An Introduction to Wavelet Analysis

1.FOURIER ANALYSIS

Math.View

- $L^2(0, 2\pi)$: all measurable functions f:

$$\int_0^{2\pi} |f(x)|^2 dx < \infty$$

Extend these functions periodically to $\mathbf{R}$: $L^2(0, 2\pi)$ becomes vector space of all functions, period 2π, square integrable.

$$f \in L^2(0, 2\pi) \qquad \Rightarrow \qquad f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} , \tag{1}$$

where c_n are the Fourier coefficients

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx . \tag{2}$$

Convergence in (1) is in $L^2(0, 2\pi)$.

- Two characteristics of (1):

(i) f is decomposed as an infinite sum of orthogonal components

$$g_n(x) = c_n e^{inx} \quad .$$

(ii) the orthonormal basis $\{w_n(x) = e^{inx}\}$ of $L^2(0, 2\pi)$ is generated by **dilation** of the single wave

$$w(x) = e^{ix} \ , \tag{3}$$

that is, $w_n(x) = w(nx)$, for all integers n .

FACT: *Every periodic function, of period 2π, square integrable, is generated by a superposition of integral dilations of the basic function $w(x) = e^{ix}$.*

2. WAVELETS

- Wavelet Uses

  signal processing, medical imaging

  pattern recognition, data compression

  numerical and data analysis

- Tools

  wavelet transform

  multiresolution analysis

  time-scale and time-frequency analysis

  best basis analysis

  matching pursuit decompositions

• $L^2(R)$: all measurable functions f:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty .$$

Clearly, $L^2(0, 2\pi)$ and $L^2(R)$ are quite different. For the later, functions have to decay to zero at infinity, **so sinusoids do not belong to it.**

We look for "waves" that generate $L^2(R)$. These have to decay to zero **fast**. That is, we look for **wavelets** to generate $L^2(R)$.

As in $L^2(0, 2\pi)$ prefer that a **single** function, ψ , generates $L^2(R)$. How to cover R, if ψ has to have fast decay?

Obvious thing : Translate ψ along R!

• Let $Z = \{0, \pm 1, \ldots\}$. The simplest way for ψ to cover R is to consider integral translations of ψ , i.e, $\psi(x-k), k \in Z$.

Consider waves with frequencies partioned in frequency bands. For computational efficiency, user power of 2 for the partitions. Consider the wavelets

$$\psi(2^j x - k), j, k \in Z \ .$$

This is obtained from $\psi(x)$ by a **binary dilation** 2^j and a **dyadic translation** $k/2^j$.

Consider an orthonormal basis generated by ψ :

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), j, k \in Z \ . \tag{1}$$

If ψ has unit length, $\psi_{jk}(.)$ also has.

• **Definition:** *A function* $\psi \in L^2(R)$ *is an orthogonal wavelet if the family defined by (1) is an orthonormal basis for* $L^2(R)$ *, i.e.*

$$< \psi_{j,k}, \psi_{\ell,m} >= \delta_{j\ell}\delta_{km} \ , \ j,k,\ell,m \in Z$$

and any $f \in L^2(R)$ *can be written as*

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(x) \ , \qquad (2)$$

where convergence in (2) is in $L^2(R)$ *.*

The representation (2) is a **wavelet series** and the **wavelet coefficients** are given by

$$c_{j,k} =< f, \psi_{j,k} >= \int_{-\infty}^{\infty} f(x)\overline{\psi_{j,k}(x)}dx \ . \qquad (3)$$

Examples

- Example 2.1: Haar

$$\psi^{(H)}(t) = \begin{cases} 1, & 0 \le t < 1/2 \\ -1, & 1/2 \le t < 1 \\ 0, & \text{otherwise}, \end{cases}$$

$$\psi^{(H)}_{j,k}(t) = \begin{cases} 2^{j/2}, & t \in [2^{-j}k, 2^{-j}(k+\frac{1}{2})) \\ -2^{j/2}, & t \in [2^{-j}(k+\frac{1}{2}), 2^{-j}(k+1)) \\ 0, & \text{otherwise}. \end{cases}$$

- Example 2.2: Daublets, Symmlets, Coiflets

- Example 2.3.

  Morlet(Modulated Gaussian):

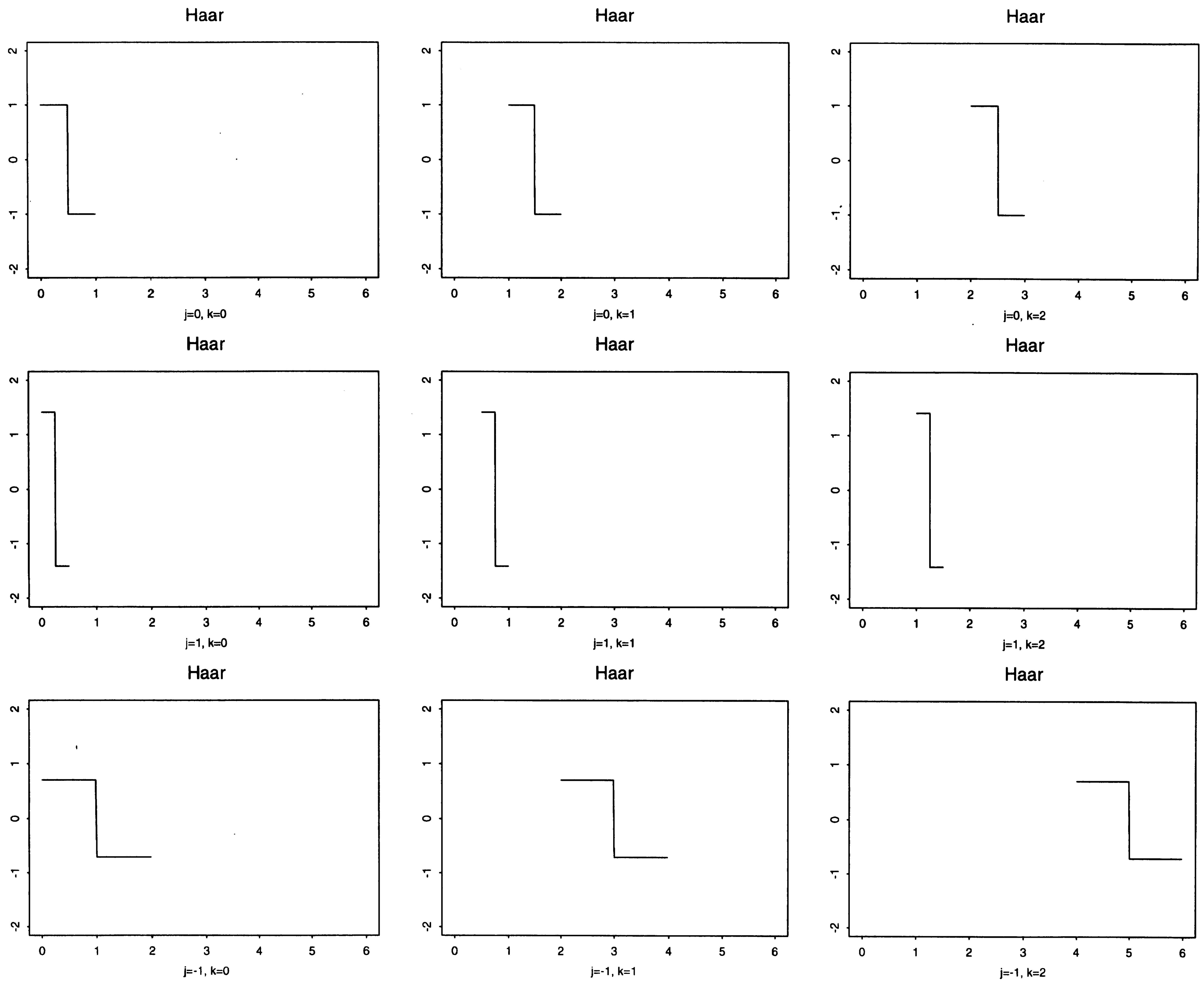
$$\psi(t) = e^{i\omega_0 t} e^{-t^2/2}.$$

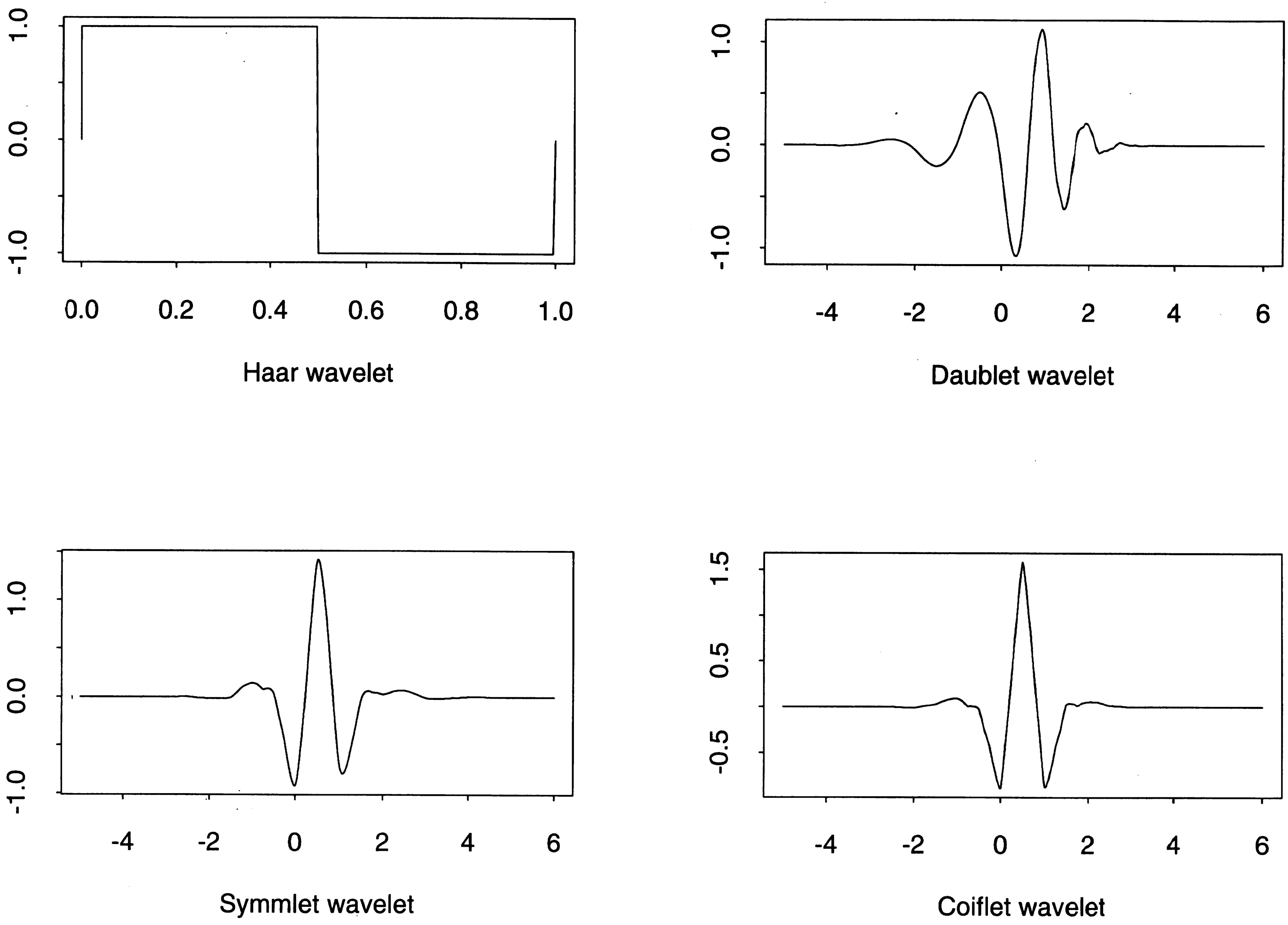
  Mexican hat:

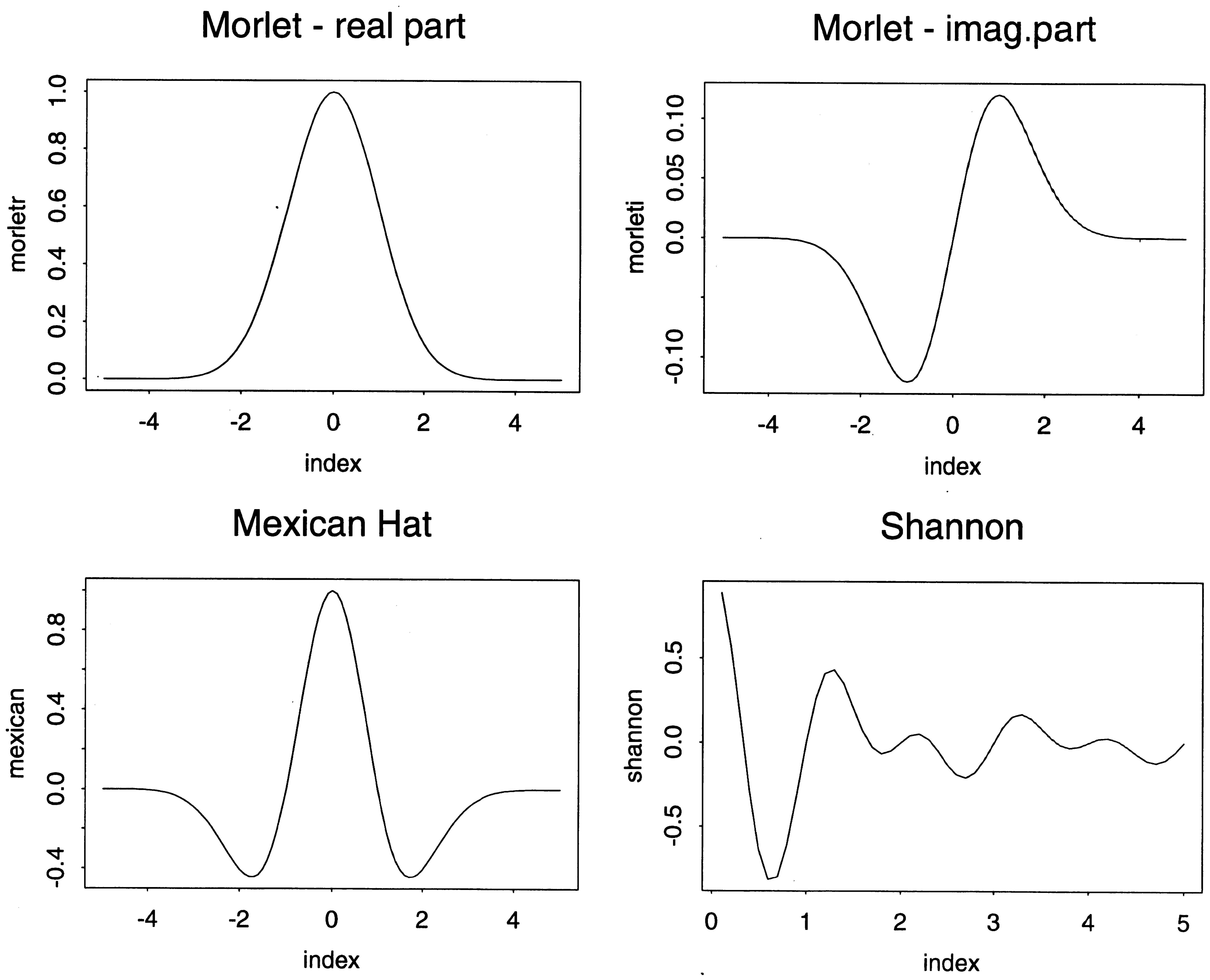
$$\psi(t) = (1 - t^2) e^{-t^2/2}.$$

  Shannon:

$$\psi(t) = \frac{\sin(\frac{\pi t}{2})}{\frac{\pi t}{2}} \cos(\frac{3\pi t}{2}).$$


Haar
j=0, k=0
Haar
j=0, k=1
Haar
j=0, k=2
Haar
j=1, k=0
Haar
j=1, k=1
Haar
j=1, k=2
Haar
j=-1, k=0
Haar
j=-1, k=1
Haar
j=-1, k=2


1.0
0.0
-1.0
0.0 0.2 0.4 0.6 0.8 1.0
Haar wavelet
1.0
0.0
-1.0
-4 -2 0 2 4 6
Daublet wavelet
1.0
0.0
-1.0
-4 -2 0 2 4 6
Symmlet wavelet
1.5
0.5
-0.5
-4 -2 0 2 4 6
Coiflet wavelet


Morlet - real part
morletr
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
index
Morlet - imag.part
morleti
-0.10
0.0
0.05
0.10
-4
-2
0
2
4
index
Mexican Hat
mexican
-0.4
0.0
0.4
0.8
-4
-2
0
2
4
index
Shannon
shannon
-0.5
0.0
0.5
0
1
2
3
4
5
index

- Scaling Function

$$\phi(t) = \sqrt{2}\sum_k \ell_k \phi(2t-k) \qquad (1)$$

Generates orthonormal family of $L^2(\Re)$,

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k), \; j,k \in Z$$

Mother wavelet can be obtained as

$$\psi(t) = \sqrt{2}\sum_k h_k \phi(2t-k), \qquad (2)$$

$$h_k = (-1)^k \ell_{1-k} \qquad (3)$$

(1) and (2): **dilation equations**

(3): **quadrature mirror filter relation**

3.THE WAVELET TRANSFORM

Discrete wavelet transform(DWT):

$f = (f1, f2, \cdots, fT)'$: discrete signal with $T = 2^M$, $M > 0$ integer.

The DWT maps f to a vector of wavelet coefficients

$$w = (s_J, d_J, d_{J-1}, \cdots, d_1)', \; J \leq M$$

where

$$s_J = (s_{J,1}, s_{J,2}, \cdots, s_{J,T/2^J})'$$

$$d_J = (d_{J,1}, d_{J,2}, \cdots, d_{J,T/2^J})'$$

$$d_{J-1} = (d_{J-1,1}, d_{J-1,2}, \cdots, d_{J-1,T/2^{(J-1)}})'$$

$$\vdots$$

$$d_1 = (d_{1,1}, d_{1,2}, \cdots, d_{1,T/2})'$$

$$s_{J,k} = \sum_{t=1}^{T} f_t \phi_{J,k}(t), \; k = 1, \cdots, \frac{T}{2^J}$$

$$d_{j,k} = \sum_{t=1}^{T} f_t \psi_{j,k}(t), \; k = 1, \cdots, \frac{T}{2^j}$$

f_t can be obtained by

$$f_t = \sum_k s_{J,k} \phi_{J,k}(t) + \sum_{j=1}^{J} \sum_k d_{j,k} \psi_{j,k}(t)$$

4. MULTIRESOLUTION ANALYSIS

NONPARAMETRIC ESTIMATION WITH WAVELETS

. one of the great success stories of wavelets is in the field of nonparametric statistical estimation.

. wavelet shrinkage: removing noise by shrinking wavelet coefficients towards zero.

. The model:

$$y_i = f_i + \epsilon_i, \quad i = 1, 2, \cdots, T$$

where $\epsilon_i \backsim iidN(0, \sigma^2)$

1. Choice of Threshold

(a) Choice of scheme

(i) Hard Threshlod

$$\delta_\lambda^H(x) = \begin{cases} 0, & \text{if } |x| \le \lambda \\ x, & \text{if } |x| > \lambda. \end{cases}$$

(ii) Soft Threshold

$$\delta_\lambda^S(x) = \begin{cases} 0, & \text{if } |x| \le \lambda \\ sign(x)(|x| - \lambda), & \text{if } |x| > \lambda. \end{cases}$$

- Shrinkage procedure

  [1] Take the discrete wavelet transform of the data $y_1, \ldots, y_T$, leading to the T wavelet coefficients $y_{j,k}$, which are contaminated by noise.

  [2] Use thresholds to reduce the coefficients, making null those coefficients below a certain value. Several choices here are possible and we will discuss some of them in the next section. We obtain, in this stage, the coefficients without noise.

  [3] Take the inverse wavelet transform of the coefficients in stage [2] to get the estimates $\widehat{f}_i$.

2. Choice of Parameters

(a) Universal

(b) SureShrink

(c) Cross-Validation

(d) Ogden and Parzen(1996)

$$S_J(t) = \sum_k s_{J,k}\,\phi_{J,k}(t) \quad \text{smooth signal}$$

$$D_j(t) = \sum_k d_{j,k}\,\psi_{j,k}(t) \quad \text{detail signal}$$

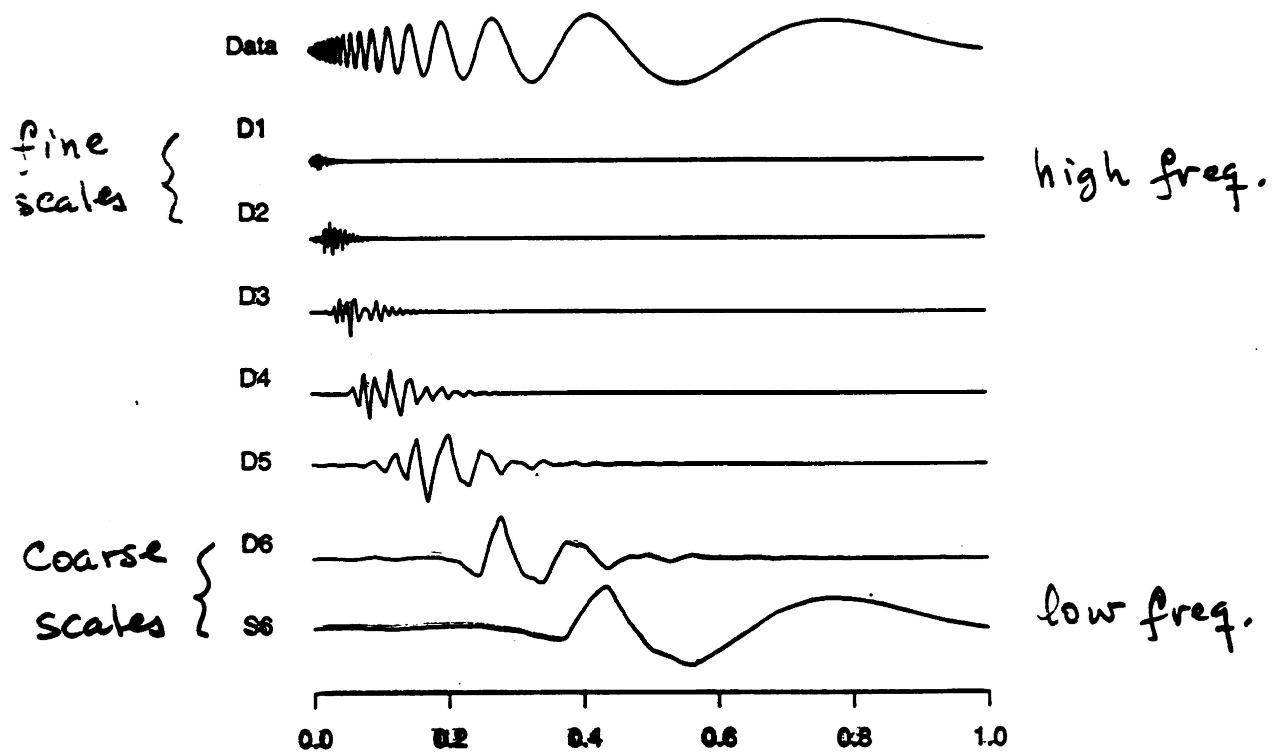
$$f(t) \simeq S_J(t) + D_J(t) + D_{J-1}(t) + \cdots + D_1(t)$$

MULTIRESOLUTION DECOMPOSITION

(orthog. comp. at diff. scales)

Mallat (1989)

Meyer (1986)



Multiresolution decomposition of the doppler signal.

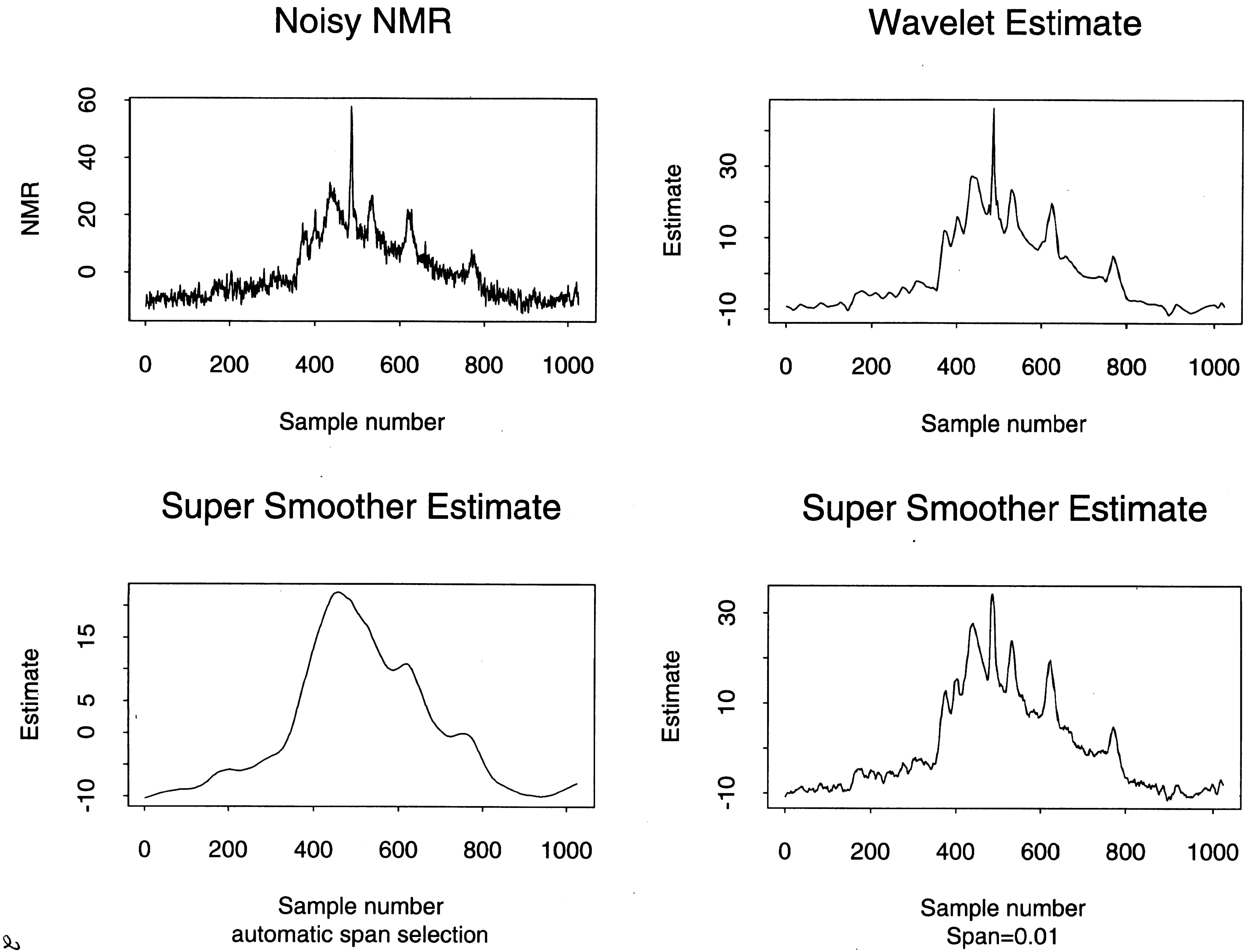
S+ WAVELETS

. Theoretical Properties of Wavelet Shrinkage:

for certain choices of the λ, the estimate $\hat{f}_i$ can almost achieve the *minimax risk* over a broad class of functions $\mathcal{F}$:

$$R(\hat{f}_i, f) \approx inf_{\hat{f}} sup_{f \in \mathcal{F}} R(\hat{f}, f)$$

Wavelet shrinkage gives nearly the best possible estimate of f_i making a minimum of assumptions about the underlying nature of $\mathcal{F}$.


Noisy NMR
NMR
60
40
20
0
0
200
400
600
800
1000
Sample number
Wavelet Estimate
Estimate
30
10
-10
0
200
400
600
800
1000
Sample number
Super Smoother Estimate
Estimate
15
5
0
-10
0
200
400
600
800
1000
Sample number
automatic span selection
Super Smoother Estimate
Estimate
30
10
-10
0
200
400
600
800
1000
Sample number
Span=0.01

Software

- Wavethresh

  Nason(1993)

  StatLib, ftp

  S-PLUS, UNIX version

- S+ WAVELETS

  Bruce and Gao(1994)

  UNIX and WINDOWS

- WaveLab

  Buckheit et al.(1995)

  Macintosh, UNIX, WINDOWS