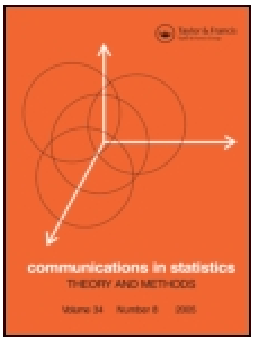

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

Wavelet Like Behavior of Slepian Functions and Their Use in Density Estimation

Gilbert G. Walter & Xiaoping Shen

Communications in Statistics—Theory and Methods, 34: 687–711, 2005

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1081/STA-200052105

Wavelet Like Behavior of Slepian Functions and Their Use in Density Estimation

GILBERT G. WALTER[1] AND XIAOPING SHEN[2]

[1]Department of Mathematics, University of Wisconsin, Milwaukee, Wisconsin, USA
[2]Department of Mathematics, Ohio University, Athens, Ohio, USA

Slepian functions (Prolate Spheroidal Wave Functions) are obtained by maximizing the energy of a σ-bandlimited function (normalized with total energy 1) on a prescribed interval $[-\tau, \tau]$. The solution to this problem leads to an eigenvalue problem $\lambda f(t) = \int_{-\tau}^{\tau} \{\sin \sigma(t-x)/\pi(t-x)\} f(x)dx$, whose solutions, in turn, form an orthogonal sequence $\{\varphi_n\}$. This sequence is a basis of the Paley-Wiener space B_σ of σ-bandlimited functions. For $\sigma = \pi$, integer translates of the Slepian functions of order 0, $\{\varphi_0(t-n)\}$ form a Riesz basis of the same space. Furthermore, by using φ_0 as a scaling function we can construct a wavelet theory based on them. Two methods of density estimations thus naturally arise; one based on the orthogonal system $\{\varphi_n\}$ and the other on the scaling functions $\{\varphi_0(t-n)\}$. The former gives more rapid convergence, while the latter avoids Gibbs phenomenon, is locally positive, and allows the use of thresholding methods. Both approaches exhibit a strong localization property.

Keywords Kernel density estimation; Probability density; Prolate spheroidal wave functions; Slepian semi-wavelets; Wavelet.

Mathematics Subject Classification Primary 60E05, 62F10; Secondary 42A10, 42A15, 42C40.

1. Introduction and Notation

The continuous *prolate spheroidal wave functions* (PSWFs), or as we shall call them, *Slepian functions*, because of a "lucky accident" (Slepian, 1983; Walter, 1992b), were found to be quite useful for analog signal processing. However, the digital revolution left them in the dust since they did not seem naturally adapted to discrete analysis. Yet they have many desirable, even unique, properties that originally made them useful and could lead to new statistical applications. The simplest such involves non parametric density estimation and regression.

Received November 12, 2003; Accepted July 30, 2004
Address correspondence to Xiaoping Shen, Department of Mathematics, Ohio University, Athens, OH 45701, USA; E-mail: shen@math.ohiou.edu

In its simplest form, such density estimation begins with an i.i.d. sample $X_1, X_2, \ldots, X_N$ of the density $f(x)$. This is used to construct an approximation $\hat{f}_h(x)$ to $f(x)$ which converges to it in some way. Many methods have been proposed for constructing such an estimator; the most widely used are kernel methods and orthogonal series methods. Wavelet methods (for details of wavelet theory, see for example, Daubechies, 1992; Walter and Shen, 2000) are a modification of the latter in which a further step, that of thresholding, is added to reduce noise. The two methods discussed in this work are based, respectively, on the Slepian functions themselves and on a family of wavelets based on them.

These functions are closely related to the Shannon sampling theorem (Shannon, 1949) given by the formula

$$f(t) = \sum_{n=-\infty}^{\infty} f(nT)\frac{\sin\sigma(t-nT)}{\sigma(t-nT)}, \quad T = \pi/\sigma, \; t \in R \tag{1}$$

which holds for σ-bandlimited signals with finite energy. This theorem has become a well-known part of both the mathematical and engineering literature (Djokovic and Vaidyanathan, 1997; Higgins, 1996; Vaidyanathan, 2001; Walter and Shen, 2003; and Zayed, 1993). It also falls naturally into a "wavelet" setting since the sinc function appearing in (1), $S(t) = \sin\pi t/\pi t$, is a particular example of a "scaling function" appearing in wavelet theory (Walter, 1992a).

The first Slepian function, which we denote by $\varphi_{0,\sigma,\tau}$, is defined as the one having the maximum energy concentration of a σ-bandlimited function on the interval $(-\tau, \tau)$; that is $\varphi_{0,\sigma,\tau}$ is the function of total energy 1 $(= \|\varphi_{0,\sigma,\tau}\|^2)$ such that

$$\int_{-\tau}^{\tau} |f(t)|^2 dt$$

is maximized. Then, $\varphi_{1,\sigma,\tau}$ is the function with the maximum energy concentration among those functions orthogonal to $\varphi_{0,\sigma,\tau}$, etc. By repeating this we get a sequence $\{\varphi_{n,\sigma,\tau}(t)\}$ which turns out to be an orthonormal basis of the space of σ-bandlimited functions on the real line. There are several other ways of characterizing them:

- as the eigenfunctions of an integral operator whose kernel is the sinc function of (1):

$$\int_{-\tau}^{\tau} \varphi_{n,\sigma,\tau}(x)\frac{1}{T}S\left(\frac{t-x}{T}\right)dx = \lambda_{n,\sigma,\tau}\varphi_{n,\sigma,\tau}(t); \tag{2}$$

- as the eigenfunctions of a differential operator.

$$(\tau^2 - t^2)\frac{d^2\varphi_{n,\sigma,\tau}}{dt^2} - 2t\frac{d\varphi_{n,\sigma,\tau}}{dt} - \sigma^2 t^2 \varphi_{n,\sigma,\tau} = \mu_{n,\sigma,\tau}\varphi_{n,\sigma,\tau}. \tag{3}$$

1.1. *Some Properties of Slepian Functions*

In addition to these basic properties, a number of others are derivable from them (see Slepian, 1983 and references therein). We list a few of them which we use later beginning with another integral equation over the real line

$$\int_{-\infty}^{\infty} \varphi_{n,\sigma,\tau}(x)\frac{1}{T}S\left(\frac{t-x}{T}\right)dx = \varphi_{n,\sigma,\tau}(t). \tag{4}$$

This leads to a dual orthogonality

$$\int_{-\tau}^{\tau} \varphi_{n,\sigma,\tau}(x)\varphi_{m,\sigma,\tau}(x)dx = \lambda_{n,\sigma,\tau}\delta_{nm},$$
$$\int_{-\infty}^{\infty} \varphi_{n,\sigma,\tau}(x)\varphi_{m,\sigma,\tau}(x)dx = \delta_{nm}. \tag{5}$$

In fact, they constitute an orthogonal basis of $L^2(-\tau, \tau)$, and an orthornormal basis of the subspace B_σ of $L^2(-\infty, \infty)$, the Paley–Wiener space of all σ-bandlimited functions. The latter gives rise to two discrete orthogonality conditions of the form

$$T\sum_{n=0}^{\infty} \varphi_{n,\sigma,\tau}(kT)\varphi_{n,\sigma,\tau}(mT) = \delta_{mk},$$
$$T\sum_{k=-\infty}^{\infty} \varphi_{n,\sigma,\tau}(kT)\varphi_{m,\sigma,\tau}(kT) = \delta_{mn}. \tag{6}$$

They lead to a series formula for a σ-bandlimited function (as well as the usual orthogonal series expansion)

$$f(t) = \sum_{n=0}^{\infty} b_n \varphi_{n,\sigma,\tau}(t)$$

where

$$b_n = \sum_{k=-\infty}^{\infty} Tf(kT)\varphi_{n,\sigma,\tau}(kT).$$

The Fourier transform has support on $(-\sigma, \sigma)$ and is given by

$$\hat{\varphi}_{n,\sigma,\tau}(\omega) = (-1)^n \sqrt{\frac{2\pi\tau}{\sigma\lambda_{n,\sigma,\tau}}}\varphi_{n,\sigma,\tau}\left(\frac{\tau\omega}{\sigma}\right)\chi_\sigma(\omega). \tag{7}$$

A Slepian function and its Fourier transform is shown in Figure 1.

These formulae can be used to find the relation between these functions at different scales. By a straightforward change of scale in the integral Eq. (2), we find that

$$\varphi_{n,\sigma\tau,1}(x) = \sqrt{\tau}\varphi_{n,\sigma,\tau}(\tau x). \tag{8}$$

These Slepian functions cannot, however, be used as orthogonal series density estimators on $(-\infty, \infty)$ since they are not a complete basis of $L^2(-\infty, \infty)$. Nonetheless, by using a modification of the usual approach in which the series approximation is preceded by a filtering operation, we will be able to get such an orthogonal series estimator. This will enable us to use some of their unique properties in Sec. 3.

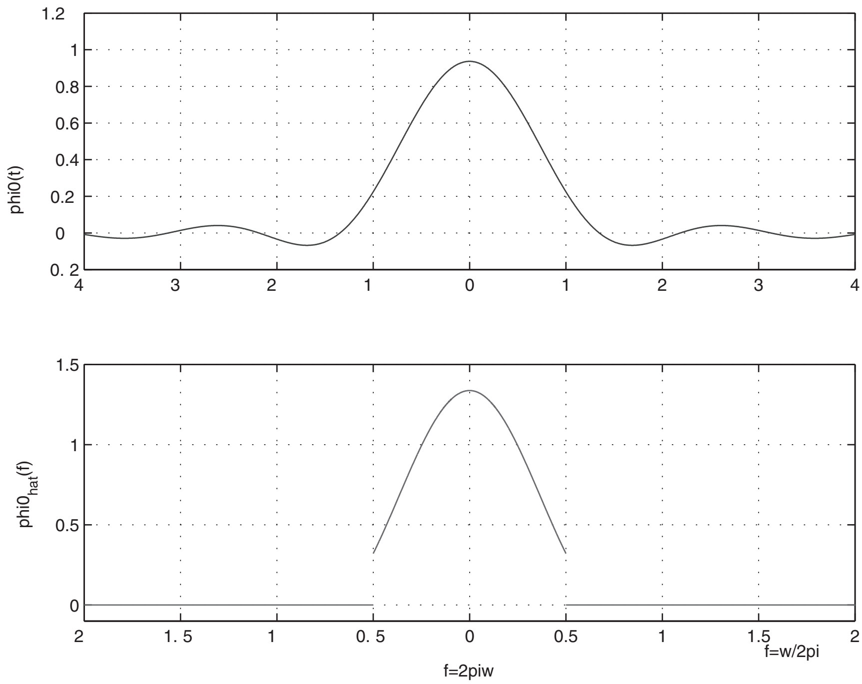

Figure 1. The Slepian function, $\varphi_{0,\pi,1}(x)$ and its Fourier transform $\hat{\varphi}_{0,\pi,1}(\omega)$.

1.2. *A Multiresolution Analysis*

We can also use the expression (8) to get a relation between, say, π and 2π-bandlimited Slepian functions. We can express the former as (Walter and Shen, 2003)

$$\varphi_{n,\pi,\tau}(t) = \sum_{k=0}^{\infty} h_{n,k}\varphi_{k,2\pi,\tau}(t) \tag{9}$$

where

$$h_{n,k} = \frac{1}{2}\sum_{p=-\infty}^{\infty} \varphi_{n,\pi,\tau}(p/2)\varphi_{k,2\pi,\tau}(p/2). \tag{10}$$

This may be considered as one form of a *dilation equation*, which relates the various subspaces in a *multiresolution analysis* (MRA) which appears in wavelet theory. In fact, the Paley–Wiener spaces constitute such an MRA $\{V_m\}$, where $V_m = B_{2^m\pi}$, the space of $2^m\pi$-bandlimited functions.

Proposition 1. *Let $f \in V_0$ with Slepian series in V_0 given by*

$$f(t) = \sum_{k=0}^{\infty} a_k^0 \varphi_{k,\pi,\tau}(t)$$

and with Slepian series in V_1 *given by*

$$f(t) = \sum_{k=0}^{\infty} a_k^1 \varphi_{k,2\pi,\tau}(t);$$

then the coefficients are related by

$$a_n^0 = \sum_{k=0}^{\infty} h_{n,k} a_k^1 \tag{11}$$

where the $h_{n,k}$ *are given by* (10) *and*

$$a_k^1 = \sum_{n=0}^{\infty} h_{n,k} a_n^0. \tag{12}$$

The formulae in this proposition extend to other scales as well, but they only work if our function belongs to the space V_m at the coarsest scale. If we wish to emulate the decomposition and reconstruction algorithms of wavelet theory, we need to find an associated basis of the orthogonal complement of V_0 in V_1. This is an interesting but not too difficult undertaking and will be touched on below.

In the remainder of this subsection we change our notation slightly to avoid so many subscripts. We shall use the notation φ_n^m for the Slepian function with bandwidth $\sigma = 2^m\pi$ and fixed concentration interval $(-\tau, \tau)$

$$\varphi_n^m(t) := \varphi_{n,2^m\pi,\tau}(t). \tag{13}$$

These $\{\varphi_n^m\}_{n=0}^{\infty}$ constitute an orthonormal basis of V_m which in turn have the usual properties of an MRA:

1. $\cdots \subseteq V_m \subseteq V_{m+1} \subseteq \cdots \subseteq L^2(R)$
2. $\overline{\bigcup V_m} = L^2(R)$,
3. $\bigcap V_m = \{0\}$.

The dilation equation relating the various scales is based on (8)

$$\varphi_n^m(t) = \sum_{k=0}^{\infty} h_{n,k}^m \varphi_k^{m+1}(t) \tag{14}$$

where the coefficients are given by Walter and Shen (2002)

$$\begin{aligned} h_{n,k}^m &= \int_{-\infty}^{\infty} \varphi_n^m(t)\varphi_k^{m+1}(t)dt \\ &= \sum_{j=-\infty}^{\infty} 2^{-m-1}\varphi_n^m(2^{-m-1}j)\varphi_k^{m+1}(2^{-m-1}j). \end{aligned} \tag{15}$$

The next step is to find a basis of the space consisting of the orthogonal complement of V_m in V_{m+1}, which we denote by W_m. Let ψ_n^m be given by

$$\begin{aligned} \psi_{2n}^m(t) &:= \left(\cos\frac{3\pi}{2}2^m t\right)\varphi_n^{m-1}(t), \\ \psi_{2n+1}^m(t) &:= \left(\sin\frac{3\pi}{2}2^m t\right)\varphi_n^{m-1}(t), \end{aligned} \tag{16}$$

which, on taking the Fourier transform becomes

$$\hat{\psi}_{2n}^{m}(\omega) = 2^{-1}\left[\hat{\varphi}_{n}^{m-1}\left(\omega - \frac{3\pi}{2}2^{m}\right) + \hat{\varphi}_{n}^{m-1}\left(\omega + \frac{3\pi}{2}2^{m}\right)\right],$$
$$\hat{\psi}_{2n+1}^{m}(\omega) = (2i)^{-1}\left[\hat{\varphi}_{n}^{m-1}\left(\omega - \frac{3\pi}{2}2^{m}\right) - \hat{\varphi}_{n}^{m-1}\left(\omega + \frac{3\pi}{2}2^{m}\right)\right]. \tag{17}$$

The first three Slepian functions and their associated "wavelets" are shown in Figure 2.

Proposition 2. *Let ψ_n^m be given by (16), let $W_m = V_m^{\perp} \cap V_{m+1}$; then $\{\psi_n^m\}_{n=0}^{\infty}$ is an orthonormal basis of W_m and $\{\psi_n^m\}_{n=0}^{\infty}|_{m=-\infty}^{\infty}$ is an orthornormal basis of $L^2(R)$.*

We omit the proof which is based on the convergence properties of the Slepian functions.

Just as in the case of wavelets, we can express any V_m as an orthogonal direct sum

$$V_m = V_k \oplus W_k \oplus \cdots \oplus W_{m-1}, \quad k < m,$$

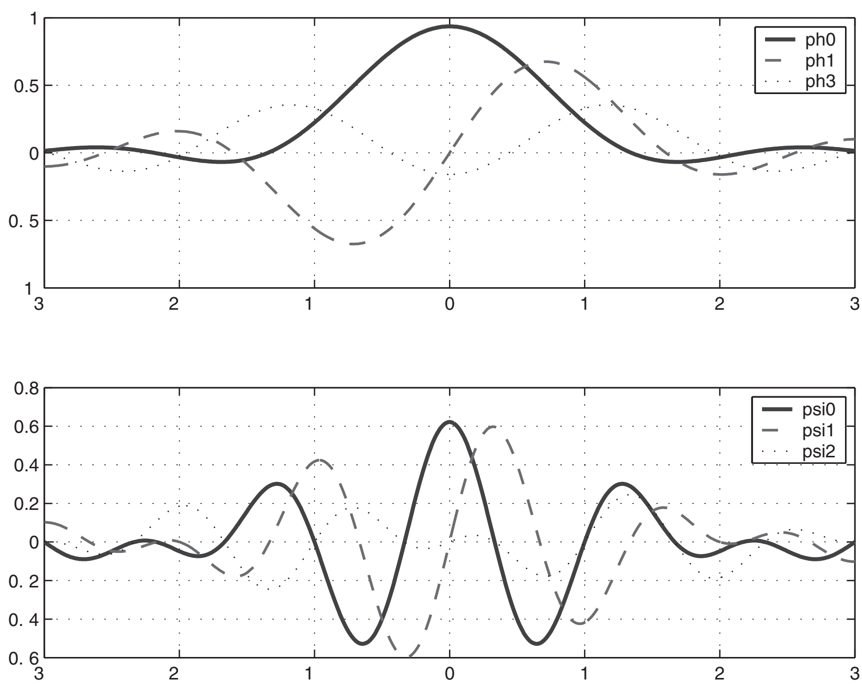

Figure 2. The first three consecutive scaling functions (top) and associated wavelets for PS wavelet (bottom).

and by taking limits

$$\bigoplus_{k=-\infty}^{\infty} W_k = L^2(R).$$

The approximation of a function in $L^2(R)$ by an element of V_m is given by the expansion

$$f_m = \sum_{k=0}^{\infty} \langle f, \varphi_k^m \rangle \varphi_k^m, \tag{18}$$

the projection onto V_m. The projection onto W_m is given by

$$g_m = \sum_{k=0}^{\infty} \langle f, \psi_k^m \rangle \psi_k^m. \tag{19}$$

We can express the approximation in V_m as the sum:

$$f_m = f_k + g_k + g_{k+1} + \cdots + g_{m-1}.$$

This is the same sort of expression encountered in wavelet theory and makes it possible to use nonlinear threshold methods for these Slepian functions as well. The procedure, if it works, would involve ignoring the coefficients $\langle f, \psi_k^m \rangle$ below a certain threshold and then reconstructing each g_m from the remaining coefficients.

2. Prolate Spheroidal Wavelets

Prolate spheroidal wavelets (PS wavelets) were introduced in Walter and Shen (2004) with a scaling function based on the Slepian function $\varphi_{0,\pi,\tau}(t)$, whose bandwidth $\sigma = \pi$, but with any concentration interval $(-\tau, \tau)$. Then the following was shown to hold.

Proposition 3. *Let $\phi(t) = \varphi_{0,\pi,\tau}(t)$ be a π-bandlimited PSWF with concentration interval $(-\tau, \tau)$; then $\{\phi(t-n)\}$ is a Riesz basis of B_π.*

Note that we have changed notation from one form of the Greek letter "phi" to another in going from the Slepian functions to the scaling functions. Later we shall introduce another index to confuse the issue even more. This basis $\{\phi(t-n)\}$ is not orthogonal as are most of the standard scaling function bases, but Riesz bases share many of the properties of orthogonal bases. In fact, they constitute a biorthogonal family with the dual basis whose elements $\tilde{\phi}(t)$ are defined in terms of the Fourier transform as

$$\hat{\tilde{\phi}}(\omega) := \frac{\hat{\phi}(\omega)}{\sum_k |\hat{\phi}(\omega - 2\pi k)|^2}. \tag{20}$$

Since in this case, $\hat{\phi}(\omega)$ has compact support on $[-\pi, \pi]$, and since it is positive there, this series converges and the quotient is non negative as well.

In the construction of wavelets, one usually begins with a scaling function ϕ, whose integer translates are a Riesz basis of a space V_0 which in our case is again the Paley–Wiener space B_π of π-bandlimited functions no matter what the choice of τ. The MRA $\{V_m\}$ is exactly as in the last section, but now we have a different basis. In the usual wavelet theory, the basis of V_m is taken to be $\{\phi(2^m t - n)\}$, which also works for this case. However, we wish to preserve the concentration interval $(-\tau, \tau)$ at all scales and hence will use a different basis for V_m, $\{\varphi_{0,2^m\pi,\tau}(t - n2^{-m})\}$ consisting of translates of a Slepian function with bandwidth $\pi 2^m$ and concentration interval $(-\tau, \tau)$. We shall call the elements of this new basis semi-scaling functions to distinguish them from the standard scaling function which are independent of scale. This still leads to dilation equations whose coefficients change slightly with scale in contrast to the usual case in which they do not.

The mother wavelet is usually given by another dilation equation, but in our case we shall also modify this definition and shall use one related to the maximization problem associated with the Slepian functions.

Definition 1. The Slepian mother semi-wavelet at the scale m is given by

$$\psi_m(t) := \cos(3\pi 2^{m-1} t)\varphi_{0,2^{m-1}\pi,\tau}(t), \tag{21}$$

where the Slepian father semi-wavelet (scaling function) at scale m is denoted by

$$\phi_m(t) := \varphi_{0,2^m\pi,\tau}(t), \tag{22}$$

for $m = 0, \pm 1, \pm 2, \ldots$.

These semi-wavelets $\psi_m(t)$ are orthogonal to each other for different scales, but at a fixed scale their translates give a Riesz basis of their closed linear span. W_m, which is composed of all continuous functions whose Fourier transforms have support in $[-2^{m+1}\pi, -2^m\pi] \cup [2^m\pi, 2^{m+1}\pi]$. Figure 3 shows the first three consecutive scaling functions (top) and associated mother wavelets (bottom) for Slepian semi-wavelets.

These definitions enable us to make the usual wavelet types of approximation, both linear and nonlinear. The importance of these prolate spheroidal wavelets for density estimation arises from their positivity and convergence properties. The scaling function is positive on the interval of concentration $[-\tau, \tau]$ but is small outside this interval. It is an entire function of exponential type and thus is quite smooth. The associated approximation kernel is also locally positive which avoids Gibbs phenomenon, while at the same time leading to a satisfactory rate of convergence of the approximation. Details will be found in Sec. 4.

3. Density Estimation with Slepian Functions

Density estimation with Slepian functions is formally similar to traditional methods of density estimation. However, it has a number of unique properties not shared by other methods. We take, as usual, a sample of a density $f(x)$, $X_1, X_2, \ldots, X_N$ and

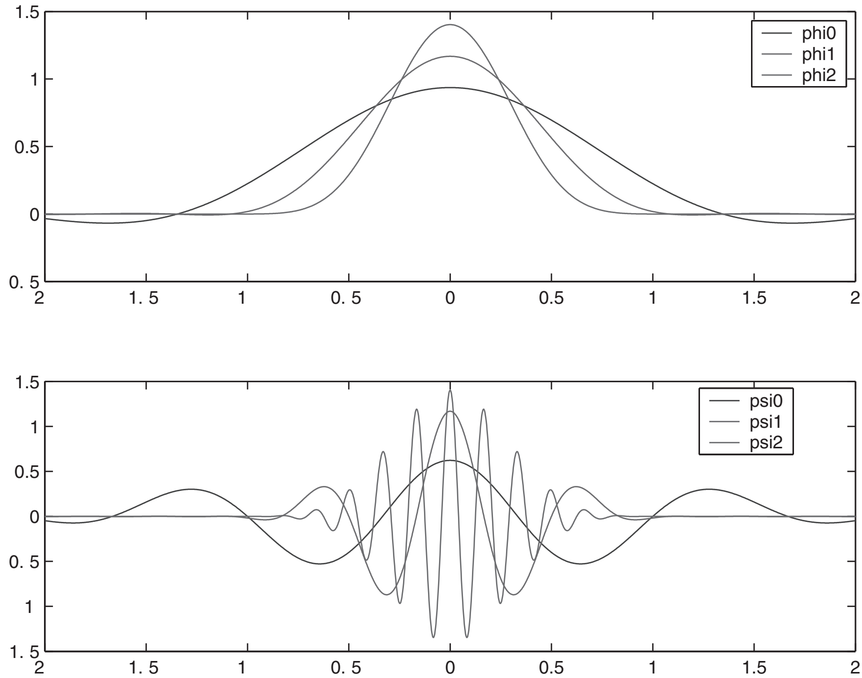

Figure 3. The first three consecutive scaling functions (top) and associated mother wavelets (bottom) for PS semi-wavelets.

use it to try to estimate $f(x)$. Two of the standard procedures are kernel estimators and orthogonal series estimators. Both begin with the empiric distribution given by

$$f^*(x) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - X_i), \tag{23}$$

where δ is the generalized function which has a point mass at 0. This is then smoothed by means of an integral operator to get the estimator

$$\bar{f}(x) := \int k(x, y)f^*(y)dy. \tag{24}$$

If $k(x, y) = K_h(x - y)$, where $K_h(x) = 1/hK(x/h)$ and K is itself a density function, we get the kernel estimator. On the other hand, if $k(x, y) = \sum_{n=0}^{m} \phi_n(x)\phi_n(y)$, where $\{\phi_n\}$ is a complete orthonormal system, then (24) is the standard orthogonal series estimator. The estimator given by the Slepian functions is a combination of these two.

Since the Slepian functions have two parameters, σ, the bandwidth and τ, which gives the concentration interval, we must first choose them. The former will vary with the size of the sample and thus requires only an initial choice. The latter would normally be clear from the type of data or we could choose it from the sample in an obvious way. Rather than allowing σ to take on all real values we shall restrict it to discrete values of the form $\sigma = 2^m\pi$, where m is an integer.

Definition 2. Let $X_1, X_2, \ldots, X_N$ be an i.i.d. sample of a density $f(x)$, $x \in R$; let $\{\varphi_n\}$ be a orthonormal Slepian system with bandwidth parameter σ and concentration parameter τ, $\varphi_n(x) = \varphi_{n,\sigma,\tau}(x)$. The *Slepian estimator* of the density $f(x)$ is given by

$$\bar{f}_{k,\sigma,\tau,N}(x) := \frac{1}{N}\sum_{n=0}^{k}\sum_{i=1}^{N}\varphi_n(x)\varphi_n(X_i). \tag{25}$$

Hidden in this formula is the smoothing kernel estimator of the form (24) which gives us a projection onto the subspace of $L^2(R)$ consisting of σ-bandlimited functions, i.e., the Paley–Wiener space B_σ. The kernel is just the sinc function

$$k(t,x) = \frac{1}{T}S\left(\frac{t-x}{T}\right) = \sum_{n=0}^{\infty}\varphi_n(t)\varphi_n(x), \quad T = \pi/\sigma. \tag{26}$$

The series on the right side of (25) would, in the case of a complete system, converge to the Dirac delta $\delta(t-x)$, but in our case the system is not complete in $L^2(R)$. The estimator in the definition is then obtained by truncating this series, making it not much different than linear wavelet estimators which also involve a projection followed the truncation of a series.

The bias of the estimator (25) is just the error in the expansion of the function f in terms of this orthonormal system $\{\varphi_{n,\sigma,\tau}\}$ plus the error in the projection onto B_σ,

$$E(f^{\#}_{k,\sigma,\tau,N}(x) - f(x)) = \sum_{n=k+1}^{\infty}\varphi_{n,\sigma,\tau}(x)\langle\varphi_{n,\sigma,\tau}, f\rangle + \int_{-\infty}^{\infty}\frac{1}{T}S\left(\frac{t-x}{T}\right)f(t)dt - f(x). \tag{27}$$

If σ now is taken to be $2^m\pi$, then the error in the projection has been shown (Walter and Shen, 2004, Lemma 9), to satisfy

$$\|f - f_m\|_\infty^2 \le \frac{\pi^{-1-2\alpha}2^{m(1-2\alpha)}}{2(2\alpha-1)}\|f\|_\alpha^2 \tag{28}$$

for $f \in H^\alpha$, the Sobolev space for $\alpha > 1/2$ and $f_m(t) = \int 2^m S(2^m(t-x))f(x)dx$. Thus, it remains only to calculate a bound on the tail of the series,

$$\sum_{n=k+1}^{\infty}\varphi_{n,\sigma,\tau}(x)\langle\varphi_{n,\sigma,\tau}, f\rangle.$$

We shall be interested in the behavior of this series when it is restricted to the interval $(-\tau,\tau)$, and therefore consider the square integral of this tail over that interval. It is

$$\begin{aligned}\int_{-\tau}^{\tau}\left|\sum_{n=k+1}^{\infty}\varphi_{n,\sigma,\tau}(x)\langle\varphi_{n,\sigma,\tau}, f\rangle\right|^2 dx &= \sum_{n=k+1}^{\infty}|\langle\varphi_{n,\sigma,\tau}, f\rangle|^2\int_{-\tau}^{\tau}|\varphi_{n,\sigma,\tau}(x)|^2 dx\\ &= \sum_{n=k+1}^{\infty}\lambda_{n,\sigma,\tau}|\langle\varphi_{n,\sigma,\tau}, f\rangle|^2\end{aligned}$$

because of the orthogonality of the $\varphi_{n,\sigma,\tau}$ on the interval $(-\tau, \tau)$. Here the $\lambda_{n,\sigma,\tau}$ are the eigenvalues of the integral operator (2) which are positive but decay very rapidly to 0 monotonically. Thus we obtain the error on the tail to be

$$\int_{-\tau}^{\tau} \left| \sum_{n=k+1}^{\infty} \varphi_{n,\sigma,\tau}(x) \langle \varphi_{n,\sigma,\tau}, f \rangle \right|^2 dx \leq \lambda_{k+1,\sigma,\tau} \|f\|^2. \tag{29}$$

This gives us the following.

Proposition 4. *Let $X_1, X_2, \ldots, X_N$ be an i.i.d. sample of a density $f \in H^\alpha$ for $\alpha > 1/2$, with the support of $f \subset [-\tau, \tau]$; then the asymptotic bias of the estimator $\bar{f}_{k,m}(x)$ given by* (25) *for $\sigma = 2^m \pi$ satisfies*

$$\left\{ \int_{-\tau}^{\tau} |E(\bar{f}_{k,m}(x) - f(x))|^2 dx \right\}^{1/2} \leq \sqrt{\lambda_{k+1}} \|f\| + \sqrt{\frac{\tau \pi^{-1-2\alpha} 2^{m(1-2\alpha)}}{(2\alpha - 1)}} \|f\|_\alpha,$$

for $k = 0, 1, \ldots,$ and $m \in R$.

The first term can be ignored for k large, since λ_k converges to 0 very rapidly; in fact, $\lambda_k < \varepsilon$ for $k > 2\sigma\tau/\pi + (\frac{1}{\pi^2} \log \frac{1-\varepsilon}{\varepsilon} + 10) \log(\sigma\tau)$ (Xiao et al., 2000). Thus, for $\varepsilon = 10^{-10}$ and $\sigma = 2^m \pi$, for example, we need at most $2^{m+1}\tau + m(\frac{10}{\pi^2} \log 10 + 10)$ $\log(2\pi\tau)$ in the series (See Table 1).

The variance expression is also easy to calculate. We find that

$$\begin{aligned} \int_{-\tau}^{\tau} E|(f^{\#}_{k,m}(x))|^2 dx &= \int_{-\tau}^{\tau} E\left| \frac{1}{N} \sum_{n=0}^{k} \sum_{i=1}^{N} \varphi_n(x) \varphi_n(X_i) \right|^2 dx \\ &= \frac{1}{N} \sum_{n=0}^{k} \lambda_n \int_{-\tau}^{\tau} |\varphi_n(t)|^2 f(t) dt \leq \frac{1}{N} \int_{-\tau}^{\tau} \sum_{n=0}^{k} |\varphi_n(t)|^2 f(t) dt \\ &\leq 2^m/N \end{aligned}$$

since the eigenvalues are all less than 1 and the expansion coefficients of the reproducing kernel $2^m S(2^m(t-x))$ are exactly $\varphi_n(t)$. The inequality therefore comes

Table 1
The relation between band width of the Slepian function $\varphi_{0,\sigma,1}(x)$ and the number of its "non zero" eigenvalues

Bandwidth (Ω)	$c = 2\pi\Omega$	No. of $\lambda_k > 10^{-4}$	λ_0
0.15915	1.00000	4	0.57258
0.63662	2.00000	5	0.88056
0.50000	3.14159	6	0.98105
0.63694	4.00000	7	0.99589
1.11408	7.00000	10	0.99999
1.27329	8.00000	11	1.00000
1.59154	10.0000	12	1.00000
3.92699	24.6740	23	1.00000

from Bessel's inequality

$$\sum_{n=0}^{k} |\varphi_n(t)|^2 \leq \int_{-\infty}^{\infty} |2^m S(2^m(t-x))|^2 dx = 2^m.$$

Thus we have a bound on the integrated mean square error,

$$MISE \leq 2^m/N + C_1 + C_2 2^{m(1/2-\alpha)}, \tag{30}$$

where C_1 can be made arbitrarily small by choosing k sufficiently large, and the whole expression can be made small by balancing as usual the first and last terms.

This is not too different than many other estimators. However, there are other properties of the Slepian functions that may confer an advantage. In particular, there is the sampling property in (6). Let us assume that m is taken so large that the sample values are approximately dyadic rational numbers, i.e., each can be expressed as $X_i = 2^{-m}N_i$, where N_i is an integer. If x has same form, then the estimator is approximately

$$\begin{aligned} f^{\#}(x) &= \frac{1}{N} \sum_{n=0}^{k} \sum_{i=1}^{N} \varphi_n(x)\varphi_n(X_i) \\ &= \frac{2^{-m}}{N} \sum_{n=0}^{k} \sum_{i=1}^{N} \varphi_n(2^{-m}p)\varphi_n(2^{-m}N_i) \\ &\approx \frac{1}{N} \sum_{i=1}^{N} \delta_n(p-N_i) = \frac{\#(N_i = p)}{N} \end{aligned}$$

which is exactly the histogram. Hence the estimator can be considered a smoothed histogram (very smoothed) composed of entire functions.

This is particularly useful when applied to the non parametric regression estimator

$$r^{\#}(x) = \frac{\int y f^{\#}(x,y)dy}{f^{\#}(x)}.$$

Thus, the estimator becomes

$$\begin{aligned} r^{\#}(x) &\approx \frac{\sum_{i=1}^{N} \delta(p-N_i)Y_i}{\sum_{i=1}^{N} \delta(p-N_i)} \\ &= \frac{\sum_{N_i=p} Y_i}{\#(N_i = p)}, \quad x = p2^{-m}. \end{aligned}$$

4. Slepian Semi-Wavelets in Density Estimation

The Slepian semi-wavelets with scaling function $\phi_m(x)$ (see (22)) at scale m can be used to define a linear wavelet estimator of a density in the usual way for wavelets in spite of the fact they are not orthogonal. This requires only that a weight be incorporated into the definition to get a proper approximation kernel, which is

$$K_m(x, t) = \sum_n \frac{\phi_m(x - n2^{-m})\phi_m(t - n2^{-m})}{2^m|\hat{\phi}_m(0)|^2}. \tag{31}$$

The linear semi-wavelet estimator is

$$f_m^l(x) := \frac{1}{N}\sum_{k=1}^{N} K_m(x, X_i) \tag{32}$$

which can also be expressed as

$$f_m^l(x) = \sum_n \phi_m(x - n2^{-m}) \sum_{i=1}^{N} \frac{1}{N} \frac{\phi_m(X_i - n2^{-m})}{2^m|\hat{\phi}_m(0)|^2},$$

and can be shown to be integrated mean square consistent as $m \to \infty$.

That it is asymptotically unbiased follows from a result in Walter and Shen (2004), where it was shown that the series

$$f_m^s(x) = \sum_n \frac{\phi_m(x - n2^{-m})f(n2^{-m})}{2^m\hat{\phi}_m(0)}$$

converges to $f(x)$ in the sense of L^2 provided that $f \in H^p$, a Sobolev space, for $p > 2$. Hence we need only show that $Ef_m^l(x) - f_m^s(x)$ converges to 0 in the sense of L^2 to show the estimator is asymptotically unbiased. This difference is

$$\begin{aligned}
&Ef_m^l(x) - f_m^s(x)\\
&= \sum_n \frac{\phi_m(x - n2^{-m})}{2^m\hat{\phi}_m(0)}\left\{\int \frac{f(t)\phi_m(t - n2^{-m})}{\hat{\phi}_m(0)}dt - f(n2^{-m})\right\}\\
&= \sum_n \frac{\phi_m(x - n2^{-m})}{2^m\hat{\phi}_m(0)}\left\{\frac{1}{2\pi}\int \frac{\hat{f}(\omega)\hat{\phi}_m(\omega)e^{in2^{-m}\omega}}{\hat{\phi}_m(0)} - \hat{f}(\omega)e^{in2^{-m}\omega}d\omega\right\}\\
&= \sum_n \frac{\phi_m(x - n2^{-m})}{2^m\hat{\phi}_m(0)}\left\{\frac{1}{2\pi}\int_{-2^m\pi}^{2^m\pi} \hat{f}(\omega)e^{in2^{-m}\omega}\left(\frac{\hat{\phi}_m(\omega)}{\hat{\phi}_m(0)} - 1\right)d\omega\right.\\
&\qquad\left. - \frac{1}{2\pi}\left(\int_{2^m\pi}^{\infty} + \int_{-\infty}^{-2^m\pi}\right)\hat{f}(\omega)e^{in2^{-m}\omega}d\omega\right\}.
\end{aligned} \tag{33}$$

Thus we need an estimate of $\frac{\hat{\phi}_m(\omega)}{\hat{\phi}_m(0)} - 1$ on the interval $[-2^m\pi, 2^m\pi]$. But we have from the formula for the Fourier transform that $\frac{\hat{\phi}_m(\omega)}{\hat{\phi}_m(0)} = \frac{\hat{\varphi}_{0,2^m\pi.\tau}(\omega)}{\hat{\varphi}_{0,2^m\pi.\tau}(0)} = \frac{\varphi_{0,2^m\pi.\tau}(\tau\omega/2^m\pi)}{\varphi_{0,2^m\pi.\tau}(0)}$ on this interval and since the Slepian function is entire, it can be differentiated infinitely often. The first derivative at 0 is zero while the second satisfies, by (3) $\tau^2\varphi''_{0,2^m\pi.\tau}(0) = \mu_0\varphi_{0,2^m\pi.\tau}(0)$. Therefore we have

$$\frac{\hat{\phi}_m(\omega)}{\hat{\phi}_m(0)} - 1 = O((\omega/2^m\pi)^2) \quad \text{as } m \to \infty.$$

Thus the first integral in (33) may be expressed as

$$\begin{aligned}&\left|\int_{-2^m\pi}^{2^m\pi} \hat{f}(\omega)e^{in2^{-m}\omega}\left(\frac{\hat{\phi}_m(\omega)}{\hat{\phi}_m(0)} - 1\right)d\omega\right|^2 \\ &\leq 2^{-4m}C\int_{-2^m\pi}^{2^m\pi} |\hat{f}(\omega)|^2(\omega^2+1)^p d\omega \int_{-2^m\pi}^{2^m\pi}\left(\frac{(\omega)^4}{(\omega^2+1)^p}\right)d\omega \\ &\leq C'2^{-3m}\end{aligned}$$

since $p > 2$. The last integral in (33) also satisfies the same sort of inequality. This gives us the convergence to 0 as $m \to \infty$ of (33).

The variance can similarly be shown to be dominated by a constant multiple of $2^m\ /N$, which enables us to get the desired mean squared consistency.

Proposition 5. *Let $f \in H^p$ for $p > 2$; let f_m^l be the estimator given by* (32). *Then the IMSE of f_m^l converges to* 0 *as $m \to \infty$.*

This is estimator is very similar to the usual kernel estimators used in density estimation except for the change of scale. We consider some other properties of this kernel $K_m(x, t)$ given by (31). We can convert the semi-scaling function series into a series involving the standard Slepian functions and then use their properties to study this kernel.

Lemma 1. *The kernel $K_m(x, t)$ given by* (31) *is a positive definite convolution kernel satisfying $K_m(x, t) = K_m(x - t, 0) > 0$ for $|x - t| \leq 1/2^{m+1}$.*

Since, by the definition and by an application of (8)

$$\phi_m(x) = \varphi_{0,2^m\pi,\tau}(x) = 2^{m/2}\varphi_{0,\pi,2^m\tau}(2^m x),$$

it follows that

$$K_m(x, t) = \sum_n \frac{\varphi_{0,\pi,2^m\tau}(2^m x - n)\varphi_{0,\pi,2^m\tau}(2^m t - n)}{2^{-m}|\hat{\varphi}_{0,\pi,2^m,\tau}(0)|^2}$$

But since the Fourier transform of $\varphi_{0,\pi,2^m\tau}$ has support on $[-\pi, \pi]$, by the Fourier integral theorem we have

$$\varphi_{0,\pi,2^m\tau}(2^m x - n) = \frac{1}{2\pi}\int \hat{\varphi}_{0,\pi,2^m\tau}(\omega)e^{i\omega(2^m x - n)}d\omega,$$

i.e., $\varphi_{0,\pi,2^m\tau}(2^m x - n)$ is the Fourier series coefficient of $f(x,\omega) = \hat{\varphi}_{0,\pi,2^m\tau}(\omega)e^{i\omega 2^m x}$ and hence by Parseval's formula,

$$\begin{aligned} K_m(x,t) &= \sum_n \frac{\varphi_{0,\pi,2^m\tau}(2^m x - n)\varphi_{0,\pi,2^m\tau}(2^m t - n)}{2^{-m}|\hat{\varphi}_{0,\pi,2^m\tau}(0)|^2} \\ &= \frac{2^m}{2\pi|\hat{\varphi}_{0,\pi,2^m\tau}(0)|^2}\int_{-\pi}^{\pi} f(x,\omega)\overline{f(t,\omega)}d\omega \\ &= \frac{2^m}{2\pi|\hat{\varphi}_{0,\pi,2^m\tau}(0)|^2}\int_{-\pi}^{\pi} |\hat{\varphi}_{0,\pi,2^m\tau}(\omega)|^2 e^{i\omega 2^m(x-t)}d\omega \\ &= K_m(x-t,0). \end{aligned} \tag{34}$$

Since $|\hat{\varphi}_{0,\pi,2^m\tau}(\omega)|^2$ is an even function, it follows that the exponential function can be replace by the cosine from which the second conclusion follows. The fact that it is positive definite follows from Bochner's theorem and the fact that the Fourier transform is non-negative.

The kernel is actually positive on a larger interval than that given by the proposition. Indeed, Figure 4 shows the kernel on the square $[-4,4]\times[-4,4]$, while Figure 5 shows the kernel as one variable function at level $m = 0$. The kernel is computed by using a truncated version of (34):

$$K(x,0) \approx \sum_{n=0}^{5} \varphi_{0,\pi,1}(n)\varphi_{0,\pi,1}(x+n)) + \sum_{n=1}^{5} \varphi_{0,\pi,1}(n)(\varphi_{0,\pi,1}(x-n).$$

Some function values for the Slepian function $\varphi_{0,\pi,1}$at integers are listed in Table 2, which shows the truncation for (34) is reasonable. From Figure 5 we observe that the first zero crossing for $\varphi_{0,\pi,1}$ occurs at $x = 1.34$, while for the kernel it occurs at $x = 3.09$ approximately. Table 3 gives some extreme values for the (approximated) kernel.

Figure 6 shows the comparison among the Slepian semi-wavelet kernel, and the other three popular positive kernels—the Epanechnikov kernel and the Biweight kernel (top) and the Gaussian kernel (middle). The bottom panel gives a detailed looking outside of the concentration interval. The Epanechnikov kernel and

Table 2
Slepian function $\varphi_{0,\pi,1}$ at integers

n	$\varphi_{0,\pi,1}(n)$	n	$\varphi_{0,\pi,1}(n)$
0	0.936577	6	−0.003456
1	0.224680	7	0.002533
2	−0.033684	8	−0.001936
3	0.014224	9	0.001528
4	−0.007868	10	−0.001237
5	0.004997		

Table 3
Minimum of the kernel

x	$min\{K(x,0)\}$	sign
[−3.6, 3.6]	−0.0145	−
[−3.5, 3.5]	−0.0142	−
[−3.4, 3.4]	−0.0125	−
[−3.3, 3.3]	−0.0094	−
[−3.2, 3.2]	−0.0052	−
[−3.1, 3.1]	−2.9689e-4	−
[−3.0, 3.0]	0.0125	+

Biweight kernel are given by

$$K_e(t) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}t^2), & |t| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases}$$

and

$$K_b(t) = \begin{cases} \frac{15}{16}(1 - t^2)^2, & |t| < 1 \\ 0 & \text{otherwise} \end{cases}$$

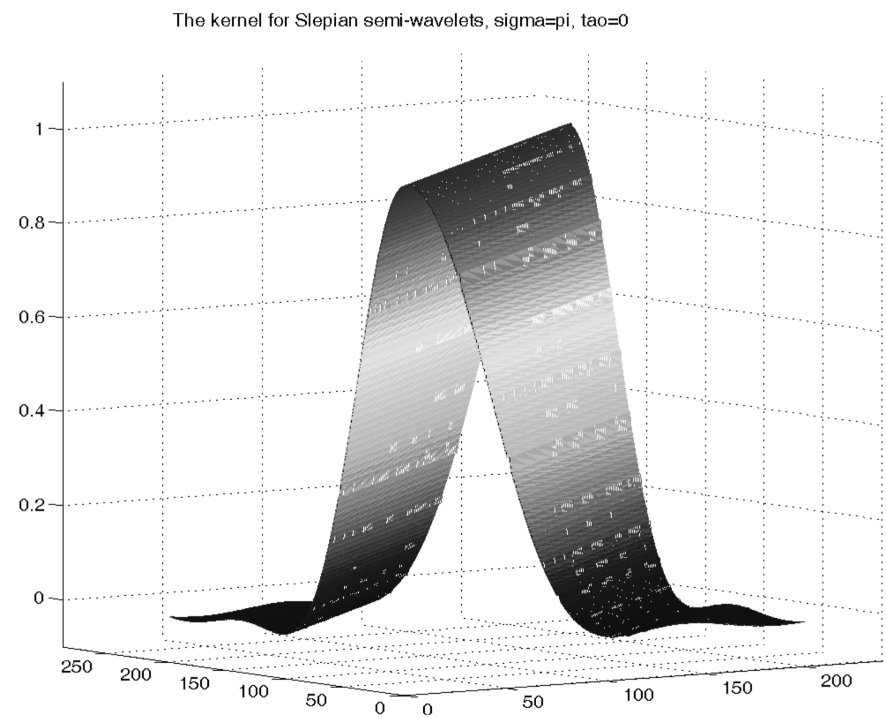

Figure 4. The kernel of the Slepian semi-wavelets at level $m = 0, \sigma = \pi, \tau = 1$.

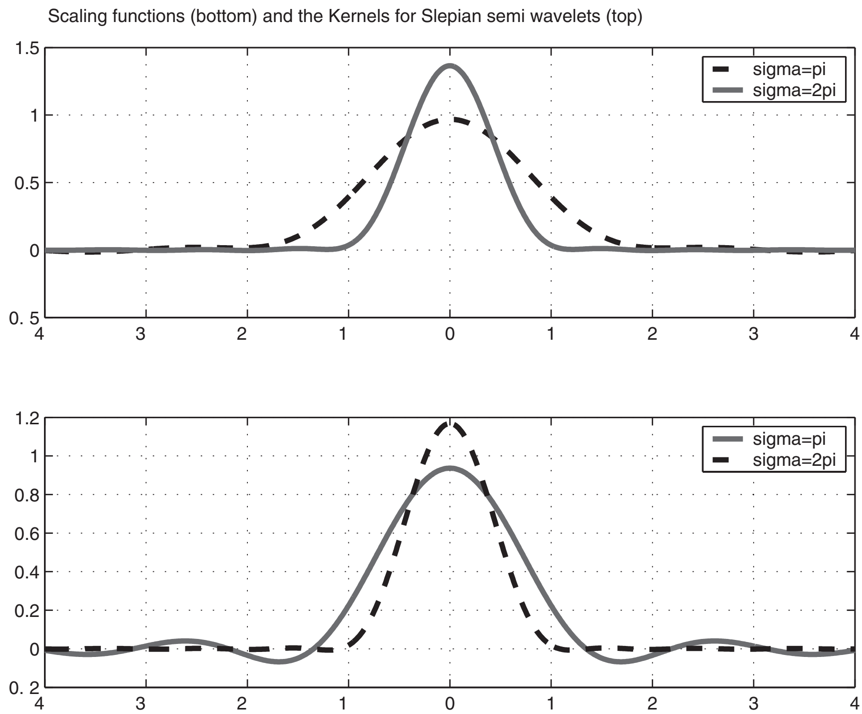

Figure 5. The Slepian semi-wavelet kernel (top) and associated scaling functions (bottom), as 1-d functions on the interval $[-4, 4]$, $\sigma = \pi$, $\tau = 1$.

respectively. Notice that all three are positive kernels and the first two kernels are compactly supported.

In order to study the behavior of $\varphi_0 = \varphi_{0,\pi,\tau}$ outside of the interval $[-\tau, \tau]$, we first use the fact that the L^2 norm is given by

$$\int_{|t|>\tau} |\varphi_0(t)|^2 dt = \int_{-\infty}^{\infty} |\varphi_0(t)|^2 dt - \int_{-\tau}^{\tau} |\varphi_0(t)|^2 dt$$
$$= 1 - \lambda_0^2 = \varepsilon.$$

As we have seen, ε is quite small provided τ is relatively large. For $\pi\tau$ equal to 2, $\varepsilon = 0.0006$, and is much smaller for larger values. Since φ_0 vanishes at infinity, we have, for $x \geq \tau$,

$$\varphi_0^2(x) = -2\int_x^{\infty} \varphi_0(t)\varphi_0'(t)dt \leq 2\left\{\int_x^{\infty} \varphi_0^2(t)dt\right\}^{1/2}\left\{\int_x^{\infty} \varphi_0'^2(t)dt\right\}^{1/2}$$

by Schwarz's inequality. The first integral on the right is dominated by ε, while the second integral satisfies

$$\int_x^{\infty} \varphi_0'^2(t)dt \leq \int_{-\infty}^{\infty} \varphi_0'^2(t)dt = \frac{1}{2\pi}\int_{\pi}^{\pi} |i\omega\hat{\varphi}_0(\omega)|^2 d\omega$$
$$\leq \frac{\pi}{2}\int_{\pi}^{\pi} |\hat{\varphi}_0(\omega)|^2 d\omega = \pi^2 \int_{-\infty}^{\infty} \varphi_0^2(t)dt = \pi^2.$$

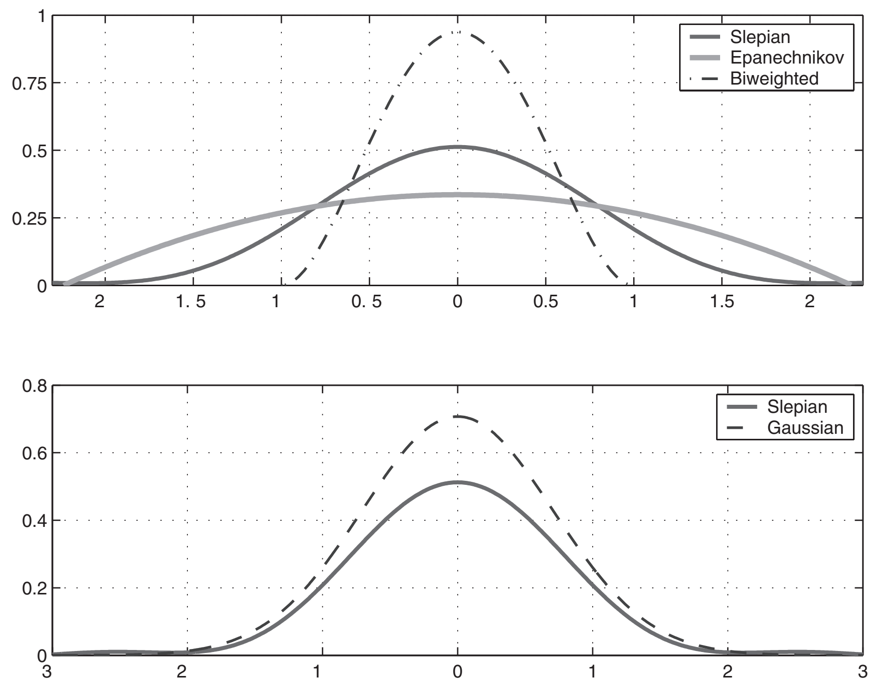

Figure 6. The Slepian semi-wavelet kernel compared to the Epanechnikov kernel and the Biweight kernel (top), and to the Gaussian kernel (middle). The figure in the bottom shows the detailed look of the two kernels in the middle panel.

By combining these two inequalities we obtain the following:

Lemma 2. *Let* $|x| \geq \tau$, *then* $|\varphi_{0,\pi,\tau}(x)|^2 \leq 2\pi\sqrt{\varepsilon}$. $|\varphi_{0,2^m\pi,\tau}(x)|^2 \leq 2\sqrt{(\varepsilon/2)2^m\pi} = \sqrt{\varepsilon\pi 2^{m+1}}$.

Thus we have shown that not only is the energy outside of the interval $[\tau, \tau]$ small, the function $\varphi_{0,\pi,\tau}$ is also uniformly small. This gives further justification for truncating the series used to define $K_m(x, t)$ in (31). Notice that the shape of the kernel is very similar to that of the Gaussian kernel. The difference is that it is an entire function of exponential type and decays like $1/x$ as $x \to \infty$. Thus it can be used to better estimate densities with longer tails. But since it is negligible outside of a finite interval, it is also suitable for densities with compact support.

The chief advantage however, is in the fact that it is a wavelet estimator and hence one can use thresholding methods to remove noise and to get an efficient representation.

5. Examples

In this section, we illustrate how the Slepian semi-wavelet kernel density estimators are used to estimate the density functions of univariate data. The results are compared to those for standard wavelet estimators.

Table 4
Eruption lengths (in minutes) of 107 eruptions of Old Faithful Geyser

1.73	4.08	4.25	1.83	3.72	4.37	3.92	2.33	3.33	4.18	4.50
4.60	4.62	4.50	3.43	1.85	4.70	3.20	4.57	3.73	4.58	1.80
4.08	2.25	4.20	4.93	3.78	1.68	4.58	3.58	1.67	3.50	3.70
3.83	3.43	4.42	3.50	4.13	1.75	3.50	3.70	4.63	4.62	2.50
3.73	4.00	4.65	3.77	1.90	4.35	3.80	4.25	1.83	4.03	2.27
4.83	4.40	4.12	1.83	4.73	1.77	3.80	3.58	2.03	1.97	2.93
4.33	4.10	1.88	4.50	1.85	4.25	1.80	3.67	2.72	4.60	4.63
4.43	3.10	1.82	4.00	2.00	4.10	1.95	1.90	4.03	4.00	4.00
3.92	4.00	4.07	3.52	3.95	4.05	1.77	4.13	1.73	3.75	1.97
2.93	3.93	4.33	1.67	3.68	1.90	4.28	4.53			

We consider three examples:

Example 1. Density estimate of the Old Faithful geyser data. The data for the eruption lengths (in minutes) of 107 eruptions of Old Faithful geyser can be found in Table 4. Both this data set and the data set in next example can be found in Silverman (1986, p. 8). Figure 7 shows the data set and its histogram estimate.

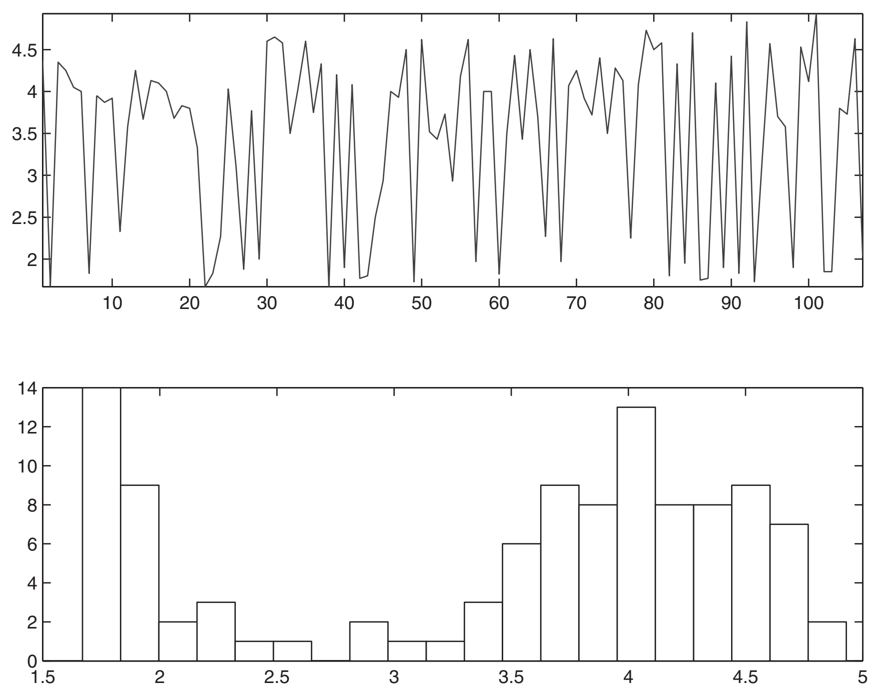

Figure 7. The eruption length (in minutes) of the Old Faithful geyser at Yellow Stone National Park (top) and its histogram estimator (bottom).

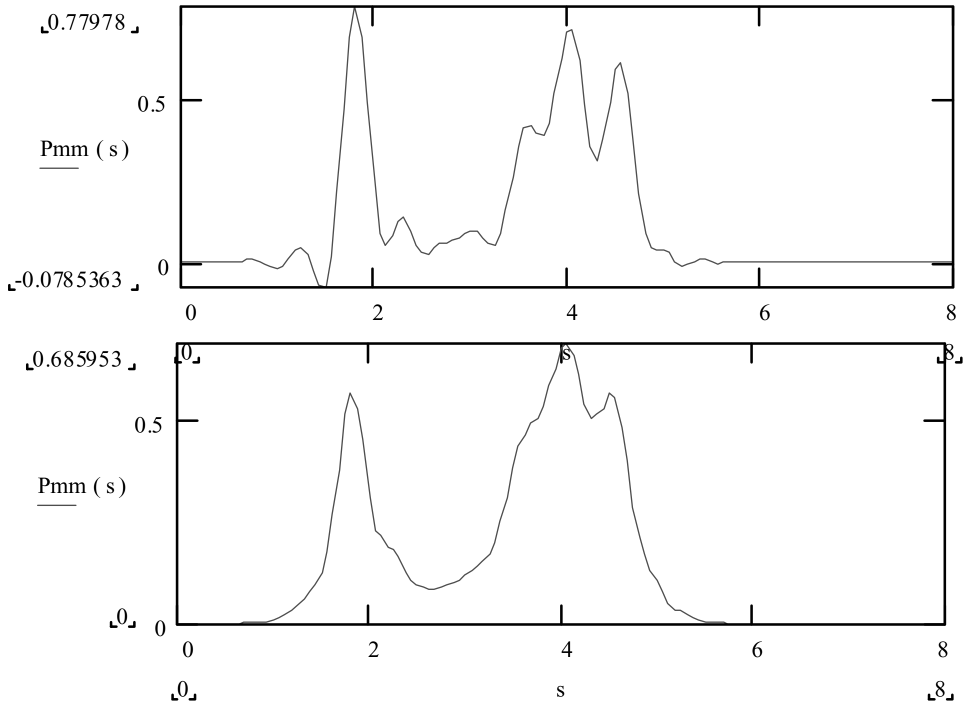

Figure 8. Density estimate for the Old Faithful geyser data using estimators associated with reproducing kernels for ${}_4\varphi$ (top), and using ${}_8\varphi$ (bottom), window width 0.25 ($m = 2$).

Figure 8 shows two estimates associated with Daubechies 4 and 8 wavelets (Walter and Shen, 1999). Figure 9 shows the Slepian semi-wavelet estimates. The left column is for the window width $= 0.5$ ($m = 1$) and the right column is for the window width $= 0.25$ ($m = 2$). The two rows correspond to different values of τ. We should point out in Figure 9 that the estimators may take negative values, these are very small; indeed, the minimum values for the density estimates are shown in the Table 5. By adjusting the parameter τ in the density estimator, the negative values can be reduced significantly (compared to Figure 8).

Example 2. Density estimate for suicide study data. The data for the lengths of treatment spells (in days) of 86 control patients in suicide study can be found in Table 6. Figure 10 shows the data set and its histogram estimator. The Slepian semi-wavelet estimator with different window width $m = 0.25$ (left column) and 0.5 (right column) and concentration $\tau = 1$ (top row) and $\tau = 2$ (bottom row) are shown in Figure 11. The results show that the Slepian semi-wavelet estimators can pick up the tail very well.

Example 3. Density estimate for INTEL CORP (NasdaqNM: INTC) stock closing price (close price adjusted for dividends and splits). While this is not necessarily

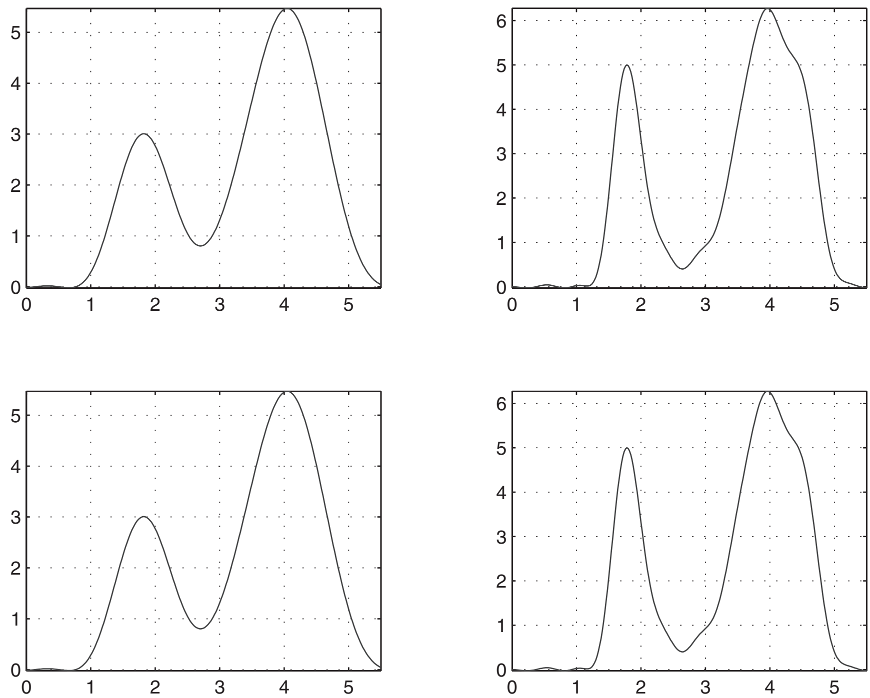

Figure 9. Density estimate for the Old Faithful geyser data by using the Slepian semi wavelet kernel, window width 0.5 (left column) and window width 0.25 (right column).

IID data, it is reasonably close (as can be attested to by anyone who has tried to predict its price). Figure 12 shows the data set (top) and its histogram (bottom). The Slepian semi-wavelet estimator with different window width $m = 0.25$ (left column) and 0.5 (right column) and concentration $\tau = 1$ (top row) and $\tau = 2$ (bottom row) are shown in Figure 13. The data source is from http://finance.yahoo.com/.

6. Conclusions

In this work we have introduced two types of density estimators, one an orthogonal series estimator based on Slepian functions and the other a wavelet type estimator based on Slepian wavelets. The former converges rapidly to the

Table 5
The minimum values of the estimator

Window Width	0.5	0.25
$\tau = 1$	−.199156572e-1	−.251416005e-1
$\tau = 2$	−.883816073e-5	−.199120488e-5

Table 6
Lengths of treatment spells (in days) of 86 control patients in suicide study

1	14	76	17	18	32	54	82	103	153	311
1	93	49	31	21	34	56	83	111	163	314
1	144	31	49	21	35	56	84	112	167	322
5	30	93	79	22	36	62	84	119	175	369
7	40	14	103	25	37	63	84	122	228	415
8	75	737	147	27	38	65	90	123	231	573
8	740	256	257	27	39	65	91	126	235	609
13	242	134	256	30	39	67	92	129		

density whenever it is sufficiently smooth, but may take negative values and have excessive oscillations arising from Gibbs phenomenon. The latter may not converge as rapidly, but has only negligible negative values and avoids Gibbs phenomenon. Both methods require fewer computations than do other methods because of the unique properties of the Slepian functions. The estimators based on them are entire functions of exponential type, but nonetheless do have a strong localization property.

If we compare these methods to traditional methods we find that all continuous orthogonal series methods, including wavelet methods, give rise to Gibbs phenomenon and may have negative values and thus share the shortcomings of the first estimator without its computational advantages. The kernel estimators with positive kernels do avoid the excessive oscillations and negativity, but are not wavelet estimators and thus cannot take advantage of their data reduction and thresholding properties.

The examples show that the negative values can occur for the standard wavelet kernel density estimators, while the values for the Slepian semi-wavelet estimators can be adapted to be practically non negative. The latter are also seem to be smoother than the former. The results of the suicide study (in Figures 10 and 11) also show that the Slepian semi-wavelet estimator can pick up densities with long tails. This is contrast to other to estimators which sometimes "eat" the tails of long-tailed random variables.

Table 7
Summary statistics for intel closing price

Min	13.16
1st Qu	17.37
Mean	23.17
Median	19.56
3rd Qu	29.94
Max	35.55
Total N	252
Std dev.	6.95

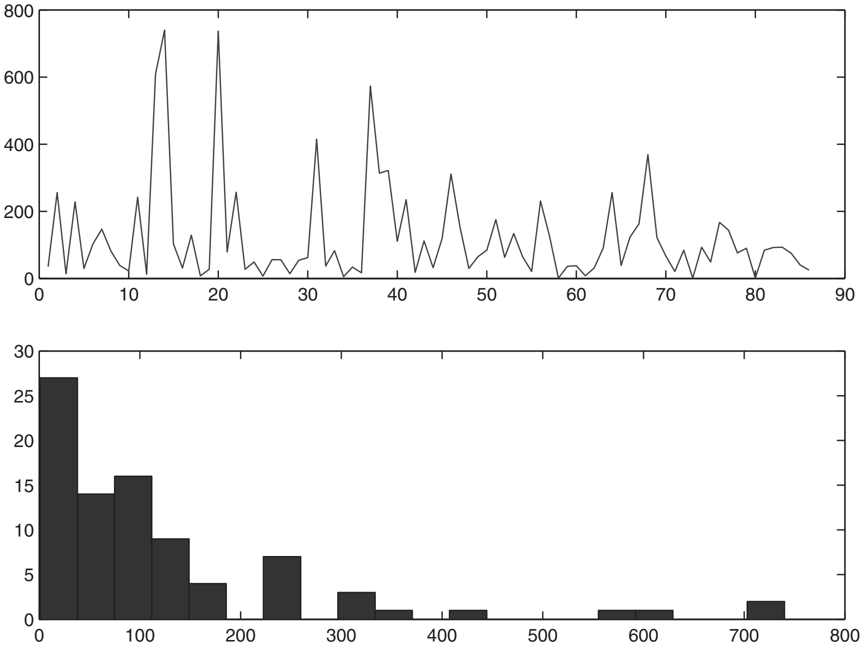

Figure 10. The suicide study data (in days) (top) and its histogram estimator (bottom).

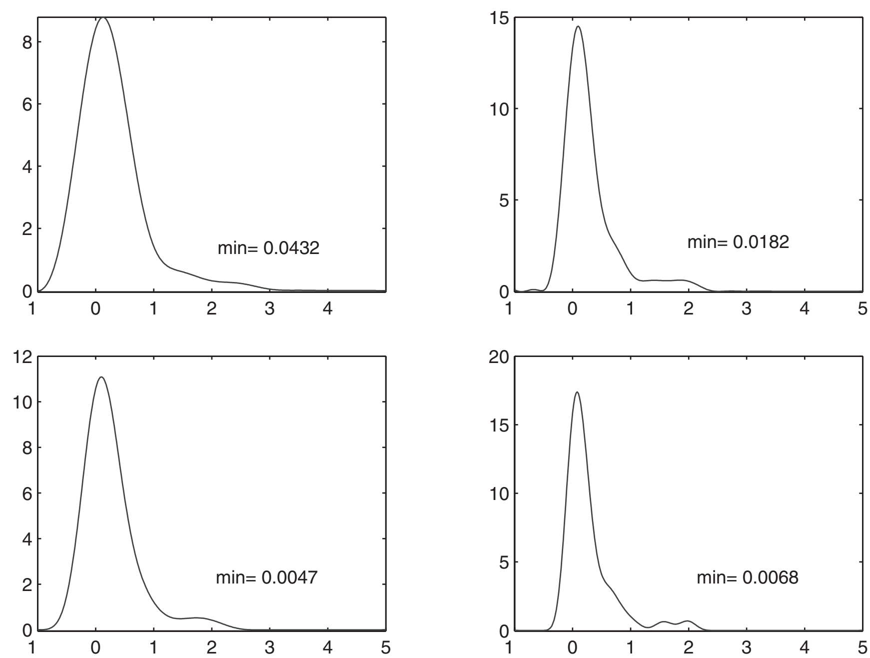

Figure 11. Density estimate for the suicide study data by using the Slepian semi wavelet kernel, window width 0.5, and window width 0.25.

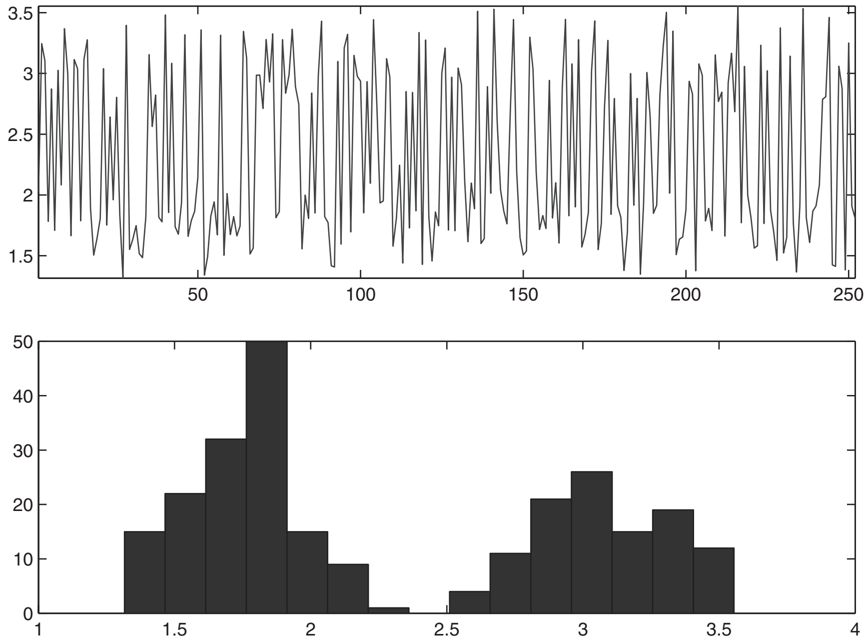

Figure 12. The intel stock closing price (top) and its histogram estimator (bottom).

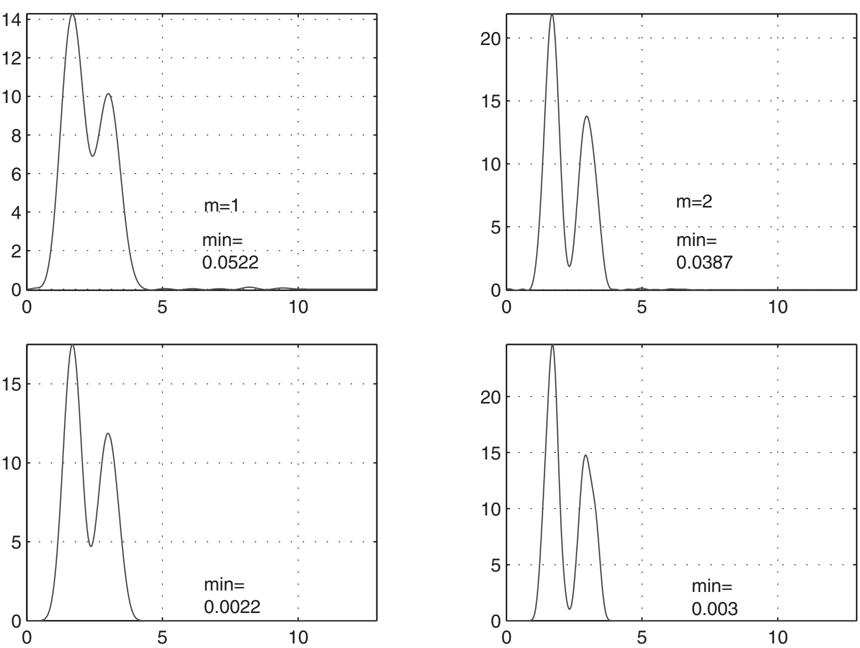

Figure 13. Density estimate for the data set of intel stock closing price by using the Slepian semi wavelet kernel, window width 0.5 (left column), and window width 0.25 (right column).

References

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia: SIAM.

Djokovic, I., Vaidyanathan, P. P. (1997). Generalized sampling theorems in multiresolution subspaces. *IEEE Trans. Signal Proc.* 45(3):583–599.

Higgins, J. R. (1996). *Sampling Series in Fourier Analysis and Signal Theory*. Oxford: Clarendon Press.

Papoulis, A. (1977). *Signal Analysis*. New York: McGraw-Hill.

Shannon, C. E. (1949). Communication in the presence of noise. *Proc. IRE* 37:10–21.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London, UK: Chapman and Hall.

Slepian, D. (1983). Some comments on Fourier analysis, uncertainty, and modeling. *SIAM Rev.* 25:379–393.

Vaidyanathan, P. P. (2001). Sampling theorems for non-bandlimited signals: theoretical impact and practical applications. *Proc SAMPTA 2001* 17–26.

Walter, G. G. (1992a). A sampling theorem for wavelet subspaces. *IEEE Trans. Inform. Theor.* 38:881–884.

Walter, G. G. (1992b). Differential operators which commute with characteristic functions with applications to a lucky accident. *Complex Variables* 18:7–12.

Walter, G. G., Shen, X. (2002a). Positive sampling in wavelet subspaces. *Appl. Comp. Harmonic Anal.* 12(1):150–165.

Walter, G. G., Shen, X. (2004). Wavelets based on prolate spheroidal wave functions. *Journal of Fourier Analysis and Applications* 10(1):1–26.

Walter, G. G., Shen, X. (2003). Sampling with prolate spheroidal functions. *J. Sampling Theor. Signal Image Proc.* 2(1):25–52.

Walter, G. G., Shen, X. (2000). *Wavelets and Other Orthogonal Systems*. 2nd ed. Boca Raton, FL: CRC Press.

Walter, G. G., Shen, X. (1999). Continuous non-negative wavelets and their use in density estimation. *Commun. Statist.* 28:1–18.

Xiao, H, Rokhlin, V., Yarvin, N. (2000). Prolate spheroidal wavefunctions, quadrature and interpolation. Special issue to celebrate Pierre Sabatier's 65th birthday (Montpellier). *Inverse Probl.* 17(4):805–838.

Zayed, A. (1993). *Advances in Shannon's Sampling Theory*. Boca Raton, FL: CRC Press.