n \hat{r}_n(t), \qquad t \in [-T, T]. \tag{2.10}$$

Recall that $\{r_k(t)\}$ with all $k \geq 0$ forms an orthogonal basis of $L^2[-T, T]$. To prove (2.10), we need to show that

$$\langle f, r_k \rangle_T = \langle \sum_{n\in \mathbf{K}} a_n \hat{r}_n, r_k \rangle_T \qquad \text{for } k \geq 0, \tag{2.11}$$

where the notation $\langle a, b \rangle_T = \int_{-T}^{T} a(t)b(t)dt$ is used.

We first consider the case $k \in \mathbf{K}$. To verify the equality in (2.11), we have $\langle r_n, r_k \rangle_T = \lambda_k \delta_{nk}$ from Property 2 in Lemma 2. By using Property 3 in Lemma 2 and (2.9), it is easy to see that (2.11) holds for $k \in \mathbf{K}$.

For the case $k \notin \mathbf{K}$, we know

$$\langle \sum_{n\in \mathbf{K}} a_n \hat{r}_n, r_k \rangle_T = \langle \sum_{n\in \mathbf{K}} a_n r_n, r_k \rangle_T = 0, \qquad \text{for } k \notin \mathbf{K},$$

since the functions $\{r_k(t)\}$ with $k \geq 0$ form an orthogonal basis in $[-T, T]$. Furthermore, we have

$$\begin{aligned} \langle f(t), r_k(t) \rangle_T &= \int_{-T}^{T} f(t) r_k(t) dt \\ &\overset{(1)}{=} \int_{-T}^{T} \left(\int_{-\infty}^{\infty} f(s) Q_J(s,t) ds \right) r_k(t) dt \\ &= \int_{-\infty}^{\infty} f(s) \left(\int_{-T}^{T} r_k(t) Q_J(s,t) dt \right) ds \overset{(2)}{=} 0. \end{aligned}$$

In the above derivation, equality (1) is based on the fact $f(t) \in \mathcal{P}_J$ and equality (2) is due to that any $r_k(t)$ with $k \notin \mathbf{K}$ is in the null space of the integral operator HH^* defined by (2.4). □

By using the orthogonal projection, it follows that the coefficients a_n for $n \in \mathbf{K}$ in (2.9) can be written as

$$a_n = \langle f, \hat{r}_n \rangle = \frac{1}{\lambda_n} \langle f, r_n \rangle_T = \frac{1}{\lambda_n} \int_{-T}^{T} f(t) r_n(t) dt.$$

The coefficients a_n, $n \in \mathbf{K}$, can be completely determined with the knowledge of the segment of $f(t)$ over the interval $[-T,T]$. It means that $f_1(t)$ can be uniquely determined from $f(t)$ on $[-T,T]$. In the band-limited case, since $f_2(t) = 0$ for $t \in [-T,T]$ as proved above, we have $f_2(t) = 0$ for $t \in R$ by using the analytic property of the function $f_2(t)$. This is however not true for general wavelet bases. As discussed above, a scale-limited signal $f(t) \in \mathcal{P}_J$ can be written as $f(t) = f_1(t) + f_2(t)$, where $f_1(t)$ and $f_2(t)$ are the projections of $f(t)$ onto $\mathcal{U}_J$ and $\mathcal{U}_J^{\perp}$, respectively. The component $f_1(t)$ can be uniquely determined from the values of $f(t)$ in $[-T,T]$ while the component $f_2(t)$ cannot since $f(t)$ contains no information of $f_2(t)$ in $[-T,T]$.

In the context of signal extrapolation, the value of a signal $f(t)$ is observed in $[-T,T]$. We first assume that the function $f(t)$ has a certain finest resolution (or scale) J. It is clear to see that, for a certain scale $J > 0$, we can determined only a finite portion of $f(t)$. Let us assume that the observation of $f(t)$ in the interval $[-T,T]$ can uniquely determine the value of $f(t)$ up to $[-\Pi,\Pi]$ in the time domain. Given T, the value of Π depends on the regularity of the signal and the scale parameters J. The mathematical relationship between these parameters is still open. However, for wavelets with compact support, we can estimate Π based on Theorem 1. Let $f_J(t_1)$ be the projection of $f(t_1)$ onto $\mathcal{P}_J$, which is the convolution of $f(t_1)$ with $\phi(2^J t - t_1)$. It is obvious that if the projection $f_J(t_1)$ does not intersect with the observation interval $[-T,T]$, it cannot be determined from the observations in $[-T,T]$. Hence, the extrapolation interval $[-\Pi,\Pi]$ can be obtained by shifting the scaling function $\phi_J(t)$ along the time axis while maintaining a nonzero intersection of $[-T,T]$ and the shifted $\phi_J(t)$.

3 Reconstructability of Discrete-time Scale-limited Signals

We first introduce some notation and concepts required for the discrete case. To get the discrete sequence, we sample the function $f(t)$ via $x[k] = f(kT_s) = f(\frac{k}{2^{J_s}})$, where $T_s = 2^{-J_s}$ is the sampling period. Since the scaling function $\phi(t)$ behaves like a lowpass filter, we have

$$x[k] = f(\frac{k}{2^{J_s}}) \approx \int_{-\infty}^{\infty} f(t)\phi(t - \frac{k}{2^{J_s}})dt = 2^{-J_s/2}\int_{-\infty}^{\infty} f(t)\phi_{J_s,k}(t)dt = 2^{-J_s/2}c_{J_s,k},$$

Where we assume that J_s is large enough for the above approximation to be valid. Given $c_{J_s,k}$ (or $x[k]$), we can perform the finite scale wavelet transform by computing the wavelet coefficients $c_{J,k}$ and $b_{j,k}$ with $J \le j \le J_s$ recursively. The operator which transforms $c_{J_s,k}$ to $c_{J,k}$ and $b_{j,k}$ with $J \le j \le J_s$denoted by $\mathbf{D}_{J,J_s}$. Since $\mathbf{D}_{J,J_s}$ is linear, it can be represented in a matrix form. When the wavelet basis is real and orthogonal, its inverse operator $\mathbf{D}_{J,J_s}^{-1}$ is the transpose of $\mathbf{D}_{J,J_s}$. We call $\mathbf{D}_{J,J_s}$ and $\mathbf{D}_{J,J_s}^{-1}$ the discrete wavelet tranform (DWT) and the inverse discrete wavelet transform

(IDWT), respectively. A sequence $x[k]$ is called scale-limited if its wavelet coefficients $b_{j,k} = 0$, $J \leq j \leq J_s$, for a certain J. We use $\tilde{\mathcal{P}}_J$ to denote the set of scale-limited discrete signals.

Then, the operator

$$\mathbf{D}_{J,J_s}^{-1}\mathbf{T}_J\mathbf{D}_{J,J_s},$$

be the projection operator which project the sequence onto the subspace $\tilde{\mathcal{P}}_J$, where the operator $\mathbf{T}_J$ truncate the wavelet coefficients $b_{j,k}$, for $J \leq j \leq J_s$. Let P be the dimension of the subspace $\tilde{\mathcal{P}}_J$, which is the number of possible non-zero wavelet transform coefficients of the sequence $x[k] \in \tilde{\mathcal{P}}_J$. Then, without loss of generality, we can express $\mathbf{T}_{J,}$ as

$$\mathbf{T}_J = \mathbf{S}_P^T\mathbf{S}_P,$$

where $\mathbf{S}_P = \{l_{ij}\}$ is a $P \times \infty$ matrix operator

$$s_{ij} = \begin{cases} 1, & \text{if } i = j, \text{ and } 1 \leq i, j \leq P, \\ 0, & \text{otherwise.} \end{cases},$$

Let $\mathbf{W} = \mathbf{D}_{J,J_s}^{-1}\mathbf{S}_P^T$, a $\infty \times P$ matrix Then,

$$\mathbf{D}_{J,J_s}^{-1}\mathbf{T}_J\mathbf{D}_{J,J_s} = \mathbf{D}_{J,J_s}^{-1}\mathbf{S}_P^T\mathbf{S}_P\mathbf{D}_{J,J_s} = \mathbf{W}\mathbf{W}^T \tag{3.1}$$

In analogy with the continuous-time case, the discrete scale-limited time-concentrated operator operator can be defined as

$$\tilde{H}\, y[n] = \sum_{k=-\infty}^{\infty} \mathbf{w}_n\mathbf{w}_k^T y[k], \qquad \text{for } n = -N, \cdots, -1, 0, 1, \cdots, N,$$

and $\tilde{H}\tilde{H}^*$ be the matrix operator mapping from $\mathbf{R}^{2N+1}$ to iterself, that is,

$$\tilde{H}\tilde{H}^*\, y[n] = \sum_{k=-N}^{N} \mathbf{w}_n\mathbf{w}_k^T y[k] \qquad \text{for } n = -N, \cdots, -1, 0, 1, \cdots, N, \tag{3.2}$$

where $\mathbf{w}_n$, a $1 \times P$ vector, is the n-th row vector of matrix $\mathbf{W}$. Here, we are interested in the eigensystem of the operator $\tilde{H}\tilde{H}^*$, i.e.

$$\sum_{k=-N}^{N} \mathbf{w}_n\mathbf{w}_k^T u_i[k] = \lambda_i u_i[n] \qquad \text{for } n = -N, \cdots, -1, 0, 1, \cdots, N. \tag{3.3}$$

Let us arrange the λ_i in order of decreasing magnitude, and vectors $\mathbf{u}_i \in \mathbf{R}^{2N+1}$ are the corresponding orthogonal eigenvectors. With the symmetric property of the operator, we can construct $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{2N+1}\}$ as an orthogonal basis of $\mathbf{R}^{2N+1}$ with the orthogonality,

$$\sum_{k=-N}^{N} u_i[k]u_j[k] = \lambda_i\delta_{i,j} \tag{3.4}$$

Now, let

$$\tilde{\mathrm{K}} = \{k : \lambda_k \neq 0, \quad \text{and} \quad k \in \{0, 1, 2, \cdots 2N+1\}\}.$$

For $i \in \tilde{\mathrm{K}}$, to extend $\mathbf{u}_i[n]$, with $|n| \leq N$, in (3.3) to all integer, we obtain

$$\hat{u}_i[n] = \frac{1}{\lambda_i} \sum_{k=-N}^{N} \mathbf{w}_n \mathbf{w}_k^T u_i[k], \qquad \text{with } n \in \mathbf{Z}. \tag{3.5}$$

It follows that $\hat{u}[n] \in \tilde{\mathcal{P}}_J$.

Let us consider the orthogonal property for $\hat{u}[n] \in \tilde{\mathcal{P}}_J$. By (3.5), we have

$$\sum_{n=-\infty}^{\infty} \hat{u}_i[n]\hat{u}_j[n] = \sum_{n=-\infty}^{\infty} \frac{1}{\lambda_i \lambda_j} \left(\sum_{k=-N}^{N} \mathbf{w}_n \mathbf{w}_k^T u_i[k] \right) \left(\sum_{l=-N}^{N} \mathbf{w}_n \mathbf{w}_l^T u_j[l] \right) \tag{3.6}$$

By using the definition of $\mathbf{W}$, we have

$$\sum_{n=-\infty}^{\infty} \mathbf{w}_n^T \mathbf{w}_n = \mathbf{W}^T \mathbf{W} = \mathbf{S}_P \mathbf{D}_{J,J_s} \mathbf{D}_{J,J_s}^{-1} \mathbf{S}_P^T = \mathbf{S}_P \mathbf{S}_P^T = \mathbf{I}_P, \tag{3.7}$$

where $\mathbf{I}_P$ is an identity matrix of dimension $P \times P$. Therefore, it follows that

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \mathbf{w}_n \mathbf{w}_k^T \mathbf{w}_n \mathbf{w}_l^T &= \mathbf{w}_1 \mathbf{w}_k^T \mathbf{w}_1 \mathbf{w}_l^T + \mathbf{w}_2 \mathbf{w}_k^T \mathbf{w}_2 \mathbf{w}_l^T + \cdots \\ &= \mathbf{w}_k \mathbf{w}_1^T \mathbf{w}_1 \mathbf{w}_l^T + \mathbf{w}_k \mathbf{w}_2^T \mathbf{w}_2 \mathbf{w}_l^T + \cdots \\ &= \mathbf{w}_k \sum_{n=-\infty}^{\infty} \mathbf{w}_n^T \mathbf{w}_n \mathbf{w}_l^T = \mathbf{w}_k \mathbf{w}_l^T. \end{aligned}$$

Then, (3.6) can be rewritten as

$$\begin{aligned} \sum_{k=-N}^{N} \sum_{l=-N}^{N} \frac{1}{\lambda_i \lambda_j} \mathbf{w}_k \mathbf{w}_l^T u_i[k] u_j[l] &= \frac{1}{\lambda_i} \sum_{k=-N}^{N} u_i[k] \left(\frac{1}{\lambda_j} \sum_{l=-N}^{N} \mathbf{w}_k \mathbf{w}_l^T u_j[l] \right) \\ &= \frac{1}{\lambda_i} \sum_{k=-N}^{N} u_i[k] u_j[k] \end{aligned}$$

Therefore, the orthogonality of the eigenvectors $\hat{\mathbf{u}}_i$ can be written as

$$\sum_{k=-\infty}^{\infty} \hat{u}_i[k]\hat{u}_j[k] = \delta_{i,j}. \tag{3.8}$$

As in the continuous-time case, The space generated by the orthonormal basis $\hat{u}_k$, for $k \in \tilde{\mathrm{K}}$, is denoted by

$$\tilde{\mathcal{U}}_J \triangleq \left\{ \text{closed linear span of } \{\hat{u}_k\} \text{ in } l^2(\mathrm{Z}) \right\}.$$

For the discrete extrapolation problem, we are given a segment of scale-limited sequence $x[k] \in \tilde{\mathcal{P}}_J$ with $|k| \le N$, where N is a certain positive integer. For the consistency with continuous-time case, we assume that $N = \frac{T}{2^{J_s}}$. Since the operator $\tilde{H}\tilde{H}^*$ has finite dimension, the result in Section 2 can be directly used in the discrete-time case. Therefore, we can decompose $x[k] = x_1[k] + x_2[k]$, where $x_1[k] = \sum_{i \in \tilde{K}} a_i \hat{u}_i$ can be completely recovered via extrapolation while $x_2[k] \in \tilde{\mathcal{P}}_J$ and $x_2[k] = 0$ for $|k| \le N$. Since the sequence $x[k]$ contains no information about $x_2[k]$, we cannot reconstruct the component $x_2[k]$ via extrapolation.

4 Convergence of Scale-limited Extrapolation Algorithm

We first examine the continuous-time case. Let $f(t) \in \mathcal{P}_J$ be a scale-limited function. We use $\mathbf{P}_J$, $\mathbf{P}_T$, $\mathbf{Q}_J$, $\mathbf{Q}_T$ to denote the projection operators which project functions onto the subspaces $\mathcal{P}_J$, $\mathcal{P}_T$, $\mathcal{P}_J^\perp$ and $\mathcal{P}_T^\perp$, respectively. Given the value of $f(t)$ for $|t| \le T$, the *generalized PG algorithm* proposed by Xia, Kuo and Zhang [18] is to recover $f(t)$ from its segment $g(t) = \mathbf{P}_T f(t)$ via the following iteration:

$$\begin{aligned} f^{(0)}(t) &= g(t), \\ f^{(l+1)}(t) &= \mathbf{Q}_T \mathbf{P}_J f^{(l)}(t) + f^{(0)}(t), \quad l = 0, 1, 2, \cdots. \end{aligned} \tag{4.1}$$

This is also known as the scale-limited extrapolation algorithm. In band-limited extrapolation, Papoulis [11] used the PSWF to prove the convergence of the PG algorithm. Here, we use theorem 1 to investigate the convergence of the generalized PG algorithm. Note that the convergence of the generalized PG algorithm has been proved by Xia et. al. in [18]. However, the convergence proof given there only applies to the wavelet basis with analytic scaling function. Thus, their proof does not include many well-known wavelet bases such as the Daubechies basis and coiflet. In the following, we will give a more generic convergence proof, which turns out to be a direct consequence of Theorem 1.

Corollary 1 *Let $\phi(t)$ be an orthogonal scaling function. If*

$$Q_J(s,t) \triangleq \sum_k \phi_{Jk}(s)\phi_{Jk}(t). \tag{4.2}$$

is continuous and positive definite in the region $[-T,T] \times [-T,T]$. Given a segment $g(t) = \mathbf{P}_T f(t)$ of the scale-limited signal $f(t) \in \mathcal{P}_J$, for $|t| \le T$, then the solution of the generalized PG algorithm will converge to a point $f^\dagger$, which is the orthogonal projection of $f(t)$ onto the subspace $\mathcal{U}_J$.

Proof: We decompose the signal $f(t)$ to be $f(t) = f_1(t) + f_2(t)$ where $f_1(t) \in \mathcal{U}_J$ and $f_2(t) \in \mathcal{U}_J^\perp$. It is clear from Theorem 1 that $\mathbf{P}_T f(t) = \mathbf{P}_T f_1(t) = g(t)$. By using the fact that $f_1(t) \in \mathcal{U}_J$ can be uniquely determined in $\mathcal{U}_J$ from its segment $g(t)$ for $t \in [-T, T]$ and Theorem 3 in [18], we conclude that the generalized PG algorithm converges to $f_1(t)$ which is the orthogonal projection of $f(t)$ onto $\mathcal{U}_J$. □.

The solution is further identified as a minimum norm solution in [7], i.e.

$$\|f_1\| = \min\{\|h\| : h \in \mathcal{P}_J, \mathbf{P}_T h = \mathbf{P}_T f\}. \tag{4.3}$$

Next, we examine the discrete-time case. Define the projection operator

$$\mathbf{T}_{\mathcal{N}} x[n] = \begin{cases} x[n], & n \leq N, \\ 0, & n > N. \end{cases}$$

Given a segment $\mathbf{y} = \mathbf{T}_{\mathcal{N}} x[k]$, $|k| \leq N$, of a scale-limited sequence $x[k] \in \tilde{\mathcal{P}}_J$, the discrete extrapolation problem is to recover $x[k]$ for $|k| \geq N$. Since $x[k] \in \tilde{\mathcal{P}}_J$ is a scale-limited sequence, we have

$$\mathbf{y} = \mathbf{T}_{\mathcal{N}} \mathbf{D}_{J,J_s}^{-1} \mathbf{T}_J \mathbf{D}_{J,J_s} \mathbf{x} = \mathbf{T}_{\mathcal{N}} \mathbf{D}_{J,J_s}^{-1} \mathbf{S}_P^T \mathbf{S}_P \mathbf{D}_{J,J_s} \mathbf{x}. \tag{4.4}$$

The extrapolation is equivalent to solving for $\mathbf{x}$ with a given $\mathbf{y}$. For a scale-limited signal $x[k] \in \tilde{\mathcal{P}}_J$, we can rewrite it as

$$x[k] = x_1[k] + x_2[k] = \sum_{i \in \tilde{\mathrm{K}}} a_i \hat{u}_i[k] + x_2[k],$$

where

$$x_1[k] \in \tilde{\mathcal{P}}_J, \quad \text{and} \quad \mathbf{y} = \mathbf{T}_{\mathcal{N}} x[k] = \mathbf{T}_{\mathcal{N}} x_1[k].$$

Due to the orthogonality property, $x_1[k]$ has to be a minimum norm solution for (4.4). In a similar manner, one can prove that the following *discrete generalized PG algorithm* provides an iterative procedure to solve (4.4):

$$\begin{aligned} x^{(0)}[n] &= \mathbf{T}_{\mathcal{N}} x[n], \\ x^{(l+1)}[n] &= \mathbf{T}_{\mathcal{N}} x[n] + (\mathbf{I} - \mathbf{T}_{\mathcal{N}}) \mathbf{D}_{J,J_s}^{-1} \mathbf{T}_J \mathbf{D}_{J,J_s} x^{(l)}[n], \quad l = 0, 1, 2, \cdots, \end{aligned} \tag{4.5}$$

where I is the identity operator. In fact, the above algorithm leads to a basic decent method for the solution of (4.4) and converges to the minimum norm solution $x_1[k]$. For details, we refer to [7].

5 Noisy Data Extrapolation via Regularization

The scale-limited extrapolation as given in (2.3) is a special case of the Fredholm integral equation of the first kind. It is well known that this problem is essentially an ill-posed one and the generalized PG algorithm fails to compute a meaningful solution when observed data are corrupted by noise.

To determine an approximation solution of an ill-posed problem, which is less sensitive to the noise but still close to the original signal, the idea of utilizing the energy constraints of signal and noise was examined by Xu [19] in the band-limited extrapolation case. These constraints are usually referred to as "regularizers". In this section, we formulate the scale-limited extrapolation with a noise energy constraint, derive a regularized solution by using the eigenfunctions of the scale-limited time-concentrated operator, and propose a practical algorithm.

We adopt the following norm notation

$$\|f(t)\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt, \quad \text{and} \quad \|f(t)\|_T^2 = \int_{-T}^{T} |f(t)|^2 dt.$$

The problem of extrapolating noisy data can be stated as the recovery of a scale-limited signal $f(t) \in \mathcal{P}_J$ based on observed noisy data

$$g(t) = f(t) + n(t), \qquad t \in [-T, T].$$

where $n(t)$ is zero-mean white noise with energy $\|n(t)\|_T^2 \leq \epsilon^2$. We consider a constrained minimum norm solution in $\mathcal{U}_J$, i.e.

$$f^* = \min_{f(t) \in \mathcal{U}_J} \|f(t)\|^2, \quad \text{and} \quad \|f^* - g\|_T^2 \leq \epsilon^2. \tag{5.1}$$

By using the Lagrangian multiplier, this is equivalent to the minimization problem:

$$\min_{f \in \mathcal{U}_J} \left\{ \|f\|^2 + \mu \left(\|f - g\|_T^2 - \epsilon^2 \right) \right\}, \tag{5.2}$$

where μ is the regularization parameter. By using the definition in [16, page 51], $\|f\|^2$ is a stabilizing functional and we call the solution of (5.2) a "regularized solution".

Theorem 2 *Let $r_k(t), k \geq 0$, be the eigenfunctins of the scale-limited time-concentrated operator derived in Section 2 and*

$$g(t) = \sum_{k \geq 0} e_k r_k(t) \qquad \textit{for } t \in [-T, T].$$

Then, the regularized solution to the noisy extrapolation problem is of the form

$$f^*(t) = \sum_{k \in \mathrm{K}} \frac{\lambda_k \mu e_k}{1 + \lambda_k \mu} \hat{r}_k(t), \tag{5.3}$$

where the regularization parameter μ is the solution of

$$\sum_{k \in \mathrm{K}} \frac{\lambda_k e_k^2}{(1 + \lambda_k \mu)^2} = \epsilon^2. \tag{5.4}$$

Proof: For any $f(t) \in \mathcal{U}_J$, we know from theorem 1 that

$$f(t) = \sum_{k \in \mathbf{K}} d_k \hat{r}_k(t). \tag{5.5}$$

Based on the Property 2 in Lemma 2, we have

$$\|f - g\|_T^2 = \sum_{k \in \mathbf{K}} \lambda_k (d_k - e_k)^2. \tag{5.6}$$

By applying the Lagrangian multiplier to (5.2), we are led to the minimization of the functional

$$L(d_k, \mu) = \sum_{k \in \mathbf{K}} d_k^2 + \mu \left[\sum_{k \in \mathbf{K}} \lambda_k (d_k - e_k)^2 - \epsilon^2 \right].$$

The solution can be obtained via

$$\frac{\partial L(d_k, \mu)}{\partial d_k} = 0 \implies d_k = \frac{\lambda_k \mu e_k}{1 + \lambda_k \mu}. \tag{5.7}$$

$$\frac{\partial L(d_k, \mu)}{\partial \mu} = 0 \implies \sum_{k \in \mathbf{K}} \lambda_k (d_k - e_k)^2 = \epsilon^2. \tag{5.8}$$

By substiuting (5.7) into (5.5) and (5.8), we have the results given in (5.3) and (5.4), respectively. □

Note that the choice of regularization parameter μ depends on the noise energy. When noise is small, we need a large μ value so that the regularized solution will be close to the observed signal. On the other hand, for a larger noise level, we need a smoothness constraint on the solution to stabilize the ill-posed problem. This however reduces the accuracy of the regularized solution. Thus, the choice of an appropriate parameter plays an important role in a regularization procedure.

The regalirized form as given in (5.3) in not practical in numerical implementation due to the expensive cost in computing the eigenfunctions and their corresponding eigenvalues. It is therefore important to seek an iterative method to compute the regularized solution numerically. Based on the regularization theory [14], we can adopt the following iterative process:

$$\text{Initialization} \quad f_0(t) = 0. \tag{5.9}$$

For $n = 0, 1, 2, \cdots$,

$$h_{n+1}(t) = (1 - \alpha) f_n + \alpha \mu (g - \mathbf{P}_T f_n), \tag{5.10}$$

$$f_{n+1}(t) = \int_{-\infty}^{\infty} h_{n+1}(s) Q_J(s, t) ds. \tag{5.11}$$

By using the eigenfunctions of the scale-limited time-concentrated operator discussed in Section 2 to analyze the above iterative procedure, we obtain the following theorem.

Theorem 3 *Assume that μ is given and*

$$0 < \alpha < \frac{2}{1+\mu\lambda_k}.$$

The iteration (5.9)-(5.11) converges to the regularized solution in Theorem 2. That is, $\lim_{n\to\infty} f_n(t) = f^*(t)$.

Proof: By using (5.9)-(5.11) with $n = 0$, we have

$$f_1(t) = < g(t), Q_J(s,t) > .$$

Recall that the $r_k(t)$ with $k \geq 0$ form an orthogonal basis of $L^2[-T,T]$. Therefore, we can write $g(t) = \sum_{k\geq 0} e_k r_k(t)$ for $t \in [-T,T]$. Then, since $< r_k(t), Q_J(s,t) >_T = 0$ for $k \notin \mathbf{K}$, we have

$$\begin{aligned} f_1(t) &= \alpha\mu \sum_{k\in\mathbf{K}} e_k < r_k(t), Q_J(s,t) >_T, \\ &= \alpha\mu \sum_{k\in\mathbf{K}} e_k \lambda_k \hat{r}_k(t). \end{aligned}$$

Thus, $f_1(t) \in \mathcal{U}_J$. Furthermore, we apply the fact $< \hat{r}_k(t), Q_J(s,t) >= \hat{r}_k(t)$ and $< \hat{r}_k(t), Q_J(s,t) >_T = \lambda_k \hat{r}_k(t)$ in (5.10) and (5.11). It follows that if $f_n(t) \in \mathcal{U}_J$, $f_{n+1}(t)$ also belongs to $\mathcal{U}_J$. By induction, we have $f_n(t) \in \mathcal{U}_J$ for $n = 1, 2, \cdots$. Hence,

$$f_n(t) = \sum_{k\in\mathbf{K}} d_{k,n} \hat{r}_k(t).$$

Based on the iteration (5.10) and (5.11), we can derive

$$\begin{aligned} d_{k,n+1} &= (1-\alpha) d_{k,n} + \alpha\mu(\lambda_k e_k - \lambda_k d_{k,n}) \\ &= [1 - \alpha(1+\mu\lambda_k)]\, d_{k,n} + \alpha\mu\lambda_k e_k, \end{aligned}$$

for $n = 0, 1, \cdots$ with $d_{k,0} = 0$. Therefore,

$$\begin{aligned} d_{k,n+1} &= \sum_{i=0}^{n} [1 - \alpha(1+\mu\lambda_k)]^i (\alpha\lambda_k\mu) e_k \\ &= \frac{\lambda_k \mu e_k}{1+\lambda_k\mu} \left[1 - (1 - \alpha(1+\mu\lambda_k))^{n+1}\right]. \end{aligned}$$

If $|1-\alpha(1+\mu\lambda_k)| < 1$ or equivalently $0 < \alpha < 2/(1+\mu\lambda_k)$, we have $\lim_{n\to\infty}(1-\alpha(1+\mu\lambda_k))^{n+1} = 0$ which implies $\lim_{n\to\infty} f_n(t) = f^*(t)$. □

Another implementational issue is the computation of the regularization parameter μ required in (5.10). To compute μ with (5.4) is expensive since the eigenvalues of the scale-limited time-concentrated operator are needed. Thus, we seek an approximating regularization parameter close to the one given by (5.4). We can express the regularized solution $f(t)$ in (5.3) explicitly as a function of μ and t and examine

$$\begin{aligned} N(\mu) &= \|f(\mu,t)\|^2 = \sum_{k\in\mathbf{K}} \frac{(\lambda_k \mu e_k)^2}{(1+\lambda_k\mu)^2}, \\ E(\mu) &= \|f(\mu,t)-g(t)\|_T^2 = \sum_{k\in\mathbf{K}} \frac{\lambda_k e_k^2}{(1+\lambda_k\mu)^2}. \end{aligned}$$

Note that $N(\mu)$ and $E(\mu)$ are monotonically increasing and decreasing functions of μ, respectively. The continuous curve consisting of $(N(\mu), E(\mu))$, $\mu \geq 0$, is called the L-curve. The importance of this curve was first discussed by Miller [9]. Recently, Hansen and O'Leary [5] used the L-curve to determine the regularization parameter. It was shown in [5] that the L-curve is concave and there exists a sharp corner on this curve which gives the optimal regularization parameter.

Here, our basic idea to estimate μ is to first determine its upper and lower bounds and to obtain an initial estimation based on these bounds. Then, we exploit the monotonicity and concavity of the L-curve and apply a linear search method. With this approach, a good approximation of μ can be obtained by only a few iterations. To compute the lower bound, we have

$$M \equiv \frac{\|g(t)\|_T^2}{\epsilon^2} = \frac{\sum_{k\in\mathbf{K}} \lambda_k e_k^2}{\sum_{k\in\mathbf{K}} \frac{\lambda_k e_k^2}{(1+\lambda_k\mu)^2}} \leq (1+2\mu+\mu^2) = (\mu+1)^2,$$

where Property 4 in Lemma 2 is used in the above inquality. For the upper bound, we have

$$\|f(\mu,t)\|^2 = \sum_{k\in\mathbf{K}} \frac{(\lambda_k\mu e_k)^2}{(1+\lambda_k\mu)^2} = \sum_{k\in\mathbf{K}} \frac{\lambda_k e_k^2}{(1+\lambda_k\mu)^2}\lambda_k\mu^2 \geq \lambda_{min}\mu^2\epsilon^2,$$

so that

$$\mu^2 \leq \frac{\|f(\mu,t)\|^2}{\lambda_{min}\epsilon^2}.$$

In analogy with the continuous-time case, we can derive a discrete-time regularizarion algorithm. We simply summarize the main result below. Given noisy observations $y[n] = x[n] + \eta[n]$ with $|n| \leq N$ and bounded noise energy i.e. $\sum_{n=-N}^{N} \eta^2[n] \leq \epsilon^2$, the regularized extrapolation method is equivalent to the minimization of the functional

$$\sum_{n=-\infty}^{\infty} (x[n])^2 \quad \text{under the constraint of} \quad \sum_{n=-N}^{N} (y[n]-x[n])^2 \leq \epsilon^2.$$

The discrete regularized solution is

$$x^*[n] = \sum_{k\in\check{\mathbf{K}}} \frac{\lambda_k\mu e_k}{1+\lambda_k\mu}\hat{u}_k[n].$$

where μ denotes the regularization parameter. The iterative algorithm (5.9)-5.11) can be modified as $\mathbf{x}_0 = 0$ and, for $n = 0, 1, 2, \cdots$,

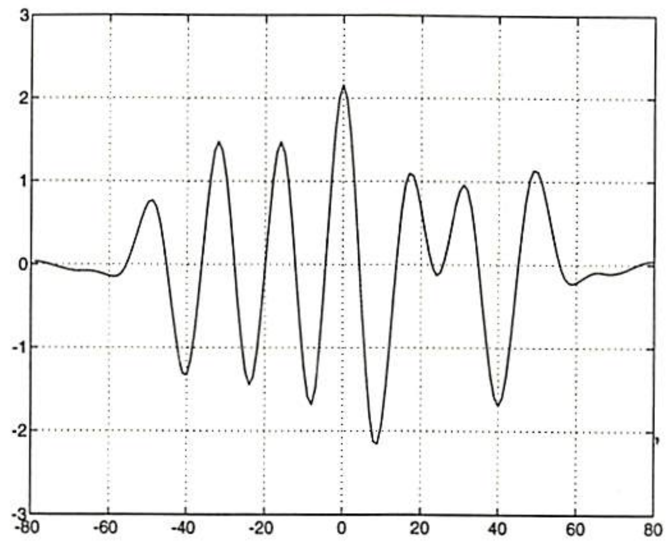
$$\begin{aligned} \mathbf{z}_{n+1} &= (1-\alpha)\mathbf{x}_n + \alpha\mu(\mathbf{y} - \mathbf{P}_T\mathbf{x}_n), \\ \mathbf{x}_{n+1} &= \mathbf{D}_{J,J_s}^{-1}\mathbf{T}_J\mathbf{D}_{J,J_s}\mathbf{z}_{n+1}. \end{aligned}$$

6 Experimental Results

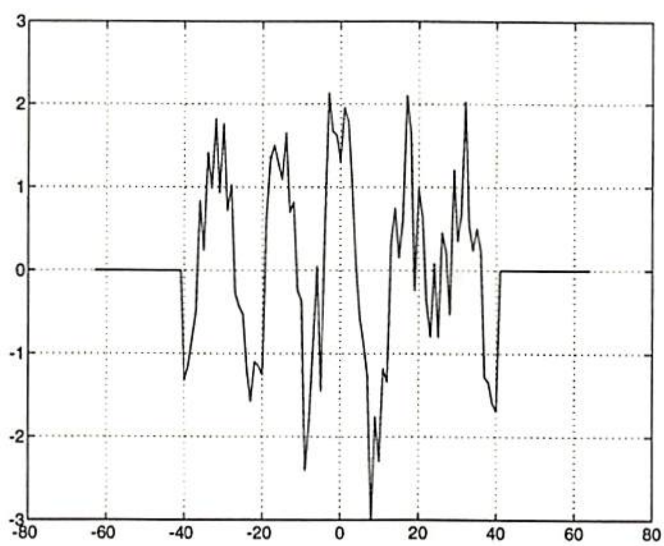
Numerical examples are given in this section to illustrate the performance of the proposed algorithms. We use the orthogonal and compactly supported coiflet of order $N = 10$ [3] as the wavelet basis for signal modeling. The coiflet mother wavelet is nearly symmetric around the y-axis so that the filter bank implementation consists of almost linear-phase filters. The high order of vanishing moments implies the smoothness of the waveform and the compact support property makes the implementation easy. Consider a scale-limited sequence $x[n]$ generated by randomly choosing the wavelet coefficients $c_{J,k}$ with $J = 1$ and $-3 \leq k \leq 4$ while setting other wavelet coefficients to zero for the coiflet basis functions. A synthesized clean signal observed at the scale $J_s = 4$ is plotted in Fig. 1 (a). Then, the signal is corrupted by zero-mean additive white Gaussian noise with $SNR = 8$ and we assume that 81 (i.e. $M = 40$) observed noisy data points are available as given in Fig. 1(b). In this experiment, we avoid the signal modeling problem by assuming that the scale-limited information is partially known a prior, i.e. only $c_{J,k}$ with $J = 1$ and $-3 \leq k \leq 4$ are nonzeros and the wavelet basis is coiflet. However, the exact values of these cofficients $c_{J,k}$ are not known. The extrapolated results by using the regularization approach with the regularization parameter $\mu = 10$, 50 and 10^5 are shown in Fig. 2. It is clear that for $\mu = 10$ we have a oversmoothed result. In contrast, the solution for $\mu = 10^5$ is divergent and the case with $\mu = 50$ gives the best result.

7 Conclusion

Instead of using the traditional Fourier-based technique, the scale-limited signal model based on the wavelet representation is investigated for signal extrapolaton. We proved properties of scale-limited extrapolaton by examining the eigenfunctions of the scale-limited time-concentrated operator. Based on the results, we also proved the convergence of the generalized PG extrapolation algorithm and developed a regularization solution for noisy data extrapolation.



(a)



(b)

Figure 1: Test Problem: (a) the original signal and (b) observed noisy data with $M = 40$ and $SNR = 8dB$.

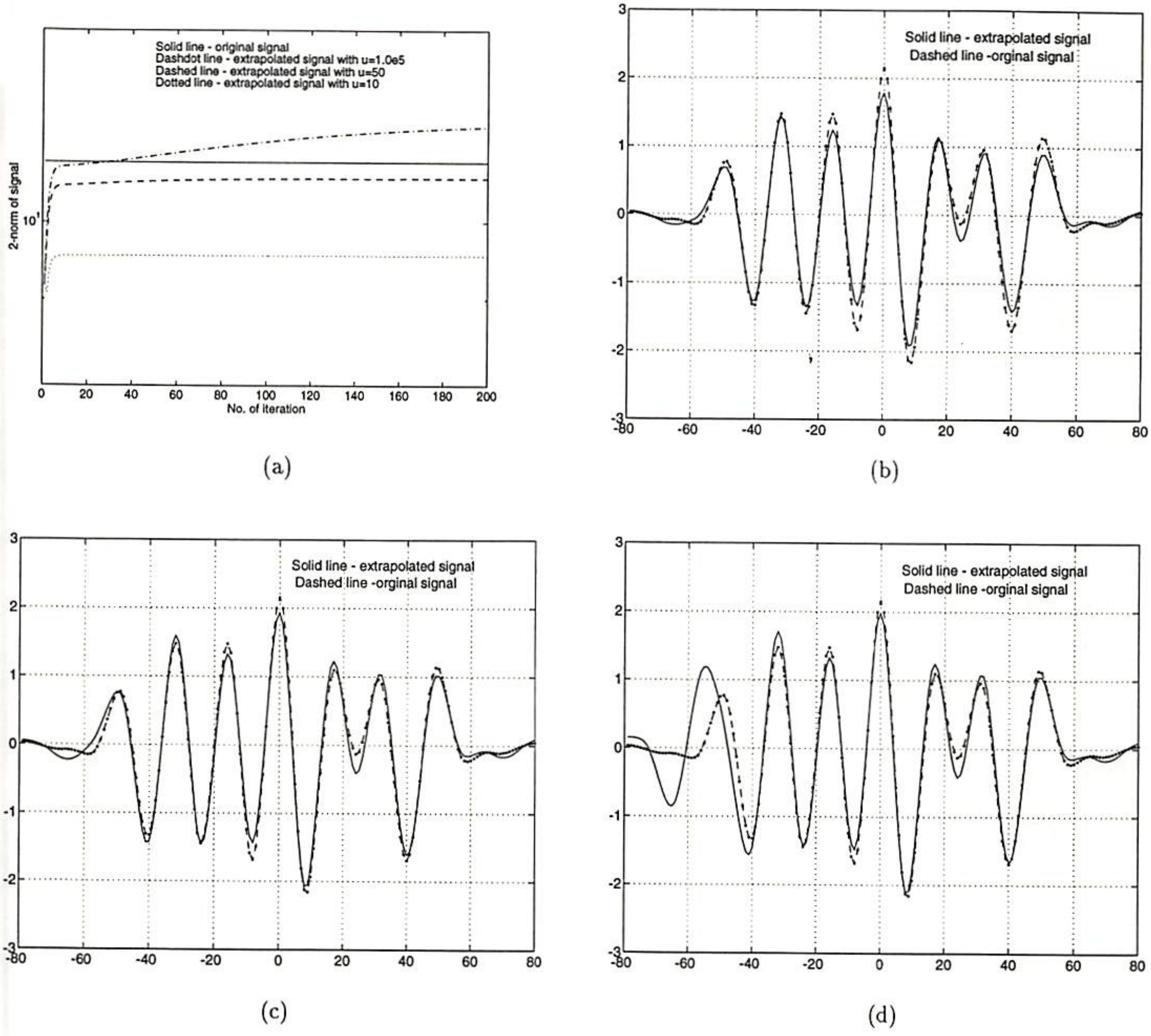


Figure 2: Results of extrapolated signals with regularization: (a) the 2-norm of extrapolated signals and the extrapolated signals with (b) $\mu = 10$, (c) $\mu = 50$, and (d) $\mu = 10^5$ in 200 iterations.

References

[1] T. P. Bronez, "Spectral estimation of irregularly sampled multidimensional process by generalized prolate spheroidal Sequences," *IEEE Trans. on Signal Processing*, Vol. vol. 36, no. 12, pp. 1862–1873, 1988.

[2] C. K. Chui, *An Introduction to Wavelets*, New York: Academic Press, 1992.

[3] I. Daubechies, *Ten Lectures on Wavelets*, Philadephia: SIAM, 1992.

[4] R. W. Gerchberg, "Super-resolution through error energy reduction," *Optica Acta*, Vol. 21, pp. 709–720, 1974.

[5] P. Hansen and D. O'leary, "The use if the L-curve in the regularization of the discrete ill-posed problems," *SIAM J. Sci. Comput.*, Vol. 14, No. 6, pp. 1487–1503, 1993.

[6] A. K. Jain and S. Ranganath, "Extrapolation algorithm for discrete signal with application in spectral estimation," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 29, pp. 830–845, Aug. 1981.

[7] L.-C. Lin and C. C. J. Kuo, "On the convergence of signal extrapolation using wavelets." Submitted to SIAM Matrix Analysis and Applications.

[8] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. on Pattern Anal. and Mach.Intell.*, Vol. 11, pp. 674–693, 1989.

[9] K. Miller, "Least squares methods for ill-posed problems with a prescribed bound," *SIAM J. Math. Anal.*, Vol. 1, pp. 52–74, 1970.

[10] A. H. Naylor and G. R. Sell, *Linear operator theory in engineering anf science*, New York: Springer-Verlag, 1982.

[11] A. Papoulis, "A new algorithm in spectral analysis and band limited extrapolation," *IEEE Trans. on Circuits and Systems*, Vol. 22, pp. 735–742, 1975.

[12] D. S. H. O. Pollak and H. J. Landau, "Prolate spheroidal wave functions I II," *Bell Syst. Tech. J.*, Vol. vol. 40, pp. 43–84, Jan. 1961.

[13] D. Slepian, "Some comments on Fourier analysis, uncertainty and modeling," *SIAM Review*, Vol. vol. 25, no. 3, pp. 379–393, July 1983.

[14] H. Stark, *Image Recovery: Theory and Application*, New York: Academic Press Inc., 1987.

[15] B. J. Sullivan and B. Liu, "On the use of singular value decomposition and decimation in discrete-time Band-limited signal extrapolation," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 32, pp. 1201–1212, Dec. 1984.

[16] A. N. Tikhonov and V. Y. Arsenin, *Solution of ill-posed problems*, Washington, D. C.: V. H. Winston and Sons, 1977.

[17] G. Walter, "A sampling theorem for wavelet subspace," *IEEE Trans. on Information Theory*, Vol. 38, No. 2, pp. 881–884, 1992.

[18] X. G. Xia, C.-C. J. Kuo, and Z. Zhang, "Signal Extrapolation in Wavelet Subspaces." to appear in SIAM J. on Scientific Computing.

[19] W. Xu and C. Chamzas, "On the extrapolation of band-limited functions with energy constraints," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 31, No. 5, pp. 1222–1234, 1983.