SIAM J. SCI. COMPUT.
Vol. 16, No. 1, pp. 50–73, January 1995

005

SIGNAL EXTRAPOLATION IN WAVELET SUBSPACES*

XIANG-GEN XIA†, C.-C. JAY KUO‡, AND ZHEN ZHANG§

Abstract. The Papoulis–Gerchberg (PG) algorithm is well known for band-limited signal extrapolation. The authors consider the generalization of the PG algorithm to signals in the wavelet subspaces in this research. The uniqueness of the extrapolation for continuous-time signals is examined, and sufficient conditions on signals and wavelet bases for the generalized PG (GPG) algorithm to converge are given. A discrete GPG algorithm is proposed for discrete-time signal extrapolation, and its convergence is investigated. Numerical examples are given to illustrate the performance of the discrete GPG algorithm.

Key words. extrapolation, Papoulis–Gerchberg algorithm, wavelets

AMS subject classifications. 41, 65D, 65F, 94

1. Introduction. Band-limited signal interpolation (or sampling) and extrapolation have many applications in both mathematics and engineering, including radio astronomy, radar target detection, geophysical exploration, medical image processing, and communication theory. The Shannon sampling theorem provides a signal interpolation formula from discrete samples of a band-limited signal if the sampling rate is above the Nyquist rate. In the '70s, Papoulis [14] and Gerchberg [9] developed an algorithm for extrapolating a band-limited signal outside a known interval. There have been many extensions and modifications of these two fundamental signal interpolation and extrapolation schemes [3], [4], [11], [15], [18]–[22], [24], [28]–[32]. However, all of them were derived from the Fourier transform viewpoint. Wavelet theory has been extensively studied for the last several years [5], [6], [8]. It provides various attractive multiresolution bases for signal representation with a good time-frequency localization property. In particular, if the scaling function is chosen to be the sinc function, the corresponding wavelet subspaces are those formed by band-limited signals. By extending the Shannon sampling theorem for band-limited signals, Walter [23] derived a general sampling theorem applicable to signals in wavelet subspaces. In this research, we are interested in generalizing the Papoulis–Gerchberg (PG) algorithm from band-limited signals to signals in the wavelet subspaces.

Let us use the following two simple examples to illustrate the nature of the extrapolation problem. Consider first the Haar wavelet, where the scaling function $\phi_H(t) = \chi_{[0,1)}(t)$, which is 1 when $0 \le t < 1$ and 0 otherwise. The wavelet subspaces

$$V_j = \left\{ f(t) : f(t) \text{ is constant in each interval } \left[\frac{k}{2^j}, \frac{(k+1)}{2^j}\right), k \in \mathbf{Z} \right\}, \quad \text{where } j \in \mathbf{Z},$$

consist of piecewise constant functions on intervals of length 2^{-j}. For a signal $f(t) \in V_j$, even though we know that the values of $f(t)$ in interval $[k_0/2^j, k_1/2^j)$, where $k_0, k_1 \in \mathbf{Z}$ and $k_0 < k_1$, there is no unique way to extend $f(t)$ outside the interval. Second, consider the sinc wavelet as mentioned earlier where $\phi_s = \frac{\sin \pi t}{\pi t}$ and the wavelet subspaces V_j are

*Received by the editors October 8, 1992; accepted for publication (in revised form) October 21, 1993. This work was supported by the National Science Foundation grant NCR-9205256, the National Science Foundation Young Investigator Award ASC-9258396, and the Presidential Faculty Fellow (PFF) Award ASC-9350309.

†Department of Electrical and Computer Engineering, Air Force Institute of Technology, AFIT/ENG, 2950 P Street, Wright-Patterson Air Force Base, Ohio 45433-7765 (xxia@laplace.afit.af.mil).

‡Signal and Image Processing Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564 (cckuo@sipi.usc.edu).

§Communication Science Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2565 (zzhang@chief.usc.edu).

$$V_j = \{f(t) : f(t) \text{ is } 2^j\pi \text{ band-limited}\}.$$

For this case, any $f(t)$ in V_j is an entire function [2] so that it can be uniquely determined by its arbitrary piece. One such extrapolation procedure is in fact provided by the PG algorithm. An important focus of this research is to study the convergence of the generalized Papoulis–Gerchberg (GPG) algorithm and the uniqueness of the extrapolated signal. Several sufficient conditions on signals and wavelet bases for convergence and uniqueness will be given in §§3 and 4. To implement extrapolation numerically, we propose a discrete GPG (DGPG) algorithm which extrapolates a scale-time limited sequence.

This paper is organized as follows. In §2, we briefly review the PG algorithm for band-limited signals and basic results of wavelet theory. We consider the extrapolation for continuous-time signals in §3 and give some examples in §4. We then focus on the extrapolation for discrete-time signals and establish a connection between the continuous and discrete cases in §5. Some numerical examples are given in §6.

2. Preliminaries. We briefly review the PG algorithm for band-limited signal extrapolation and orthogonal wavelets in this section. The following notations will be used throughout this paper. The $L^2(\mathbf{R})$ denotes all real square integrable functions (or signals) defined on $\mathbf{R}$. For $D > 0$, the $L^2[-D, D]$ denotes all signals $f(t)$ defined on $[-D, D]$ satisfying

$$\int_{-D}^{D} |f(t)|^2 dt < \infty.$$

Let $\langle\ ,\ \rangle$ and $\|\cdot\|$ denote the inner product and the norm on $L^2(\mathbf{R})$, i.e.,

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(t)g(t)dt, \qquad \text{where } f(t), g(t) \in L^2(\mathbf{R}),$$

and $\|f\|^2 = \langle f, f\rangle$. Similarly, we use $\langle\ ,\ \rangle_D$ and $\|\cdot\|_D$ to denote the inner product and the norm on $L^2[-D, D]$. For $f(t) \in L^2(\mathbf{R})$, we define

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-it\omega}dt$$

to be the Fourier transform of $f(t)$.

2.1. The PG algorithm. A signal $f(t)$ is said to be Ω band-limited if its Fourier spectrum $\hat{f}(\omega) = 0$ for $|\omega| > \Omega$. Let $f(t)$ be Ω band-limited and $f(t)$ be given for $|t| < T$ with $T > 0$. The question is to recover $f(t)$ for $|t| \geq T$. Define two projection operators P_T and P_Ω as follows:

$$P_T f(t) = \begin{cases} f(t), & |t| < T, \\ 0, & |t| \geq T, \end{cases}$$

and

$$P_\Omega \hat{f}(\omega) = \begin{cases} \hat{f}(\omega), & |\omega| < \Omega, \\ 0, & |\omega| \geq \Omega, \end{cases}$$

where P_T and P_Ω act on signals in the time and frequency domains, respectively. Let $\mathcal{F}$ and $\mathcal{F}^{-1}$ be the Fourier transform operator and its inverse, respectively, and let I be the identity operator. Then, the PG algorithm is defined by the following iterative procedure.

PG ALGORITHM.

$$f^{(0)}(t) = P_T f(t). \tag{2.1}$$

For $l = 0, 1, 2, \ldots,$

$$f^{(l+1)}(t) = P_T f(t) + (I - P_T)\mathcal{F}^{-1} P_\Omega \mathcal{F} f^{(l)}(t). \tag{2.2}$$

This can also be written as

$$f^{(l+1)}(t) = P_T f(t) + (I - P_T) P^\Omega f^{(l)}(t), \tag{2.3}$$

where $P^\Omega \triangleq \mathcal{F}^{-1} P_\Omega \mathcal{F}$, which can be viewed as the orthogonal projection from $L^2(\mathbf{R})$ onto the Ω band-limited signal space.

In [14], it was shown that $\| f^{(l)} - f \|$ converges to 0 as l goes to ∞. For the generalization and discretization of the PG algorithm for band-limited signals, we refer to [4], [10], [18]–[22], [24], [28], [30]–[32].

2.2. Orthogonal wavelets. We focus on real orthogonal wavelets in this paper, and refer to [5], [6], [8] for more detailed discussion. Let $\phi(t)$ be a real scaling function such that, for a fixed arbitrary integer j,

$$\{\phi_{jk}(t)\}_{k \in \mathbf{Z}}, \qquad \text{where } \phi_{jk}(t) = 2^{j/2}\phi(2^j t - k)$$

is an orthonormal basis of the wavelet subspace V_j, and $\{V_j\}_{j \in \mathbf{Z}}$ is a multiresolution approximation of $L^2(\mathbf{R})$, i.e., $V_j \subset V_{j+1}$ and $\overline{\bigcup_j V_j} = L^2(\mathbf{R})$. The wavelet function corresponding to $\phi(t)$ is denoted by $\psi(t)$ and $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$. The associated quadrature mirror filters can be expressed as

$$H(\omega) = \sum_k h_k e^{-ik\omega} \quad \text{and} \quad G(\omega) = \sum_k g_k e^{-ik\omega}, \tag{2.4}$$

where $g_k = (-1)^k h_{1-k}$,

$$\hat{\phi}(2\omega) = H(\omega)\hat{\phi}(\omega), \quad \text{and} \quad \hat{\psi}(2\omega) = G(\omega)\hat{\phi}(\omega).$$

Then, we have

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{jk}(t) \tag{2.5}$$

for any $f(t) \in L^2(\mathbf{R})$ and

$$f(t) = \sum_{k=-\infty}^{\infty} c_{J,k}\phi_{Jk}(t) = \sum_{j<J} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{jk}(t) \tag{2.6}$$

for any $f(t) \in V_J$, where $b_{j,k} = \langle f, \psi_{jk} \rangle$ and $c_{J,k} = \langle f, \phi_{Jk} \rangle$. Let P^J denote the orthogonal projection operator from $L^2(\mathbf{R})$ onto V_J, i.e., for any $f(t) \in V_J$,

$$P^J f(t) = \sum_{k=-\infty}^{\infty} c_{J,k}\phi_{Jk}(t). \tag{2.7}$$

The $b_{j,k}$ in (2.5) are called the wavelet series transform (WST) coefficients of $f(t)$, and (2.5) provides the inverse wavelet series transform (IWST) of $b_{j,k}$. On one hand, the WST coefficients $b_{j,k}$ with $j < J$ can be obtained from coefficients $c_{J,k}$ by the recursive formulas:

$$c_{j-1,k} = \sqrt{2}\sum_n h_{n-2k}c_{j,n}, \quad b_{j-1,k} = \sqrt{2}\sum_n g_{n-2k}c_{j,n} \tag{2.8}$$

for $j = J, J-1, J-2, \ldots$. On the other hand, we have the following synthesis formula to compute coefficients $c_{J,k}$ from $c_{J_0,k}$ and $b_{j,k}$ with $J_0 \le j < J$ via

$$c_{j+1,n} = \sqrt{2}\left(\sum_k h_{n-2k}c_{j,k} + \sum_k g_{n-2k}b_{j,k}\right) \tag{2.9}$$

for $j = J_0, J_0+1, \ldots, J-1$. By viewing $c_{J,n}$ as a sequence $x[n]$, we call (2.8) the discrete wavelet transform (DWT) of the sequence $x[n]$ and (2.9) the inverse discrete wavelet transform (IDWT) of coefficients $c_{J_0,k}$ and $b_{j,k}$. By using the orthonormality of the wavelet basis, one can prove that both DWT and IDWT preserve energy.

3. Extrapolation of continuous-time signals.

We examine the extrapolation for continuous-time signals in wavelet subspaces in this section.

3.1. The GPG algorithm.

Let $f(t) \in V_J$ for a fixed integer J. Given the value of $f(t)$ for $|t| < T$ $(T > 0)$, we are concerned with the determination of the value $f(t)$ for $|t| \ge T$. We propose the following GPG algorithm for extrapolation.

GPG ALGORITHM.

$$f^{(0)}(t) = P_T f(t). \tag{3.1}$$

For $l = 0, 1, 2, \ldots,$

$$f^{(l+1)}(t) = P_T f(t) + (I - P_T)P^J f^{(l)}(t), \tag{3.2}$$

where P^J is the projection operator defined in (2.7).

Note that the above iterative procedure reduces to the standard PG algorithm (2.3) if $\phi(t) = \frac{\sin \pi t}{\pi t}$ as shown below. For this case, $f(t) \in V_J$ implies that $f(t)$ is $2^J\pi$ band-limited and $P^J = \mathcal{F}^{-1}P_\Omega\mathcal{F} = P^\Omega$ with $\Omega = \pi 2^J$.

The GPG algorithm has the property that it reduces the error energy during the iteration process. To see this, we know from (3.2) that

$$f^{(l+1)}(t) - f(t) = (I - P_T)(P^J f^{(l)} - f)(t) = (I - P_J)P^J(f^l - f)(t)$$

for $f(t) \in V_J$. Therefore,

$$\|f^{(l+1)} - f\| \le \|I - P^J\|\|P^J\|\|f^{(l)} - f\| \le \|f^{(l)} - f\|.$$

3.2. Convergence and uniqueness results.

We say that a signal $f(t)$ can be *uniquely* determined in a signal set S from its segment $f(t)$ defined on interval $[A, B]$, if any $f(t), g(t) \in S$ with $f(t) = g(t)$ for $t \in [A, B]$, implies $f(t) = g(t)$ for $t \in \mathbf{R}$. In this subsection, we will perform some theoretical study on the convergence of the GPG algorithm, the uniqueness of extrapolated signals, and their relationship.

We first focus on the convergence issue. To do so, results from operator theory [16] are needed. Let us define

$$Q(s,t) \triangleq \sum_{k=-\infty}^{\infty} \phi(s-k)\phi(t-k), \qquad (s,t) \in \mathbf{R}^2, \tag{3.3}$$

and

$$Q_J(s,t) \triangleq 2^J Q(2^J s, 2^J t) = \sum_{k=-\infty}^{\infty} \phi_{Jk}(s)\phi_{Jk}(t). \tag{3.4}$$

The definitions (3.3) and (3.4) also appeared in [23], where $Q_J(s,t)$ was called the reproducing kernel (RK) for the reproducing kernel Hilbert space (RKHS) V_J. In particular, when $\phi(t)$ is equal to the sinc function, it is known that

$$Q_J(s,t) = \frac{\sin 2^J \pi(s-t)}{\pi(s-t)},$$

which is analytic in $\mathbf{R}^2$. When the decay of $\phi(t)$ satisfies $|\phi(t)| \leq O(1+|t|^{0.5+\epsilon})^{-1}$ for some $\epsilon > 0$, $Q(s,t)$ in (3.3) is always finite for all real s, t. In what follows, we assume that $Q_J(s,t)$ is continuous in $[-T,T]^2$ and finite in $\mathbf{R}^2$.

Some basic results are summarized below (see [16]). The following operator defined from $L^2[-T,T]$ to itself,

$$\mathcal{Q}_J\, g(t) = \int_{-T}^{T} g(s) Q_J(s,t)ds, \qquad t \in [-T,T],$$

is completely continuous and symmetric. We use λ_k and $\Phi_k(t)$ to denote the eigenvalues and normalized eigenfunctions of operator $\mathcal{Q}_J$ and arrange $|\lambda_k|$ in a descending order, i.e.,

$$\mathcal{Q}_J \Phi_k(t) = \lambda_k \Phi_k(t), \qquad t \in [-T,T], \tag{3.5}$$

for $k = 0, 1, 2, \ldots$, where $\int_{-T}^{T} |\Phi_k(t)|^2 dt = 1$ and $\infty > |\lambda_0| \geq |\lambda_1| \geq \cdots$. Then, $\{\Phi_k(t)\}_k$ is an orthogonal basis of the range space $\mathcal{Q}_J(L^2[-T,T])$ of operator $\mathcal{Q}_J$ so that we can write

$$\mathcal{Q}_J\, g(t) = \sum_{k=0}^{\infty} \lambda_k \langle g, \Phi_k \rangle_T \Phi_k(t), \quad t \in [-T,T], \quad \forall g(t) \in L^2[-T,T].$$

Moreover, if all eigenvalues λ_k are not zero, $\{\Phi_k(t)\}_k$ is an orthonormal basis of $L^2[-T,T]$ so that we have

$$g(t) = \sum_{k=0}^{\infty} \langle g, \Phi_k \rangle_T \Phi_k(t), \quad t \in [-T,T], \quad \forall g(t) \in L^2[-T,T].$$

Now, let

$$\mathbf{K} = \{k : \lambda_k \neq 0, \quad \text{and} \quad k \in \{0,1,2,\ldots\}\}.$$

For $k \in \mathbf{K}$, we can extend $\Phi_k(t)$ from $[-T,T]$ to $\mathbf{R}$ via (3.5) as

$$\Phi_k(t) = \frac{1}{\lambda_k} \int_{-T}^{T} \Phi_k(s) Q_J(s,t)ds \tag{3.6}$$

$$= \frac{1}{\lambda_k} \sum_n \left(\int_{-T}^{T} \Phi_k(s)\phi_{Jn}(s)ds \right) \phi_{Jn}(t), \tag{3.7}$$

where $t \in \mathbf{R}$.

In the following, the symbol $\Phi_k(t)$ with $k \in \mathbf{K}$ is always used to mean the extended one as given in (3.6). Properties of the extended $\Phi_k(t)$, $k \in \mathbf{K}$, are described in the following lemma.

LEMMA 1. *The extended $\Phi_k(t)$ with $k \in \mathbf{K}$ as defined in* (3.6) *are in V_J and orthogonal in $L^2(\mathbf{R})$, and the associated λ_k are greater than* 0.

Proof. Due to (3.7), $\lambda_k \neq 0$ and the fact that

$$\sum_n \left| \int_{-T}^{T} \Phi_k(s)\phi_{Jn}(s)ds \right|^2 \leq \|\Phi_k\|_T^2 = 1 < \infty,$$

we conclude that $\Phi_k(t) \in V_J$. To check the orthogonality of $\Phi_k(t)$ in $L^2(\mathbf{R})$, we have

$$\begin{aligned}\int_{-\infty}^{\infty} \Phi_{k_1}(t)\Phi_{k_2}(t)dt &= \int_{-\infty}^{\infty} \frac{1}{\lambda_{k_1}\lambda_{k_2}} \int_{-T}^{T} \Phi_{k_1}(s_1)Q_J(s_1,t)ds_1 \int_{-T}^{T} \Phi_{k_2}(s_2)Q_J(s_2,t)ds_2 \\ &= \frac{1}{\lambda_{k_1}\lambda_{k_2}} \int_{-T}^{T}\int_{-T}^{T} \Phi_{k_1}(s_1)\Phi_{k_2}(s_2) \int_{-\infty}^{\infty} Q_J(s_1,t)Q_J(s_2,t)dt ds_1 ds_2\end{aligned}$$

for $k_1, k_2 \in \mathbf{K}$. By using

$$\begin{aligned}\int_{-\infty}^{\infty} Q_J(s_1,t)Q_J(s_2,t)dt &= \int_{-\infty}^{\infty} \sum_{k_1} \phi_{Jk_1}(s_1)\phi_{Jk_1}(t) \sum_{k_2} \phi_{Jk_2}(s_2)\phi_{Jk_2}(t)dt \\ &= \sum_k \phi_{Jk}(s_1)\phi_{Jk}(s_2) = Q_J(s_1,s_2),\end{aligned} \tag{3.8}$$

we can further simplify the above expression to be

$$\begin{aligned}\int_{-\infty}^{\infty} \Phi_{k_1}(t)\Phi_{k_2}(t)dt &= \frac{1}{\lambda_{k_1}\lambda_{k_2}} \int_{-T}^{T}\int_{-T}^{T} \Phi_{k_1}(s_1)\Phi_{k_2}(s_2)Q_J(s_1,s_2)ds_1 ds_2 \\ &\overset{(1)}{=} \frac{1}{\lambda_{k_1}} \int_{-T}^{T} \Phi_{k_1}(s_1)\Phi_{k_2}(s_1)ds_1 \\ &= \begin{cases} \frac{1}{\lambda_{k_1}}, & k_1 = k_2, \\ 0, & k_1 \neq k_2, \end{cases}\end{aligned} \tag{3.9}$$

where the equality labeled with (1) is due to the fact that $\{\Phi_k(t)\}_k$ is an orthonormal basis of $L^2[-T,T]$. It is also obvious that $\lambda_k > 0$. ☐

Consider the space

$$U_J \triangleq \left\{\text{closed linear span of } \{\Phi_k(t)\} \text{ in } L^2(\mathbf{R})\right\} = P^J P_T(L^2(\mathbf{R})), \tag{3.10}$$

generated by $\Phi_k(t)$ with $k \in \mathbf{K}$. For any $g(t) = \sum_{k\in\mathbf{K}} a_k\Phi_k(t) \in U_J$, we know from (3.9) that

$$\|g\|^2 = \sum_{k\in\mathbf{K}} \frac{|a_k|^2}{\lambda_k} < \infty,$$

so that U_J is a subspace of $L^2(\mathbf{R})$. Furthermore, since $\Phi_k(t) \in V_J$ for $k \in \mathbf{K}$, $U_J \subset V_J$. Now, we are ready to state our first main result on the convergence of the GPG algorithm.

THEOREM 1. *Let* $Q_J(s,t)$ *be defined by* (3.4) *and continuous in* $[-T,T]^2$ *and* U_J *be defined by* (3.10). *If* $f(t) \in U_J$, *then* $\|f^{(l)} - f\| \to 0$ *as* $l \to \infty$, *where* $f^{(l)}(t)$ *is obtained from the* GPG *algorithm described by* (3.1)–(3.2).

Proof. We have the following representation for any $f(t) \in U_J$,

$$f(t) = \sum_{k\in\mathbf{K}} a_k \Phi_k(t), \quad \text{where } \sum_{k\in\mathbf{K}} \frac{|a_k|^2}{\lambda_k} < \infty.$$

Given $k \in \mathbf{K}$ and initialization $f_k(t) = \Phi_k(t)$, $f_k^{(0)}(t) = P_T f_k(t)$, consider the iteration

$$f_k^{(l+1)}(t) = P_T f_k(t) + (I - P_T)P^J f_k^{(l)}(t).$$

We want to prove

$$f_k(t) - f_k^{(l)}(t) = (I - P_T)(1-\lambda_k)^l \Phi_k(t) \quad \text{for } l = 0, 1, 2, \ldots. \tag{3.11}$$

The equality (3.11) is trivial for $l = 0$. By assuming (3.11) is true for l, i.e.,

$$\begin{aligned} f_k^{(l)}(t) &= f_k(t) - (I - P_T)(1-\lambda_k)^l \Phi_k(t) \\ &= \Phi_k(t) - (I - P_T)(1-\lambda_k)^l \Phi_k(t) \\ &= (1 - (1-\lambda_k)^l)\Phi_k(t) + (1-\lambda_k)^l P_T \Phi_k(t), \end{aligned}$$

it can be shown that (3.11) holds for $l + 1$. To see this, by using (2.7) let us examine

$$\begin{aligned} P^J f_k^{(l)}(t) &= \sum_n \langle f_k^{(l)}, \phi_{Jn}\rangle \phi_{Jn}(t) = \int_{-\infty}^{\infty} f_k^{(l)}(s) Q_J(s,t)\,ds \\ &= (1 - (1-\lambda_k)^l) \int_{-\infty}^{\infty} \Phi_k(s) Q_J(s,t)\,ds + (1-\lambda_k)^l \int_{-T}^{T} \Phi_k(s) Q_J(s,t)\,ds \\ &= (1 - (1-\lambda_k)^l)\Phi_k(t) + \lambda_k (1-\lambda_k)^l \Phi_k(t) \\ &= (1 - (1-\lambda_k)^{l+1})\Phi_k(t). \end{aligned}$$

Thus, we have

$$\begin{aligned} f_k(t) - f_k^{(l+1)}(t) &= (I - P_T)[\Phi_k(t) - (1 - (1-\lambda_k)^{l+1})\Phi_k(t)] \\ &= (I - P_T)(1-\lambda_k)^{l+1}\Phi_k(t), \end{aligned}$$

and (3.11) is proved by induction. A direct consequence of (3.11) is

$$f(t) - f^{(l)}(t) = \sum_{k\in\mathbf{K}} a_k [f_k(t) - f_k^{(l+1)}(t)] = (I - P_T)\sum_{k\in\mathbf{K}} (1-\lambda_k)^l a_k \Phi_k(t). \tag{3.12}$$

Therefore,

$$\|f - f^{(l)}\|^2 \le \sum_{k\in\mathbf{K}} (1-\lambda_k)^{2l} \frac{|a_k|^2}{\lambda_k}.$$

To show $\|f^{(l)} - f\| \to 0$ as $l \to \infty$, we only have to prove $0 < \lambda_k \le 1$ for $k \in \mathbf{K}$. The fact that $\lambda_k > 0$ has been proved in Lemma 1. By using (3.9), we have

$$\frac{1}{\lambda_k} = \int_{-\infty}^{\infty} |\Phi_k(t)|^2 dt \geq \int_{-T}^{T} |\Phi_k(t)|^2 dt = 1,$$

so that $\lambda_k \leq 1$, which completes the proof of Theorem 1. □

Based on Theorem 1, we have the following straightforward corollary on the uniqueness of extrapolation.

COROLLARY 1. *Let* $Q_J(s,t)$ *be defined by* (3.4) *and continuous in* $[-T,T]^2$ *and* U_J *be defined by* (3.10). *If* $f(t) \in U_J$, *then* $f(t)$ *is uniquely determined in* U_J *from the values of* $f(t)$ *with* $t \in [-T,T]$.

Although it may not be easy to check the condition $f(t) \in U_J$ practically, Theorem 1 does tell us that there exists a subspace U_J in V_J where the GPG algorithm converges. The observation is not only of theoretical interest but also provides an important step to the derivation of further results.

For a kernel $K(s,t)$ satisfying $K(s,t) = K(t,s)$ and

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i \bar{a}_j K(t_i, t_j) \geq 0, \tag{3.13}$$

where the bar denotes the complex conjugate, for any integer $N > 0$, any N points $t_i \in [-T,T]$ and any N numbers a_i, we say $K(s,t)$ is *symmetric nonnegative definite* in $[-T,T]^2$. If the inequality in (3.13) is strictly great than 0 when there is at least one $a_i \neq 0$, we say that $K(s,t)$ is *symmetric positive definite* in $[-T,T]^2$. With a symmetric kernel $K(s,t)$ in $[-T,T]^2$, we can define an operator from $L^2[-T,T]$ to $L^2[-T,T]$ like $\mathcal{Q}_J$,

$$\mathcal{K}g(t) = \int_{-T}^{T} K(s,t)g(s)ds, \qquad t \in [-T,T].$$

Then, $K(s,t)$ is positive (or nonnegative) definite if and only if all eigenvalues of the operator $\mathcal{K}$ are positive (or nonnegative). As an example, the kernel $Q_J(s,t)$ defined in (3.4) is symmetric nonnegative definite in $[-T,T]^2$ for any $\phi(t)$.

The following lemma is needed for the proof of Theorem 2.

LEMMA 2. *For any* $f(t) \in V_J$ *and* $k \in \mathbf{K}$,

$$\int_{-T}^{T} \Phi_k(s) f(s) ds = \lambda_k \int_{-\infty}^{\infty} \Phi_k(s) f(s) ds. \tag{3.14}$$

Proof. We first prove that (3.14) is true when $f(t) = \phi_{Jn}(t)$ for an arbitrary n.

$$\begin{aligned}\int_{-\infty}^{\infty} \Phi_k(s)\phi_{Jn}(s)ds &= \int_{-\infty}^{\infty} \frac{1}{\lambda_k}\int_{-T}^{T} \Phi_k(s')Q_J(s',s)ds'\phi_{Jn}(s)ds \\ &= \frac{1}{\lambda_k}\int_{-T}^{T} \Phi_k(s')\int_{-\infty}^{\infty} Q_J(s',s)\phi_{Jn}(s)dsds' \\ &= \frac{1}{\lambda_k}\int_{-T}^{T} \Phi_k(s')\phi_{Jn}(s')ds'.\end{aligned}$$

For general $f(t) \in V_J$, since $f(t) = \sum_n a_n \phi_{Jn}(t)$ and $\sum_n |a_n|^2 < \infty$, (3.14) also holds. □

The second convergence result is stated as follows.

THEOREM 2. *If $Q_J(s,t)$ is continuous and positive definite in $[-T,T]^2$ and $f(t)$ is uniquely determined in V_J by $f(t)$, $t \in [-T,T]$, then $\|f^{(l)} - f\| \to 0$ when $l \to \infty$, where $f^{(l)}(t)$ is obtained from the* GPG *algorithm* (3.1)–(3.2).

Proof. When $Q_J(s,t)$ is positive definite in $[-T,T]^2$, we have $\lambda_k > 0$ for $k = 0, 1, 2, \ldots$, i.e., $\mathbf{K} = \{0, 1, 2, 3, \ldots\}$. Therefore, by using the results summarized in the beginning of §3.2, $\{\Phi_k(t)\}_k$ is an orthonormal basis of $L^2[-T,T]$. For $f(t) \in V_J$, when $t \in [-T,T]$,

$$f(t) = \sum_{k=0}^{\infty} \langle f, \Phi_k \rangle_T \Phi_k(t).$$

By Lemma 2, when $t \in [-T,T]$,

$$f(t) = \sum_{k=0}^{\infty} \lambda_k \langle f, \Phi_k \rangle \Phi_k(t).$$

For $t \in \mathbf{R}$, let

$$\tilde{f}(t) = \sum_{k=0}^{\infty} \lambda_k \langle f, \Phi_k \rangle \Phi_k(t).$$

Then, $f(t) = \tilde{f}(t)$ for $t \in [-T,T]$ and

$$\tilde{f}(t) = \sum_{k=0}^{\infty} \langle f, \sqrt{\lambda_k}\Phi_k \rangle (\sqrt{\lambda_k}\Phi_k(t)).$$

By (3.9), $\{\sqrt{\lambda_k}\Phi_k(t)\}_k$ is orthonormal in $L^2(\mathbf{R})$. Therefore, $\|\tilde{f}\|^2 \le \|f\|^2 < \infty$. This proves $\tilde{f}(t) \in U_J$. Since $U_J \subset V_J$, $\tilde{f}(t)$ is also in V_J. By assumption, we know that $f(t)$ is uniquely determined in V_J by $f(t)$, $t \in [-T,T]$ so that $f(t) = \tilde{f}(t)$ for all real t. Thus, $f(t) \in U_J$ and, by Theorem 1, Theorem 2 is proved. □

Theorem 2 tells us that if $Q_J(s,t)$ is continuous and positive definite in $[-T,T]^2$, the uniqueness of extrapolation in V_J implies the convergence of the GPG algorithm. Instead of checking the uniqueness of extrapolation for various functions $f(t)$ of interest in V_J individually, the following theorem says that it is sufficient to check that for the scaling function $\phi(t)$ only.

THEOREM 3. *If $Q_J(s,t)$ is continuous and positive definite in $[-T,T]^2$ and the scaling function $\phi(t)$ is uniquely determined in V_J by any one of its segments $\phi(t)$, $t \in [-2^JT - k, 2^JT - k]$, $k \in \mathbf{Z}$, then $\|f^{(l)} - f\| \to 0$ as $l \to \infty$ where $f^{(l)}(t)$ is obtained from the* GPG *algorithm* (3.1)–(3.2).

Proof. Since $\phi(t)$ is uniquely determined in V_J by any one of the segments $\phi(t)$, $t \in [-2^JT - k, 2^JT - k]$, $k \in \mathbf{Z}$, $\phi_{Jk}(t)$ is also uniquely determined by $\phi_{Jk}(t)$, $t \in [-T,T]$ for any $k \in \mathbf{Z}$. Similar to the proof of Theorem 2, $\phi_{Jk}(t) \in U_J$ for all $k \in \mathbf{Z}$. This implies that $V_J \subset U_J$. Thus, $V_J = U_J$ by $U_J \subset V_J$. By Theorem 1, Theorem 3 is proved. □

As a direct consequence of Theorem 3, we have

COROLLARY 2. *Under the same conditions as stated in Theorem* 3, *if $f(t) \in V_J$, then $f(t)$ is uniquely determined in V_J by its segment $f(t)$, $t \in [-T,T]$.*

By the definition of positive definite, if $Q_J(s,t)$ is continuous and positive definite in $[-T,T]^2$ then it is also continuous and positive definite in any $[A,B]^2$ with $-T \le A < B \le T$. The observation implies the following corollary.

COROLLARY 3. *Let $Q_J(s,t)$ be continuous and positive definite in $[-T,T]^2$. If $\phi(t)$ is uniquely determined in V_J by any of its segments and $f(t) \in V_J$, then $\|f^{(l)} - f\| \to 0$ as*

$l \to \infty$ where $f^{(l)}(t)$ is obtained from (3.1)–(3.2) *with $[-T, T]$ replaced by an arbitrarily fixed $[A, B]$ with $-T \le A < B \le T$, and also $f(t)$ is uniquely determined in V_J by its segment $f(t)$, $t \in [A, B]$.*

The proof is completely similar to that of Theorem 3 except replacing $[-T, T]$ by $[A, B]$. The details are therefore omitted. It is easy to see that the Haar wavelet does not satisfy Theorem 3. We consider examples of wavelet bases satisfying the conditions in Theorem 3 in the following section.

4. Examples of wavelet bases for continuous-time signal extrapolation. We first consider a general result on scaling functions such that they satisfy the conditions in Theorem 3.

THEOREM 4. *Let $\phi(t)$ be a scaling function. If its Fourier transform satisfies*

$$\hat{\phi}(\omega) = \begin{cases} 1, & |\omega| < a\pi, \\ 0, & |\omega| > b\pi, \end{cases} \tag{4.1}$$

where a and b with $a < b$ are two positive constants and $a + b \le 2$, then it satisfies the conditions in Theorem 3. *That is, $Q_J(s, t)$ is continuous and positive definite in $[-T, T]^2$ and $\phi(t)$ is uniquely determined in V_J by any one of its segments $\phi(t)$, $t \in [-2^J T - k, 2^J T - k]$, $k \in \mathbf{Z}$ for any $T > 0$.*

Proof. We know from (4.1) that $\phi(t)$ is band-limited. Therefore, $\phi(t)$ is uniquely determined in V_J by any one of its segments $\phi(t)$, $t \in [-2^J T - k, 2^J T - k]$, $k \in \mathbf{Z}$ for any $T > 0$. To prove Theorem 4, we only need to prove the positive definiteness of $Q_J(s, t)$ in $[-T, T]^2$ for any $T > 0$. Without loss of generality, we assume $J = 1$. Thus we need to prove that $Q(s, t)$ is positive definite in $[-T, T]^2$ for any $T > 0$.

Let $g(t) \in L^2[-T, T]$ so that

$$\int_{-T}^{T} g(s) Q(s, t) ds = 0 \quad \forall t \in [-T, T].$$

That is,

$$\sum_{k=-\infty}^{\infty} \int_{-T}^{T} g(s)\phi(s-k) ds \phi(t-k) = 0 \quad \forall t \in [-T, T]. \tag{4.2}$$

Since $\phi(t)$ is band-limited, the function

$$\sum_{k=-\infty}^{\infty} \int_{-T}^{T} g(s)\phi(s-k) ds \phi(t-k)$$

is also band-limited with variable t. Equation (4.2) implies that

$$\sum_{k=-\infty}^{\infty} \int_{-T}^{T} g(s)\phi(s-k) ds \phi(t-k) = 0 \quad \forall t \in \mathbf{R}.$$

Since $\{\phi(t-k)\}$ is an orthonormal basis of V_0,

$$\int_{-T}^{T} g(s)\phi(s-k) ds = 0 \quad \forall k \in \mathbf{Z}.$$

Then, we have

$$\sum_{k=-\infty}^{\infty} \int_{-T}^{T} g(s)\phi(s-k) ds e^{ik\omega} = 0 \quad \forall \omega \in \mathbf{R},$$

or

$$\int_{-T}^{T} g(s) \sum_{k=-\infty}^{\infty} \phi(s-k)e^{ik\omega}ds = 0 \quad \forall \omega \in \mathbf{R}. \tag{4.3}$$

We now examine what $\sum_{k=-\infty}^{\infty} \phi(s-k)e^{ik\omega}$ is.

$$\begin{aligned}\sum_{k=-\infty}^{\infty} \phi(s-k)e^{ik\omega} &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{\phi}(\omega_1)e^{i\omega_1(s-k)}d\omega_1 e^{ik\omega} \\ &= \int_{-\infty}^{\infty} \hat{\phi}(\omega_1) \left(\frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{ik(\omega-\omega_1)} \right) e^{i\omega_1 s} d\omega_1 \\ &= \int_{-\infty}^{\infty} \hat{\phi}(\omega_1) \left(\sum_{k=-\infty}^{\infty} \delta(\omega-\omega_1+2k\pi) \right) e^{i\omega_1 s} d\omega_1 \\ &= \sum_{k=-\infty}^{\infty} \hat{\phi}(\omega+2k\pi)e^{i(\omega+2k\pi)s}.\end{aligned}$$

Thus, by (4.3), we have

$$\int_{-T}^{T} g(s) \sum_{k=-\infty}^{\infty} \hat{\phi}(\omega+2k\pi)e^{i(\omega+2k\pi)s}ds = 0 \quad \forall \omega \in \mathbf{R},$$

or

$$\sum_{k=-\infty}^{\infty} \hat{\phi}(\omega+2k\pi) \int_{-T}^{T} g(s)e^{i(\omega+2k\pi)s}ds = 0 \quad \forall \omega \in \mathbf{R}. \tag{4.4}$$

For $|\omega| < a\pi$, by combining (4.1) and (4.4), we have

$$\int_{-T}^{T} g(s)e^{i\omega s}ds = 0.$$

Let

$$\hat{g}_T(\omega) = \int_{-T}^{T} g(s)e^{i\omega s}ds.$$

We have $\hat{g}_T(\omega) = 0$ for all $\omega \in \mathbf{R}$, since it is time-limited. This implies that $g(s) = 0$ for $s \in [-T, T]$ almost surely and that $Q(s,t)$ is positive definite in $[-T, T]^2$ for any $T > 0$. □

We give some examples satisfying the conditions in Theorem 4 and, therefore, the conditions in Theorem 3 below.

Example 1 (The sinc wavelet). It is clear that the sinc wavelet satisfies the condition in Theorem 4. The well-known convergence result of band-limited signal extrapolation is in fact a special case of Theorem 4. □

Example 2 (The Meyer wavelets). Define $\phi(t)$ by

$$\hat{\phi}(\omega) = \begin{cases} 1, & |\omega| \le \frac{2}{3}\pi, \\ \cos(\frac{\pi}{2}\nu(\frac{3}{2\pi}|\omega|-1)), & \frac{2}{3}\pi \le |\omega| \le \frac{4}{3}\pi, \\ 0 & \text{otherwise,} \end{cases}$$

where ν is a real function in C^k or C^∞ satisfying

$$\nu(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1, & \text{if } x \geq 1, \end{cases} \tag{4.5}$$

and

$$\nu(x) + \nu(1-x) = 1, \qquad x \in \mathbf{R}. \tag{4.6}$$

Clearly, the Meyer wavelets satisfy the condition in Theorem 4.

Example 3 (The cardinal Meyer wavelets). The function $\phi(t)$ is called a cardinal scaling function and $\psi(t)$ is the associated cardinal wavelet if they satisfy the properties given in §2.2 and

$$\phi(n) = \begin{cases} 1, & n = 0, \\ 0, & n = \pm 1, \pm 2, \dots. \end{cases} \tag{4.7}$$

It is clear that condition (4.7) is equivalent to

$$\sum_{n=-\infty}^{\infty} \hat{\phi}(\omega + 2n\pi) = 1 \quad \forall \omega \in \mathbf{R}. \tag{4.8}$$

For more details about cardinal wavelets, we refer to [1] and [27].

In the following we construct real and cardinal Meyer wavelets as follows. We define $\phi(t)$ by

$$\hat{\phi}(\omega) = \begin{cases} 1, & |\omega| \leq \frac{2}{3}\pi, \\ \frac{1}{2}\left(1 + e^{i\pi\nu(\frac{3}{2\pi}\omega - 1)}\right), & \frac{2}{3}\pi < \omega < \frac{4}{3}\pi, \\ \frac{1}{2}\left(1 - e^{i\pi\nu(\frac{3}{2\pi}(\omega+2\pi)-1)}\right), & -\frac{4}{3}\pi < \omega < -\frac{2}{3}\pi, \\ 0, & |\omega| \geq \frac{4}{3}\pi, \end{cases} \tag{4.9}$$

where $\nu(x)$ is as in Example 2.

We first show that $\phi(t)$ is real. To do so, it is enough to show $\hat{\phi}(\omega) = \hat{\phi}^*(-\omega)$ for all real ω. For $\phi(t)$ with (4.9), it is enough to show $\hat{\phi}(\omega) = \hat{\phi}^*(-\omega)$ for $\frac{2}{3}\pi < \omega < \frac{4}{3}\pi$. When $\frac{2}{3}\pi < \omega < \frac{4}{3}\pi$,

$$\begin{aligned} \hat{\phi}^*(-\omega) &= \frac{1}{2}\left(1 - e^{-i\pi\nu(\frac{3}{2\pi}(-\omega+2\pi)-1)}\right) \\ &= \frac{1}{2}\left(1 - e^{-i\pi\nu(2-\frac{3}{2\pi}\omega)}\right) \\ &\overset{(1)}{=} \frac{1}{2}\left(1 - e^{-i\pi(1-\nu(-1+\frac{3}{2\pi}\omega))}\right) \\ &= \frac{1}{2}\left(1 + e^{i\pi\nu(\frac{3}{2\pi}\omega-1)}\right) \\ &= \hat{\phi}(\omega), \end{aligned}$$

where step (1) is from (4.6). This proves $\phi(t)$ is real.

We now prove that $\phi(t)$ is a scaling function. To do so, with the definition (4.9) of $\phi(t)$ and the results from [8] and [13], we only need to show

$$\sum_{n=-\infty}^{\infty} |\hat{\phi}(\omega + 2n\pi)|^2 = 1 \quad \forall \omega \in \mathbf{R}. \tag{4.10}$$

By using (4.9), we see that to prove (4.10), it is enough to prove that when $\frac{2}{3}\pi < \omega < \frac{4}{3}\pi$,

$$|\hat{\phi}(\omega)|^2 + |\hat{\phi}(\omega - 2\pi)|^2 = 1.$$

This is achieved by noticing that

$$\hat{\phi}(\omega) = \frac{1}{2}\left(1 + e^{i\pi\nu(\frac{3}{2\pi}\omega - 1)}\right),$$

and

$$\hat{\phi}(\omega - 2\pi) = \frac{1}{2}\left(1 - e^{i\pi\nu(\frac{3}{2\pi}\omega - 1)}\right).$$

To see that $\phi(t)$ is cardinal, from (4.8) it is enough to show

$$\sum_{n=-\infty}^{\infty} \hat{\phi}(\omega + 2n\pi) = 1 \quad \forall \omega \in \mathbf{R}. \tag{4.11}$$

Equation (4.11) can be proved in a similar manner as that of (4.10).

To conclude, $\phi(t)$ with (4.9) is real, cardinal, and band-limited. Clearly, $\phi(t)$ satisfies the condition in Theorem 4 and therefore the conditions in Theorem 3. Since $\hat{\phi}(\omega)$ has the same regularity as that of ν, $\phi(t)$ may have fast decay when t goes to infinity. As mentioned in [8], a candidate of $\nu(x)$ in (4.9) is

$$\nu(x) = x^4(35 - 84x + 80x^2 - 20x^3), \qquad 0 \le x \le 1, \tag{4.12}$$

which has $C^{4-\epsilon}$ regularity. For more about Meyer wavelets, we refer the reader to [8] and [13]. □

Although many wavelets as discussed above satisfy the conditions in Theorem 3, wavelets with compact support such as the Daubechies wavelets are excluded. In the next section, we consider the discrete-time signal extrapolation problem where the conditions on wavelet bases for convergence are weaker than those in Theorem 3 so that Daubechies wavelets may be included.

5. Extrapolation of discrete-time signals. We consider the discrete version of the GPG algorithm in §3 for handling discrete-time signals. This is what we need practically. Recall that the DWT of a sequence $c_{J,n} = x[n]$ can be implemented via (2.8) for a certain integer J. The discrete sequence $c_{J,n}$ is said to be (J, K) *scale-time limited* for certain integers J and $K > 0$ if its DWT coefficients (with lowest resolution J_0) satisfy that coefficients $c_{J_0,k}$ and $b_{j-1,k}$ may take nonzero values only when $|k| \le K$ and $J_0 \le j < J$. When J and K are sufficiently large, the (J, K) scale-time limited sequence provides a practical discrete-time signal model.

5.1. The DGPG algorithm. Let $x[n]$ be a (J, K) scale-time limited sequence. The values of $x[n]$, $n \in \mathcal{N}$, are given, where the cardinality $|\mathcal{N}| = N$ is finite. The extrapolation problem is to recover $x[n]$ for $n \notin \mathcal{N}$. Let $P_{\mathcal{N}}$ and $P_{J,K}$ be the following operators:

$$P_{\mathcal{N}}y[n] = \begin{cases} y[n], & n \in \mathcal{N}, \\ 0, & n \notin \mathcal{N}, \end{cases}$$

and

$$P_{J,K}d_{j,k} = \begin{cases} d_{j,k}, & |k| \le K \text{ and } J_0 \le j < J, \\ 0 & \text{otherwise.} \end{cases}$$

Let I be the identity operator and $\mathcal{D}$ and $\mathcal{D}^{-1}$ be the DWT and IDWT operators. The DGPG algorithm can be stated as follows.

THE DGPG ALGORITHM.

$$x^{(0)}[n] = P_{\mathcal{N}}x[n] \tag{5.1}$$

for $l = 0, 1, 2, \ldots,$

$$x^{(l+1)}[n] = P_{\mathcal{N}}x[n] + (I - P_{\mathcal{N}})\mathcal{D}^{-1}P_{J,K}\mathcal{D}x^{(l)}[n]. \tag{5.2}$$

Similar to the continuous-time GPG algorithm, it can be shown that the error energy of the DGPG algorithm is monotonically decreasing, i.e.,

$$\sum_n |x^{(l+1)}[n] - x[n]|^2 \le \sum_n |x^{(l)}[n] - x[n]|^2 \qquad l = 0, 1, 2, \ldots.$$

5.2. Convergence and uniqueness results. To show the convergence of the DGPG algorithm (5.1)–(5.2), some tools are needed. We introduce two operators H and G related to the quadrature mirror filters $H(\omega)$ and $G(\omega)$ in (2.4) as follows:

$$Hy[k] \triangleq \sqrt{2}\sum_n h_{n-2k}y[n], \qquad \text{and} \qquad Gy[k] \triangleq \sqrt{2}\sum_n g_{n-2k}y[n].$$

Let H^* and G^* be their duals, respectively, i.e.,

$$H^*y[n] \triangleq \sqrt{2}\sum_k h_{n-2k}y[k], \qquad \text{and} \qquad G^*y[n] \triangleq \sqrt{2}\sum_k g_{n-2k}y[k].$$

Then, from (2.9), we have

$$\begin{aligned} x[n] &= \left(H^*c_{J-1,k} + G^*b_{J-1,k}\right)[n] \\ &= \left(H^*\left(H^*c_{J-2,k} + G^*b_{J-2,k}\right) + G^*b_{J-1,k}\right)[n] \\ &= \left((H^*)^{J-J_0}c_{J_0,k} + (H^*)^{J-J_0-1}G^*b_{J_0,k} + \cdots + H^*G^*b_{J-2,k} + G^*b_{J-1,k}\right)[n]. \end{aligned}$$

We can rewrite the above equation as

$$x[n] = \mathbf{w}_n\mathbf{p}, \qquad n \in \mathbf{Z}, \tag{5.3}$$

where $\mathbf{p}$ and $\mathbf{w}_n$ are, respectively, column and row vectors of length $(2K+1)(J - J_0 + 1)$ of the form

$$\begin{aligned} \mathbf{p} &= \left(\mathbf{c}_{J_0}, \mathbf{b}_{J_0}, \mathbf{b}_{J_0+1}, \ldots, \mathbf{b}_{J-1}\right)^T, \\ \mathbf{w}_n &= \left((H^*)_n^{J-J_0}, \left((H^*)^{J-J_0-1}G^*\right)_n, \ldots, (H^*G^*)_n, G_n^*\right), \end{aligned}$$

and where

$$\begin{aligned} \mathbf{c}_{J_0} &= \left(c_{J_0,-K}, c_{J_0,-K+1}, \ldots, c_{J_0,K}\right), \\ \mathbf{b}_j &= \left(b_{j,-K}, b_{j,-K+1}, \ldots, b_{j,K}\right), \\ G_n^* &= \sqrt{2}\left(g_{-K-2n}, g_{-K+1-2n}, \ldots, g_{K-2n}\right), \\ (H^*G^*)_n &= 2\left(\sum_{n_1} h_{n_1-2n}g_{-K-2n_1}, \sum_{n_1} h_{n_1-2n}g_{-K+1-2n_1}, \ldots, \sum_{n_1} h_{n_1-2n}g_{K-2n_1}\right), \end{aligned}$$

$$\left((H^*)^j G^*\right)_n = (\sqrt{2})^{j+1} \left(\sum_{n_1}\sum_{n_2}\cdots\sum_{n_j} h_{n_j-2n}h_{n_{j-1}-2n_j}\cdots h_{n_1-2n_2}g_{-K-2n_1}, \right.$$
$$\sum_{n_1}\sum_{n_2}\cdots\sum_{n_j} h_{n_j-2n}h_{n_{j-1}-2n_j}\cdots h_{n_1-2n_2}g_{-K+1-2n_1},$$
$$\ldots,$$
$$\left. \sum_{n_1}\sum_{n_2}\cdots\sum_{n_j} h_{n_j-2n}h_{n_{j-1}-2n_j}\cdots h_{n_1-2n_2}g_{K-2n_1} \right),$$

$$(H^*)_n^{J_1} = (\sqrt{2})^{J_1} \left(\sum_{n_1}\sum_{n_2}\cdots\sum_{n_{J_1-1}} h_{n_{J_1}-2n}h_{n_{J_1-1}-2n_{J_1}}\cdots h_{n_1-2n_2}h_{-K-2n_1}, \right.$$
$$\sum_{n_1}\sum_{n_2}\cdots\sum_{n_{J_1-1}} h_{n_{J_1}-2n}h_{n_{J_1-1}-2n_{J_1}}\cdots h_{n_1-2n_2}h_{-K+1-2n_1},$$
$$\ldots,$$
$$\left. \sum_{n_1}\sum_{n_2}\cdots\sum_{n_{J_1-1}} h_{n_{J_1}-2n}h_{n_{J_1-1}-2n_{J_1}}\cdots h_{n_1-2n_2}h_{K-2n_1} \right)$$

for $1 \le j \le J - J_0 - 1$ and $J_1 = J - J_0$. Now, by letting

$$\mathcal{N} = \{m_1, m_2, \ldots, m_N \colon m_1 < m_2 < \cdots < m_N\},$$

we obtain the following linear system

$$\mathbf{x} = \mathbf{W}\mathbf{p}, \tag{5.4}$$

where

$$\mathbf{x} = (x[m_1], x[m_2], \ldots, x[m_N])^T \quad \text{and} \quad \mathbf{W} = \left(\mathbf{w}_{m_1}^T, \mathbf{w}_{m_2}^T, \ldots, \mathbf{w}_{m_N}^T\right)^T \tag{5.5}$$

are known. If $\mathbf{p}$ can be uniquely solved from (5.4), then $x[n]$ with $n \notin \mathcal{N}$ can be extrapolated from $x[n]$ with $n \in \mathcal{N}$. For $\mathbf{p}$, we have $(2K+1)(J-J_0+1) \triangleq r_0$ unknowns. Therefore, to uniquely determine $x[n]$, it is required that $N \ge r_0$ and that the rank of $\mathbf{W}$ has to be r_0. The above arguments prove the following theorem.

THEOREM 5. *Let $x[n]$ be a (J, K) scale-time limited sequence. Then, $x[n]$ can be uniquely determined from $x[n]$, $n \in \mathcal{N}$, if and only if the rank of $\mathbf{W}$ is $r_0 = (2K+1)(J-J_0+1)$.*

To extrapolate $x[n]$ outside $\mathcal{N}$ via the DGPG algorithm is equivalent to the solution of (5.4) for $\mathbf{p}$. There are two reasons to avoid solving (5.4) directly. One is that the direct computation of $\mathbf{W}$ is expensive. The other is that, even though $\mathbf{W}$ is known, to solve the linear system (5.4) is also expensive. We now go back to the convergence of the DGPG algorithm.

THEOREM 6. *Let $x[n]$ be a (J, K) scale-time limited sequence. If the rank of $\mathbf{W}$ is $r_0 = (2K+1)(J-J_0+1)$, then*

$$\sum_{n=-\infty}^{\infty} |x^{(l)}[n] - x[n]|^2 \to 0 \quad \text{as } l \to \infty. \tag{5.6}$$

On the other hand, if (5.6) is true for all (J, K) scale-time limited sequences, then the rank of the matrix $\mathbf{W}$ is r_0.

Proof. When the rank of $\mathbf{W}$ is $r_0 = (2K+1)(J-J_0+1)$, the DWT coefficients $\mathbf{p}$ is uniquely determined from $\mathbf{x}$ by solving the linear system (5.4). Also, $\mathbf{W}\mathbf{W}^T$ has rank r_0 and is a nonnegative definite matrix. Let λ_i, $i = 1, 2, \ldots, N$, be the eigenvalues of $\mathbf{W}\mathbf{W}^T$ and $\mathbf{q}_i$, $i = 1, 2, \ldots, N$, be the corresponding eigenvectors, i.e.,

$$\mathbf{W}\mathbf{W}^T\mathbf{q}_i = \lambda_i \mathbf{q}_i \qquad i = 1, 2, \ldots, N, \tag{5.7}$$

where

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{r_0} > \lambda_{r_0+1} = \cdots = \lambda_N = 0,$$

$\mathbf{q}_i$ forms an orthonormal basis of $\mathbf{C}^N$, and where $\mathbf{C}$ denotes the set of complex numbers. Let

$$(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{r_0}) \triangleq \mathbf{W}^T(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{r_0}).$$

Since $\mathbf{q}_i$, $1 \leq i \leq r_0$, are linearly independent and the matrix $\mathbf{W}^T$ has rank r_0, $\mathbf{y}_i$ with $1 \leq i \leq r_0$ forms a basis of $\mathbf{C}^{r_0}$. Therefore, there are r_0 constants a_i such that

$$\mathbf{p} = \sum_{i=1}^{r_0} a_i \mathbf{y}_i.$$

Only $\mathbf{q}_i[n]$ with $n \in \mathcal{N}$ are given in (5.7). For $1 \leq i \leq r_0$, we extend $\mathbf{q}_i[n]$ from $n \in \mathcal{N}$ to all integers via

$$\tilde{\mathbf{q}}_i[n] = \frac{1}{\lambda_i}\mathbf{w}_n \mathbf{W}^T \mathbf{q}_i, \quad n \in \mathbf{Z}. \tag{5.8}$$

By (5.3) and (5.8), we have

$$x[n] = \mathbf{w}_n \mathbf{p} = \mathbf{w}_n \sum_{i=1}^{r_0} a_i \mathbf{y}_i = \sum_{i=1}^{r_0} a_i \mathbf{w}_n \mathbf{W}^T \mathbf{q}_i = \sum_{i=1}^{r_0} a_i \lambda_i \tilde{\mathbf{q}}_i[n]. \tag{5.9}$$

Note that the reason for proving (5.9) is similar to that for proving $f(t) \in U_J$ in the proof of Theorems 2 or 3.

We now prove that $\lambda_i \leq 1$ for $1 \leq i \leq r_0$. For $1 \leq i \leq r_0$,

$$\begin{aligned}
\|\tilde{\mathbf{q}}_i\|^2 &= \sum_{n=-\infty}^{\infty} |\tilde{\mathbf{q}}_i[n]|^2 = \frac{1}{\lambda_i^2} \sum_{n=-\infty}^{\infty} |\mathbf{w}_n \mathbf{W}^T \mathbf{q}_i|^2 = \frac{1}{\lambda_i^2} \|\mathcal{D}^{-1}\mathbf{W}^T\mathbf{q}_i\|^2 \\
&= \frac{1}{\lambda_i^2}\|\mathbf{W}^T\mathbf{q}_i\|^2 = \frac{1}{\lambda_i^2}\|\mathcal{D}P_{\mathcal{N}}\tilde{\mathbf{q}}_i\|^2 = \frac{1}{\lambda_i^2}\|P_{\mathcal{N}}\tilde{\mathbf{q}}_i\|^2 \\
&\leq \frac{1}{\lambda_i^2}\|\tilde{\mathbf{q}}_i\|^2,
\end{aligned}$$

where we use the properties that both $\mathcal{D}$ and $\mathcal{D}^-1$ preserve the total energy and that $\mathbf{W}^T$ behaves like $\mathcal{D}P_{\mathcal{N}}$ when operating on (J, K) scale-time limited sequences. Thus, $\lambda_i \leq 1$ for $1 \leq i \leq r_0$. Similar to the proof of the error formula (3.12), we have

$$x[n] - x^{(l)}[n] = (I - P_{\mathcal{N}}) \sum_{i=1}^{r_0} a_i \lambda_i (1 - \lambda_i)^l \tilde{\mathbf{q}}_i[n]. \tag{5.10}$$

By $0 < \lambda_i \leq 1$, $1 \leq i \leq r_0$, and similar to the proof of the standard PG algorithm,

$$\sum_{n=-\infty}^{\infty} |x[n] - x^{(l)}[n]|^2 \to 0 \text{ as } l \to \infty.$$

This proves the first part of Theorem 6. If (5.6) is true for all (J, K) scale-time limited sequences $x[n]$, then $x[n]$ is uniquely determined by $x[n]$, $n \in \mathcal{N}$. By Theorem 5, the second part of Theorem 6 is also proved. □

Suppose that $\mathbf{W}$ has rank r_0 and $\lambda_1, \lambda_2, \ldots, \lambda_N$ are eigenvalues of the matrix $\mathbf{WW}^T$ arranged in a descending order, i.e.,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{r_0} > \lambda_{r_0+1} = \cdots = \lambda_N = 0.$$

Based on (5.10), we obtain an error estimate

$$|x^{(l)}[n] - x[n]| \leq O((1-\lambda_{r_0})^l)$$

for sufficiently large l. Even when the rank of $\mathbf{W}$ is r_0, the conventional PG (DPG) algorithm for band-limited signals [20], [22] usually converges slowly. This is because the condition number of the corresponding $\mathbf{WW}^T$ is usually quite large due to the smoothness of the sinc function. In contrast, by using suitable wavelet bases the condition number of $\mathbf{WW}^T$ is smaller so that a faster convergence rate can be achieved (see numerical experiments in §6).

5.3. Connection between continuous-time and discrete-time extrapolation. We examined continuous-time signals in the wavelet subspace V_J in §3, where each $f(t) \in V_J$ has the form

$$f(t) = \sum_{k=-\infty}^{\infty} c_{J,k}\phi_{Jk}(t) = \sum_{k=-\infty}^{\infty} c_{J_0,k}\phi_{J_0k}(t) + \sum_{J_0 \leq j < J} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{jk}(t).$$

In practice, $f(t)$ is small for large $|t|$ so that $c_{J_0,k}$ and $b_{j,k}$ are also small for large $|k|$. Thus, it is important to consider signals in the following subspace of V_J,

$$V_{J,K} \triangleq \left\{ f(t) : f(t) = \sum_{k=-K}^{K} c_{J_0,k}\phi_{J_0k}(t) + \sum_{J_0 \leq j < J} \sum_{k=-K}^{K} b_{j,k}\psi_{jk}(t) \right.$$
$$\left. \text{for some constants } c_{J_0,k}, b_{j,k} \right\}.$$

We call signals in $V_{J,K}$ as (J, K) scale-time limited. This explains the motivation of the definition of scale-time limited sequences. For $f(t) \in V_{J,K}$, we have

$$f(t) = \sum_k c_{J,k}\phi_{Jk}(t) = \sum_{k=-K}^{K} c_{J_0,k}\phi_{J_0k}(t) + \sum_{J_0 \leq j < J} \sum_{k=-K}^{K} b_{j,k}\psi_{jk}(t),$$

where

$$c_{J,k} = \langle f, \phi_{Jk} \rangle, \quad c_{J_0,k} = \langle f, \phi_{J_0k} \rangle, \quad b_{j,k} = \langle f, \psi_{jk} \rangle.$$

Since $\phi(t)$ behaves like a lowpass filter, $c_{J,k}$ is close to $f(k/2^J)$ [12], [17], [26] for sufficiently large J. Therefore, we can replace $c_{J,k}$ or $x[k]$ with samples $f(k/2^J)$ and use the discrete-time GPG algorithm to provide a good approximation for continuous-time signal extrapolation. In [25], it was proved that the discrete-time signal extrapolation converges to the continuous-time one when the sampling rate in the given interval goes to infinity.

6. Numerical examples. In this section, numerical examples are presented to illustrate the theory for discrete-time signal extrapolation developed in §5. Two wavelet bases are considered. One is the Meyer wavelet given in Example 3 of §4 with ν defined by (4.12), which is simply called the Meyer M_4 basis. The other is the Daubechies D_4 basis [7]. Note that the Meyer M_4 basis satisfies the convergence conditions given in Theorem 3 for the continuous-time signal extrapolation while the Daubechies D_4 basis does not. However, since we perform discrete-time signal extrapolation numerically, the convergence condition in Theorem 6 is more relevant. We show numerically the rank of the matrix $\mathbf{W}$ below to verify this convergence theorem for both M_4 and D_4 wavelet bases.

6.1. The Meyer M_4 basis. The test signal $x[n]$ as given in Fig. 1 is a $(J, K) = (4, 2)$ scale-time limited signal with $J_0 = 1$, whose DWT coefficients are $c_{1,k} = 0$, $b_{j,k} = 1.0$, $1 \le j \le 3$, $1 \le k \le 5$, and the wavelet basis is M_4. Since the low resolution coefficients $c_{1,k}$ are 0 for this case, $x[n]$ is in fact generated by the 15 wavelet coefficients $b_{j,k}$. Thus, $x[n]$ has $r_0 = 15$ degrees of freedom. We consider four test cases: the values $x[n]$ are given for $n \in \mathcal{N}$ with $\mathcal{N} = \{n : 1 \le n \le 15\}$, $\{n : 1 \le n \le 30\}$, $\{n : 1 \le n \le 42\}$, and $\{n : 1 \le n \le 45\}$.

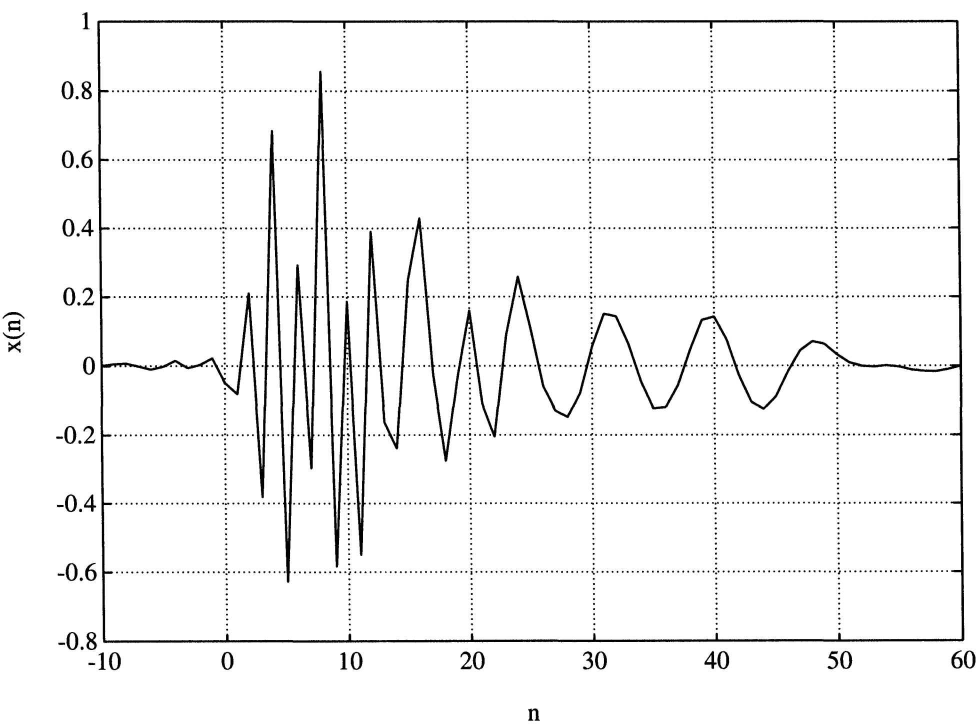

FIG. 1. *The test signal $x[n]$ with the Meyer wavelet M_4.*

We use the DGPG algorithm (5.1)–(5.2) to compute $x^{(l)}[n]$ iteratively. Let e_l denote the 2-norm of the extrapolation error at lth iteration, i.e.,

$$e_l \triangleq \left(\sum_n |x^{(l)}[n] - x[n]|^2 \right)^{1/2}.$$

The convergence history of e_l as a function of the number of iteration is plotted in Fig. 2, where the errors remain about the same for $N = 15$ while the errors decrease as the iteration proceeds for $N = 30$, 42, and 45.

We show the rank of the matrix $\mathbf{W}$ and the first 15 eigenvalues of $\mathbf{W}\mathbf{W}^T$ with the wavelet M_4 in Table 1. It is clear that the condition for Theorem 6 is satisfied in cases with $N = 30, 42,$

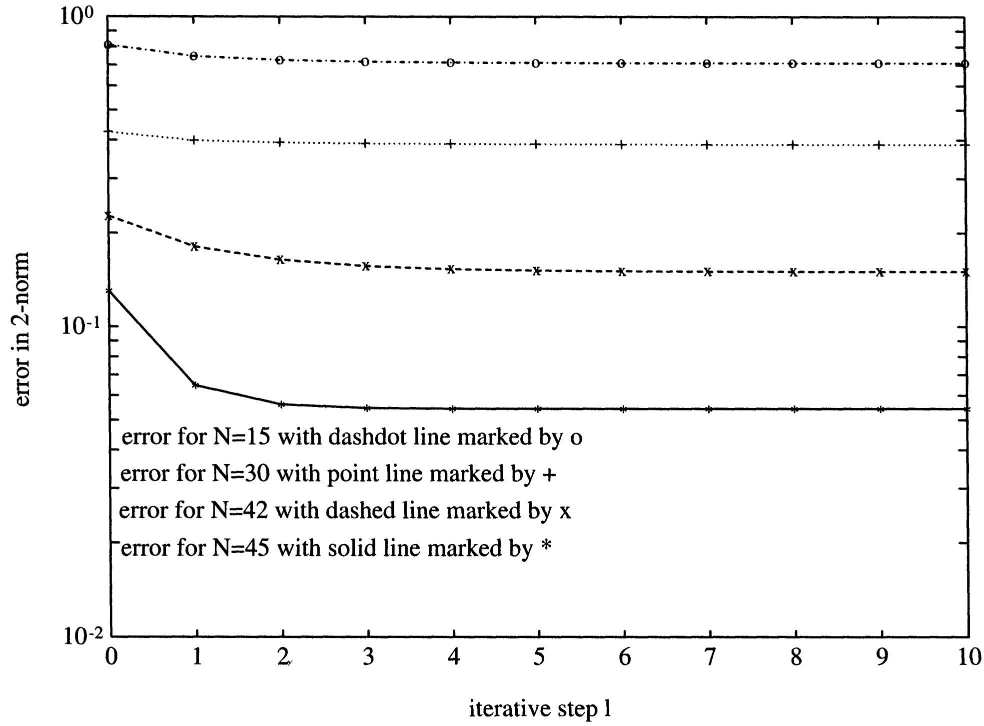

FIG. 2. *Convergence history of the 2-norm of extrapolation errors.*

TABLE 1

	First 15 eigenvalues of $\mathbf{WW}^T$					Rank of $\mathbf{W}$
	0.99997	0.99872	0.95037	0.76713	0.62626	
$N = 15$	0.31103	0.30700	0.24620	0.11131	0.04713	14
	0.00073	0.00044	0.00009	0.00001	0.00000	
	1.00000	0.99906	0.95079	0.76827	0.62761	
$N = 30$	0.32933	0.31993	0.29804	0.29778	0.29363	15
	0.15791	0.15750	0.10932	0.01303	0.00051	
	1.00000	0.99906	0.95079	0.76827	0.62761	
$N = 42$	0.32951	0.32013	0.29860	0.29792	0.29412	15
	0.16446	0.16273	0.15191	0.14297	0.04114	
	1.00000	0.99906	0.95079	0.76827	0.62761	
$N = 45$	0.32951	0.32013	0.29860	0.29792	0.29412	15
	0.16477	0.16428	0.15236	0.14580	0.10901	

and 45 so that the algorithm converges for these cases. However, the convergence rate is quite slow. This can be explained by the large condition number of the matrix $\mathbf{WW}^T$.

6.2. The Daubechies D_4 basis. The test signal $x[n]$ as given in Fig. 3 is constructed as the one given in §6.1 except that the wavelet basis is replaced by the Daubechies D_4 basis.

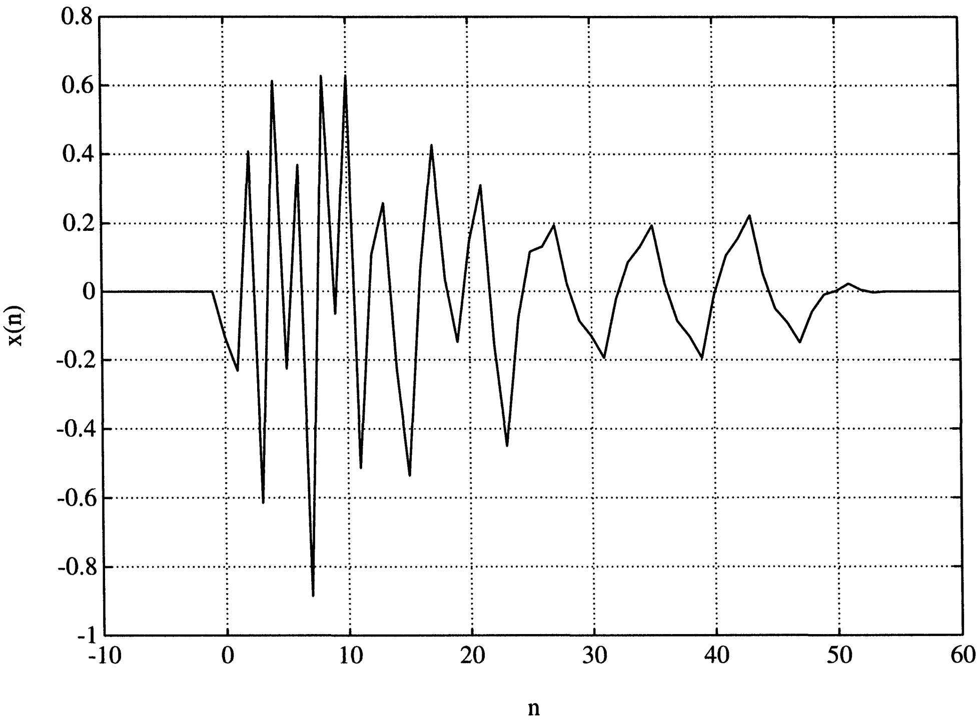

FIG. 3. *The test signal $x[n]$ with the Daubechies wavelet D_4.*

The convergence history of e_l as a function of the number of iterations is plotted in Fig. 4. One can clearly see that the errors remain about the same for $N = 15, 30$ while the errors decrease significantly as the iteration proceeds for $N = 42, 45$. In other words, the DGPG algorithm converges to the correct solution only when $N = 42$ and 45. This convergence behavior can be explained by examining Table 2, where we show the ranks and the first 15 eigenvalues of $\mathbf{WW}^T$ for all four cases, where $\mathbf{W}$ is defined by (5.5). Note that the ranks of $\mathbf{W}$ are equal to the degree of freedom of $x[n]$ only when $N = 42$ and 45. For $N = 15$ (or $N = 30$), there are 4 (or 1) zero eigenvalues among the largest 15 eigenvalues.

Note also that to determine $x[n]$ outside $\mathcal{N}$ is equivalent to the determination of wavelet coefficients $b_{j,k}$, $1 \le j \le 3$, and $1 \le k \le 5$, which are all equal to 1 by design. Let $b_{j,k}^{(l)}$ be the corresponding DWT coefficients of $x^{(l)}[n]$. The convergence of $x^{(l)}[n]$ to $x[n]$ is equivalent to the convergence of $b_{j,k}^{(l)}$ to $b_{j,k}$. We show the values of $b_{j,k}^{(5)}$ and $b_{j,k}^{(10)}$ in Tables 3 and 4. We see that $b_{j,k}^{(l)}$ converge quite fast to the true values 1 of $b_{j,k}$ for $N = 42, 45$. In contrast, there are 4 (or 1) $b_{j,k}^{(l)}$'s which do not converge to 1 for $N = 15$ (or $N = 30$) due to the zero eigenvalues existing in the matrix $\mathbf{WW}^T$.

7. Conclusions and extensions. In this paper, we proposed a new iterative algorithm which extends the well-known PG algorithm for band-limited signal extrapolation to signals in the wavelet subspaces. The convergence of the GPG algorithm and the uniqueness of

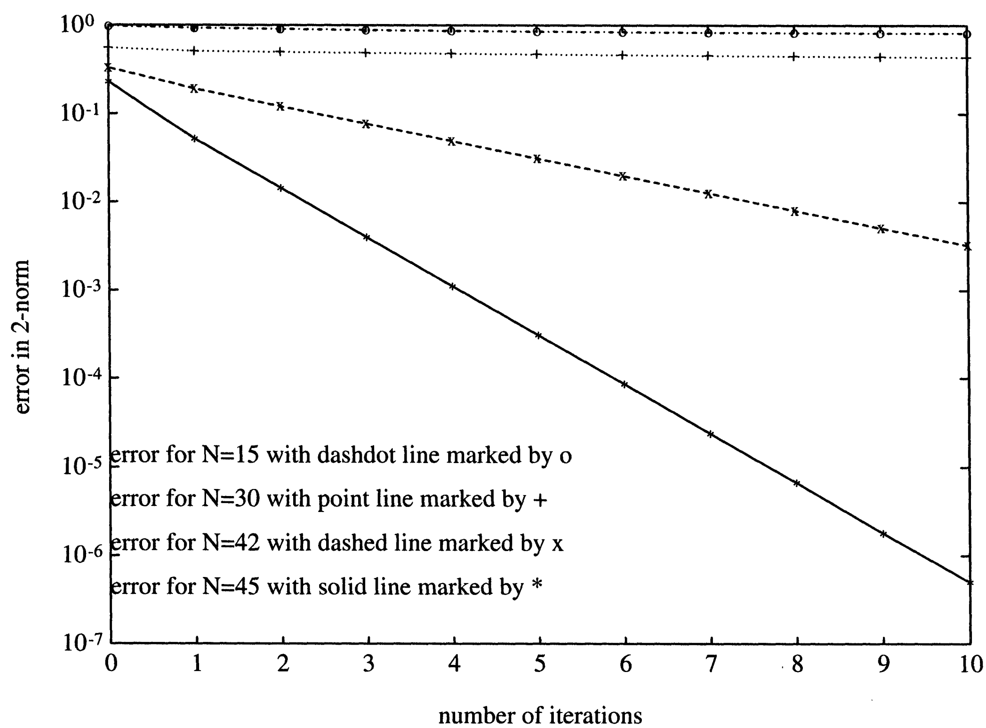

FIG. 4. *Convergence history of the* 2*-norm of extrapolation errors.*

TABLE 2

	First 15 eigenvalues of $\mathbf{W}\mathbf{W}^T$					Rank of **W**
	1.00000	1.00000	1.00000	1.00000	1.00000	
$N = 15$	1.00000	1.00000	0.97964	0.96402	0.11057	11
	0.00065	0.00000	0.00000	0.00000	0.00000	
	1.00000	1.00000	1.00000	1.00000	1.00000	
$N = 30$	1.00000	1.00000	1.00000	1.00000	1.00000	14
	1.00000	0.97843	0.79403	0.40207	0.00000	
	1.00000	1.00000	1.00000	1.00000	1.00000	
$N = 42$	1.00000	1.00000	1.00000	1.00000	1.00000	15
	1.00000	1.00000	0.99793	0.97843	0.36378	
	1.00000	1.00000	1.00000	1.00000	1.00000	
$N = 45$	1.00000	1.00000	1.00000	1.00000	1.00000	15
	1.00000	1.00000	1.00000	0.97843	0.72182	

extrapolated signals for both continuous-time and discrete-time cases are investigated. Several conditions on signals and wavelet bases to achieve convergence and uniqueness are described.

TABLE 3

	$b_{j,k}^{(5)}, 1 \leq j \leq 3, 1 \leq k \leq 5$				
	1.08760	0.59008	0.00000	0.00000	0.00000
$N = 15$	1.00013	1.00000	0.99501	0.47866	0.00000
	1.00019	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	0.92549	0.15589	0.00000
$N = 30$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	0.99550	0.92973
$N = 42$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	0.99954
$N = 45$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000

TABLE 4

	$b_{j,k}^{(10)}, 1 \leq j \leq 3, 1 \leq k \leq 5$				
	1.06050	0.85387	0.00000	0.00000	0.00000
$N = 15$	1.00009	1.00000	0.99722	0.64568	0.00000
	1.00013	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	0.93937	0.31247	0.00000
$N = 30$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	0.99953	0.99268
$N = 42$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000
$N = 45$	1.00000	1.00000	1.00000	1.00000	1.00000
	1.00000	1.00000	1.00000	1.00000	1.00000

We observe from the experiments in §§6.1 and 6.2 that the convergence of the DGPG algorithm not only depends on the wavelet basis but also the length N of the observation sequence. Roughly speaking, the uniqueness and convergence of the GPG or DGPG algorithm depend on the smoothness of the wavelet bases. If the algorithm converges, its convergence rate depends on the time-localization property of wavelet bases. Thus, the desired wavelet bases might be those which are smooth and have a fast decay in the time domain. However, a more quantitative characterization remains to be investigated. There are many interesting topics to be studied as extensions of this work. The list at least includes: (1) to search more wavelet bases satisfying the conditions in Theorem 3; (2) to establish a relationship between the conditions in Theorem 3 and the one in Theorem 6 for continuous- and discrete-time signal extrapolation; (3) to perform signal extrapolation with noisy data; (4) to investigate the properties of the eigenvectors of the kernel $Q_J(s, t)$ in §3. Since various bases can be provided by wavelet theory, it is believed that the proposed GPG algorithm should have many potential applications in signal recovery.

Acknowledgments. The authors are grateful to the referees for their patience, encouragement, and valuable suggestions. Especially, one referee pointed out that the space U_J in (3.10) is identical to the range space $P^J P_T(L^2(\mathbf{R}))$, which helps to simplify the derivation in §3.

REFERENCES

[1] A. Aldroubi and M. Unser, *Families of wavelet transforms in connection with Shannon's sampling theory and the gabor transform*, in Wavelets: A Tutorial in Theory and Applications, C. K. Chui, ed., Academic Press, New York, 1992, pp. 509–528.

[2] R. P. Boas, *Entire Functions*, Academic, New York, 1954.

[3] P. L. Butzer, *A survey of the Whittaker–Shannon sampling theorem and some of its extensions*, J. Math. Res. Exposition, 3 (1983), pp. 185–212.

[4] J. A. Cadzow, *An extrapolation procedure for band-limited signals*, IEEE Trans. Acoust., Speech Signal Process., 27 (1979), pp. 1208–1209.

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

[6] ———, ed., *Wavelets: A Tutorial in Theory and Applications*, Academic Press, New York, 1992.

[7] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 41 (1988), pp. 909–996.

[8] ———, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadephia, PA, 1992.

[9] R. W. Gerchbert, *Super-resolution through error energy reduction*, Optica Acta, 21 (1974), pp. 709–720.

[10] N. E. Hurt, *Phase Retrieval and Zero Crossings*, Kluwer Academic Publishers, Boston, 1989.

[11] A. J. Jerri, *The Shannon sampling theorem—its various extensions and applications: A tutorial review*, Proc. IEEE, 65 (1977), pp. 1565–1596.

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

[13] Y. Meyer, *Ondelettes et Operateurs*, Hermann, Paris, France, 1990.

[14] A. Papoulis, *A new algorithm in spectral analysis and band limited extrapolation*, IEEE Trans. Circuits and Systems, 22 (1975), pp. 735–742.

[15] M. D. Rawn, *Generalized sampling theorems for Bessel-type transforms of bandlimited functions and distributions*, SIAM J. Appl. Math., 49 (1989), pp. 638–649.

[16] F. Riesz and B. S. Nagy, *Functional Analysis*, 4th ed., Frederick Ungar Publishing Co., New York, 1965.

[17] O. Rioul and P. Duhamel, *Fast algorithms for discrete and continuous wavelet transform*, IEEE Trans. Inform. Theory, 38 (1992), pp. 569–586.

[18] J. L. C. Sanz and T. S. Huang, *Discrete and continuous bandlimited signal extrapolation*, IEEE Trans. Acoust., Speech Signal Process., 31 (1983), pp. 1276–1285.

[19] ———, *On the Papoulis-Gerchberg algorithm*, IEEE Trans. Circuits and Systems, (1983), pp. 907–908.

[20] ———, *Some aspects of band-limited signal extrapolation: Models, discrete approximations, and noises*, IEEE Trans. Acoust., Speech Signal Process., 31 (1983), pp. 1492–1501.

[21] ———, *A unified Hilbert space approach to iterative linear signal restoration*, J. Opt. Soc. Amer., 73 (1983), pp. 1455–1465.

[22] H. J. Schlebusch and W. Splettstösser, *On a conjecture of J. L. C. Sanz and T. S. Huang*, IEEE Trans. Acoust., Speech Signal Process., 33 (1985), pp. 1628–1630.

[23] G. G. Walter, *A sampling theorem for wavelet subspaces*, IEEE Trans. Inform. Theory, 38 (1992), pp. 881–884.

[24] X. G. Xia, *An extrapolation for general analytic signals*, IEEE Trans. Signal Process., 40 (1992), pp. 2243–2249.

[25] X. G. Xia, C.-C. J. Kuo, and Z. Zhang, *Connection between continuous and discrete time signal extrapolation in wavelet subspaces*, USC-SIPI Report 224, Dept. of EE-Systems, University of Southern California, Los Angeles, CA, 1992.

[26] ———, *On signal sampling and wavelet coefficient computation*, USC-SIPI Report 215, Dept. of EE-Systems, University of Southern California, Los Angeles, CA, 1992.

[27] X. G. Xia and Z. Zhang, *On sampling theorem, wavelets and wavelet transforms*, IEEE Trans. Signal Process., 41 (1993), pp. 3524–3535.

[28] W. XU AND C. CHAMZAS, *On the extrapolation of band-limited functions with energy constraints*, IEEE Trans. Acoust., Speech Signal Process., 31 (1983), pp. 1222–1234.

[29] M. ZAKAI, *Band-limited functions and the sampling theorem*, Inform. and Control, 8 (1965), pp. 143–158.

[30] X. W. ZHOU AND X. G. XIA, *On a conjecture of band-limited signal extrapolation*, Kexue Tongbao (Chinese Science Bulletin), 23 (1986).

[31] ———, *The extrapolation of high dimensional band-limited signals*, IEEE Trans. Acoust., Speech Signal Process., 37 (1989), pp. 1576–1580.

[32] ———, *A Sanz-Huang's conjecture on band-limited signal extrapolations with noises*, IEEE Trans. Acoust., Speech Signal Process., 37 (1989), pp. 1468–1472.