International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 4, July Year 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Computational Statistics on Wavelet Methods

S MD Riyaz Ahamad
Department of Mathematics, College of Global Engineering and Technology, India
Email: riyazroq@gmail.com

----------**********----------

Abstract:

Wavelet methods have exhibited significant effectiveness in function estimation by applying term-by-term thresholding to the empirical wavelet coefficients. Nevertheless, it has been demonstrated that aggregating the empirical wavelet coefficients into blocks and making simultaneous threshold decisions for all coefficients within each block presents several advantages over term-by-term wavelet thresholding, including asymptotic optimality and better mean squared error performance infinite sample situations. An empirical Bayes approach is employed to integrate information regarding neighboring empirical wavelet coefficients into function estimation, leading to the development of block wavelet shrinkage and block wavelet thresholding estimators. Simulated examples are utilized to demonstrate the performance of these estimators and to compare them with various existing non-Bayesian block wavelet thresholding estimators. The results indicate that the proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators significantly surpass the non-Bayesian counterparts in finite sample scenarios. Additionally, an application to a data set gathered from an anaesthesiological study is also presented.

Keywords — ***Empirical Bayes, Block thresholding, Maximum likelihood estimation, Non-parametric regression; and Wavelet transform***.

----------**********----------

I. INTRODUCTION

Over the past ten years, the literature on non-parametric regression has been predominantly influenced by non-linear wavelet methods. These techniques are founded on the concept of thresholding, which generally involves the assessment of each empirical wavelet coefficient individually. When a coefficient is sufficiently large in magnitude, meaning it surpasses a predetermined threshold, the associated term in the empirical wavelet expansion is either retained (or diminished towards zero); otherwise, it is excluded. The resulting estimators are usually executed via fast algorithms, rendering them particularly attractive in practical applications. Following the groundbreaking works of Donoho and Johnstone (1994) and Donoho et al. (1995), a variety of alternative wavelet methods have emerged. This assortment encompasses Bayesian term-by-term wavelet shrinkage and wavelet thresholding estimators. (To clarify terminology, a shrinkage rule reduces empirical wavelet coefficients towards zero, while a thresholding rule not only reduces but also nullifies all empirical wavelet coefficients that fall below a specified level.) These estimators have demonstrated effectiveness and have been considered to be less arbitrary than their classical counterparts. Comprehensive reviews and descriptions of the various classical and Bayesian term-by-term wavelet schemes can be found in works such as those by Ogden (1997) and Vidakovic (1999), as well as in papers published in the edited volume by Mueller and Vidakovic (1999), and in the review articles by Antoniadis (1997), Vidakovic (1998b), and Abramovich et al. (2000). Hall et al. (1997), Hall et al. (1998, 1999), Cai (1999), and Cai and Silverman (2001), among others, introduced an alternative method based on block thresholding. In this methodology, the empirical wavelet coefficients are thresholded in blocks rather than on an individual basis. The

resultant estimators have demonstrated several advantages over term-by-term wavelet thresholding estimators, including asymptotic optimality and improved mean squared error performance in finite sample contexts. The objective of this paper is to expand upon the benefits of block thresholding. I propose an empirical Bayes approach that integrates information from neighboring coefficients into wavelet function estimation, leading to block wavelet shrinkage and block wavelet thresholding estimators. Simulated examples are utilized to demonstrate the efficacy of the resulting estimators and to compare these estimators against several existing non-Bayesian block wavelet thresholding estimators. The proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators have been demonstrated to surpass the performance of non-Bayesian block wavelet thresholding estimators in finite sample circumstances. The organization of the paper is as follows. In Section 2, the wavelet paradigm is introduced within the framework of classical non-parametric regression. In Section 3, the prior model applied to the wavelet coefficients of the unknown response function is presented, leading to the development of posterior-based block wavelet shrinkage and block wavelet thresholding estimators. An empirical Bayes methodology for estimating the hyperparameters of the prior model is also discussed. Section 4 includes simulated examples to illustrate the proposed estimators, which are then compared with various existing non-Bayesian block wavelet thresholding estimators. Furthermore, an application to a dataset derived from an anaesthesiological study is presented. Section 5 concludes with final remarks.

II. WAVELET REGRESSION

regard as the classical non-parametric regression setting

$$y_i = g(t_i) + \sigma\varepsilon_i, \quad i = 1,\ldots,n, \tag{1}$$

where $t_i = i/n$, $n = 2^J$ for some positive integer J, ε_i are independent and identically distributed N(0; 1) random variables and the noise level σ may, or might not, be identified. The problem is to estimation the fundamental function g from the interpretation $\mathbf{y} = (y_1,\ldots,y_n)'$ without assuming any particular parametric structure on its form. The wavelet approach to this problem is simply described. Given a appropriate wavelet basis and primary resolution level $j_0 \geq 0$, the discrete wavelet transform (DWT) of y gives go up to an n-dimensional vector $\hat{\mathbf{d}}$ consisting of what are acknowledged as the empirical scaling coefficients $\hat{c}_{j_0k}$ $(k = 0,\ldots,2^{j_0}-1)$ and the empirical wavelet coefficients $\hat{d}_{jki}$ $(j = j_0,\ldots,J-1; k = 0,1,\ldots,2^j-1)$. In follow, the DWT (and its inverse (IDWT)) may be perform through a computationally fast algorithm developed by Mallat (1989) that require only O(n) operations. Due to the orthogonality of the DWT, it follow from (1) that

$$\hat{c}_{j_0k} = c_{j_0k} + \sigma\varepsilon_{j_0k}, \quad k = 0,1,\ldots,2^{j_0}-1, \tag{2}$$

$$\hat{d}_{jk} = d_{jk} + \sigma\varepsilon_{jk}, \quad j = j_0,\ldots,J-1, \ k = 0,1,\ldots,2^j-1, \tag{3}$$

where the ε_{jk} are themselves independent and identically distributed N(0; 1) random variables, and the c_{j_0k} and d_{jk} are, correspondingly, the right scaling and wavelet coefficients of the (unknown) vector of function values $\mathbf{g} = (g(t_1),\ldots,g(t_n))'$. The challenge is now to estimate the true wavelet coefficients dμ from the empirical ones, highlighting the sparseness within the wavelet expansion dμ. The critical insight here is that most significant information about the underlying function, g, dμ is concentrated in a few substantial wavelet coefficients, whereas smaller coefficients are likely due to noise contamination affecting all coefficients uniformly. Thus, a reliable estimation of g could be achieved by accurately identifying and preserving the largest wavelet coefficients while disregarding the smaller ones. This concept is implemented through term-by-term wavelet thresholding, where each empirical wavelet coefficient dμ is assessed against a predefined threshold and either retained or reduced towards zero accordingly.

If the magnitude exceeds a specified threshold, the item is retained; otherwise, it is disregarded. This method seeks a balance between the variance and bias components affecting the mean squared error, albeit the balance achieved is not optimal, leading to biased estimators due to the excessive removal of terms from the empirical wavelet expansion, thereby resulting in sub-optimal L2-risk rates, as documented by Hall et al. (1998, 1999). Enhancing estimation precision can be accomplished by incorporating information from adjacent empirical wavelet coefficients. Specifically, empirical wavelet coefficients at each resolution level j can be thresholded in groups rather than individually, thereby increasing the data available for accurately estimating the average empirical wavelet coefficient within a block. and creation a decision. Considering whether to retain or discard it represents a significantly more complex decision than applying a term-by-term threshold rule, thereby enhancing the accuracy of threshold determinations (refer to Hall et al., 1997; Hall et al., 1998, 1999; Cai, 1999; and Cai and Silverman, 2001). In the subsequent section, we introduce an empirical Bayes method to integrate information from neighboring coefficients into wavelet function estimation, culminating in the development of level-dependent block wavelet shrinkage and block wavelet thresholding estimators.

III. THE PRIOR MODEL

Evaluate the sequence model as outlined in equations (2) and (3). A notable feature of Bayesian wavelet shrinkage and thresholding estimation strategies is the application of diverse, largely independent prior distributions on the wavelet coefficients, as documented extensively in the literature by researchers such as Chipman et al. (1997), Abramovich et al. (1998), Clyde et al. (1998), Johnstone and Silverman (1998), Vidakovic (1998a), Abramovich and Sapatinas (1999), Clyde and George, 1999, ffi000). It is reasonable to expect that a model allowing for correlations between neighboring wavelet coefficients would be more parsimonious. Accordingly, we suggest placing priors on blocks of wavelet coefficients rather than on each one individually. More precisely, at each resolution level $j\ (j=j_0,\ldots,J-1)$, the wavelet coefficients d_{jk} are grouped into m_j non-overlapping blocks, $b_{jK}\ (K=1,\ldots,m_j)$, of length lj . (The choice of lj will be discussed in Section 4.) For each of the block b_{jK}, Assumed the following prior model:

$$b_{jK}|\gamma_{jK} \sim \mathrm{N}(0,\gamma_{jK}V_j), \tag{4}$$

$$\gamma_{jK} \sim \text{Bernoulli}\,(\pi_j), \tag{5}$$

independently for $K=1,\ldots,m_j$, where $0\leqslant\pi_j\leqslant 1$. i also assumed block independence across the different resolution levels $j\ (j=j_0,\ldots,J-1)$. according to this prior model, a block is either zero with probability $1-\pi_j$ or, with probability π_j, multivariate normally distributed with mean zero and variance–covariance matrix V_j. Thus, our prior model allows for the fact that a wavelet coefficient is more likely to contain signal if neighboring wavelet coefficients do also. To complete the prior formulation, I assumed vague priors for the scaling coefficients $c_{j0k}\ (k=0,1,\ldots,2^{j_0}-1)$ and in particular we took $c_{j0k}\sim \mathrm{N}(0,\varepsilon),\ \varepsilon\to\infty$. Through the covariance matrix V_j, we may allow for various forms of dependency between the wavelet coefficients within each block. Since the correlation between two wavelet coefficients generally weakens as the distance between them increases, we choose V_j to be the $l_j\times l_j$ matrix with elements

$$V_j[k,l]=\tau_j^2\rho_j^{|k-l|} \quad \text{where } |\rho_j|<1,\ k,l=1,\ldots,l_j. \tag{6}$$

A similar choice of covariance structure has been employed by Vidakovic and MEuller (1995) and Vannucci and Corradi (1999a,b) in a diDerent context. The above model can be regard as an lean-to of the prior mixture model of a univariate normal distribution and a point mass at zero (i.e. $l_j=1$), used in Bayesian term-by-term wavelet function estimation by, for example, Clyde et al. (1998), Abramovich et al. (1998), Johnstone and Silverman (1998), Abramovich and Sapatinas (1999) and Clyde and George (1999, 2000).

3.2. Posterior-based block shrinkage and block thresholding estimators: At each resolution level

$j\ (j=j_0,\ldots,J-1)$. I now considered the corresponding non-overlapping blocks $\hat{b}_{jK}\ (K=1,\ldots,m_j)$ of the empirical wavelet coefficients $\hat{d}_{jk}$. Combining the prior model (4) and (5) with the likelihood from $\hat{b}_{jK}\sim \mathrm{N}(b_{jK},\sigma^2 I)$, and averaging over all possible γ_{jK}, yields the marginal posterior distribution

$$b_{jK}|\hat{b}_{jK}\sim\frac{1}{1+O_{jK}}N(A_j\hat{b}_{jK},\sigma^2A_j)+\frac{O_{jK}}{1+O_{jK}}\delta(\mathbf{0}), \qquad (7)$$

where $\delta(\mathbf{0})$ is a vector of point masses at zero, $A_j=(\sigma^2V_j^{-1}+I)^{-1}$ and O_{jK} is the posterior odds ratio that $\gamma_{jK}=0$ versus $\gamma_{jK}=1$, given by

$$O_{jK}=\frac{1-\pi_j}{\pi_j}\left(\frac{\det(V_j)}{\sigma^{2j}\det(A_j)}\right)^{1/2}\exp\left\{-\frac{\hat{b}'_{jK}A_j\hat{b}_{jK}}{2\sigma^2}\right\}. \qquad (8)$$

For the jKth block, de5ne the vector $\hat{d}^{\star}_{jK}=A_j\hat{b}_{jK}$ and its elements $\hat{d}^{\star}_{jk}$. Applying different losses, one can generate various block shrinkage and block thresholding estimators. In exacting, we have:

• Posterior mean (L2-loss): it is straightforward to explain that the posterior mean of b_{jK} is given by

$$\tilde{b}_{jK}=E(b_{jK}\,|\,\hat{b}_{jK})=\frac{1}{1+O_{jK}}\hat{d}^{\star}_{jK}. \qquad (9)$$

This is a block (non-linear) shrinkage rule wherever each empirical wavelet coefficient in a block is shrunk by the similar shrinkage factor depending on all coefficients within the block.

• Marginal posterior median (L1-loss): in this case, used for the posterior distribution in (7), the marginal posterior distribution of d_{jk} is given by $d_{jk}\,|\,\hat{b}_{jK}\sim\frac{1}{1+O_{jK}}N(\hat{d}^{\star}_{jk},\sigma^2A_j[j,j])+\frac{O_{jK}}{1+O_{jK}}\delta(0)$, where $\delta(0)$ is a point mass at zero and $A_j[j,j]$ is the entry that appear on the diagonal of Aj (they are the same for all k). Hence, following the point of view of Abramovich et al. (1998) and Abramovich and Sapatinas (1999) in the Bayesian term-by-term wavelet function inference framework, the marginal posterior median of d_{jk} has the following closed form:

$$\tilde{d}_{jk}=\text{median}\,(d_{jk}\,|\,\hat{b}_{jK})=\text{sign}(\hat{d}^{\star}_{jk})\max(0,\zeta_{jK}), \qquad (10)$$

Where $\zeta_{jK}=|\hat{d}^{\star}_{jk}|-\sigma\,(A_j[j,j])^{1/2}\,\Phi^{-1}\left(\frac{1+\min(O_{jK},1)}{2}\right)$ and Φ is the standard normal cumulative distribution function. This is an individual thresholding rule where each empirical wavelet coefficient is thresholded utilizing information about neighboring coefficients within a block.

A. Hypothesis testing (0=1-loss): adjust the Bayes Factor (BF) method of Vidakovic (1998a) in the Bayesian term-by-term wavelet function estimation frame to our model, test $\mathrm{H}_0:b_{jK}=0$ versus $\mathrm{H}_1:b_{jK}\neq 0$ and refuse H0 if the posterior odds ratio $O_{jK}=P(\mathrm{H}_0|\hat{b}_{jK})/P(\mathrm{H}_1|\hat{b}_{jK})<1$, where O_{jK} is defined in (8). This imply

$$\tilde{b}_{jK}=\begin{cases}\hat{b}_{jK} & \text{if } O_{jK}<1,\\ 0 & \text{otherwise.}\end{cases} \qquad (11)$$

A block thresholding rule where the whole block of empirical coefficients is thresholded. Finally, owed to the vague priors imposed on the scaling coefficients c_{j_0k} and using (2), any of the over losses will result in estimating c_{j_0k} by their empirical counterparts $\hat{c}_{j_0k}$. Hence, estimate of the values g at the sample points can be obtained by inverting the DWT on the vector consisting of both the empirical scaling coefficients and the shrunken or thresholded empirical wavelet coefficients, obtained from one of (9), (10), or (11). We call the resultant estimators of g as the BlockPostMean, BlockPostMed and Block BF estimators, correspondingly.

B. An empirical Bayes approach to estimate the hyper parameters: In arrange to apply the Bayesian block shrinkage and block thresholding estimators describe in the earlier section, one needs to specify the hyper parameters $\pi_j,\ \tau_j,\ \rho_j\ (j=j_0,\ldots,J-1)$ and σ. idyllically, they could be obtained from some prior in order about, for example, the reliability of the indefinite function g. Such an approach has been considered in Abramovich et al. (1998) and Abramovich and Sapatinas (1999) for independent priors, and extension to our model is quite straightforward. In practice, though, one can make use of an empirical Bayes approach to estimate the hyper parameter in order to get a completely data-based procedure (see, for example, Johnstone and Silverman, 1998; Clyde and George, 1999, ffi000). In this paper, I have used the latter approach to get marginal maximum

possibility estimates of the hyper parameters π_j, τ_j and ρ_j *given* σ. More particularly, at each declaration level j, while the marginal distribution of the empirical blocks $\hat{b}_{jK}$ be a combine of two multivariate normal distributions, the marginal log-likelihood function ℓ_j is proportional to

$$\ell_j(\pi_j,\tau_j,\rho_j,\sigma) \propto \sum_{K=1}^{m_j} \log\left\{\pi_j(\det(B_j))^{-1/2}\exp\left(-\frac{1}{2}\hat{b}'_{jK}B_j^{-1}\hat{b}_{jK}\right)\right.$$
$$\left.+(1-\pi_j)\sigma^{-l_j}\exp\left(-\frac{1}{2\sigma^2}\hat{b}'_{jK}\hat{b}_{jK}\right)\right\}, \qquad (12)$$

where $B_j = \sigma^2 I + V_j$ and V_j is defined in (6). A To form a system of maximum likelihood equations, one must differentiate equation (12) with respect to the parameters and set the resulting expressions to zero; however, these equations are not explicitly solvable, necessitating the use of numerical iterative methods. The conventional approach to determine the maximum likelihood estimates (MLEs) for mixture distributions is employing the Expectation-Maximization (EM) algorithm, as developed by Dempster et al. (1977). In cases with independent priors, the application of the EM algorithm enables one to and the MLEs for π_j and τ_j given σ (as in Johnstone and Silverman, 1998 and Clyde and George, 1999) or the MLEs for π_j, τ_j and σ (as in Clyde and George, 2000). though, in our case, the M-step of the EM algorithm is not explicit, ensuing rather in an impractical application of the process. In this paper, we have measured the following approach. I have estimated σ by the median absolute deviation (as suggested by Donoho and Johnstone (1994) and frequently applied in practice)

$$\hat{\sigma} = \frac{\text{median}(|\{\hat{d}_{J-1,k}: k=0,1,\ldots,2^{J-1}-1\}|)}{0.6745}, \qquad (13)$$

and contain minimized $-\ell_j(\pi_j,\tau_j,\rho_j,\hat{\sigma})$ with respect to π_j, τ_j and ρ_j straight. The log-likelihood function was reparametrised with

$$\pi_j = \frac{1}{1+\exp(-\theta_{1j})}, \quad \tau_j = |\theta_{2j}| \quad \text{and} \quad \rho_j = \frac{2}{\pi}\arctan(\theta_{3j})$$

so that the parameter estimates would lie in the ranges $0 \leqslant \hat{\pi}_j \leqslant 1$, $\hat{\tau}_j \geqslant 0$ and $-1 < \hat{\rho}_j < 1$. The algorithm that used for the minimization of $-\ell_j$ is the Nelder–Mead simplex search method which does not require 5rst derivatives of $-\ell_j$ (see, for example, Everitt 1987, pp. 16–ffi0). The advantages of fixing the value of σ to, for example, the median absolute deviation in (13) are three-fold. Primary it reduces the dimension of the parameter space. next, it eliminates all the singularities in the probability function. Third, simulations indicate that the choice of initial values is not so crucial in finding the MLEs for π_j, τ_j and ρ_j. In exacting, a small scale simulation learn was carried out in order to examine the significance of starting values for the above inference problem. It was create that the convergence from different starting values for π_j, τ_j and ρ_j was always to be of the similar parameter estimates, a result that held for various set of random sample from the prior mixture model (4) and (5).

IV. APPLICATIONS AND COMPARISONS

This section aims to demonstrate the effectiveness of the proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators, while also facilitating a comparison with various established non-Bayesian block wavelet thresholding estimators. Simulated samples and a data set collected in an anaesthesiological study have been used for this purpose. The computational algorithms related to wavelet analysis were performed using the Wave Lab software that is freely available from `http://www-stat.stanford.edu/software/software.html`. The entire study was carried out using the Matlab programming environment.

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

A. Simulation study: The results of the simulation study will now be presented, with the remainder of this section devoted to the discussion of these results. We compare the Block Post- Mean, Block Post Med and Block BF estimators with the non-overlapping Block JS estimate or of Cai (1999) and the overlapping NeighBlock and Neigh Coe?

Estimators of Cai and Silverman (2001). These estimators have been shown to be superior in mean squared error to the Visu Shrink (see Donoho and Johnstone, 1994), Sure Shrink (see Donoho and Johnstone, 1995) and Translation Invariant (see Coifman and Donoho, 1995) estimators, which have become very popular in term-by-term wavelet thresholding. The BlockJS estimator is a block-wise James–Stein rule, with properly chosen block size and threshold value. The NeighBlock and NeighCoeD estimators are, actually, variants of the BlockJS estimator, where the empirical wavelet coefficients are considered in overlapping blocks. I referred to the papers of Cai (1999) and Cai and Silverman (ffi001) for more details. In this simulation study, I evaluated the various wavelet estimators using Daubechies's compactly supported wavelets Symmlet 8 (see Daubechies, 1992, p. 198) and Coi@et 3 (see Daubechies, 1992, p. 258), and primary resolution levels $j_0 = 3$ and 5. I have considered the Blocks, Bumps, Heavi Sine and Doppler functions of Donoho and Johnstone (1994, 1995), which constitute standard tests for wavelet estimators. These functions are supposed to caricature spatially variable signals arising in a number of scienti5c 5elds, including imaging and spectroscopy. For each test function, $M = 100$ samples were generated by adding independent random noise $\epsilon \sim \mathrm{N}(0, \sigma^2)$ to n= 256 (small sample size), 512 (moderate sample size) and 10ffi4 (large sample size) equally spaced points on [0,1].

The value of σ was taken to correspond to the values 3 (high noise level), 5 (moderate noise level) and 7 (low noise level) for the root signal-to-noise ratio (RSNR)

$$\mathrm{RSNR}(g,\sigma) = \frac{\sqrt{(1/n)\sum_{i=1}^{n}(g(t_i)-\bar{g})^2}}{\sigma}, \quad \text{where } \bar{g} = \frac{1}{n}\sum_{i=1}^{n} g(t_i).$$

The goodness-of-fit for an estimator $\hat{g}$ of g was measured by its average mean squared error (AMSE) from the M simulations, defined as

$$\mathrm{AMSE}(g) = \frac{1}{nM}\sum_{m=1}^{M}\sum_{i=1}^{n}(\hat{g}_m(t_i) - g(t_i))^2;$$

its average mean absolute deviation (AMAD) from the M simulations, de5ned as $\mathrm{AMAD}(g) = \frac{1}{nM}\sum_{m=1}^{M}\sum_{i=1}^{n}|\hat{g}_m(t_i) - g(t_i)|;$ and its average maximal absolute deviation (AMXD) from the M simulations, defined as

$$\mathrm{AMXD}(g) = \frac{1}{M}\sum_{m=1}^{M}\max_{1\leqslant i\leqslant n}|\hat{g}_m(t_i) - g(t_i)|.$$

In order to examine the effect of the block lengths on the numerical performance of the proposed wavelet estimators, we performed a preliminary simulation study. In particular, I considered two choices of block length at each resolution level j, Namely $l_j = j$ and $l_j = 2^{[\log_2(Cj)]}$, where $2^{[x]}$ is the largest dyadic integer smaller than or equal to x. For the values of C that we examined, the former choice resulted in longer blocks and it was found to be less adequate than the latter choice, whose block lengths evenly divide the empirical wavelet coefficients into non-overlapping blocks. Furthermore, I found that a value of $C \in [0.4, 0.7]$ was the most reasonable choice uniformly across various test functions, sample sizes and RSNRs. This 5nding is consistent with the remark of Hall et al. (1999) that the selection of block length is not critical if it is chosen within an appropriate range of values. Accordingly, and somewhat arbitrarily, I have taken C =0:5 and suggest $l_j = 2^{[\log_2(j/2)]}$ to be the reference of block length at each resolution level j for the proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators. Thus, there will be 8 blocks of length 1 on the 3rd level, 8 blocks of length 2 on the 4th level, 16 blocks of length 2 on the 5th level, 32 blocks of length 2 on the 6th level, 64 blocks of length 2 on the 7th level, 64 blocks of length 4 on the 8th level, 128 blocks of length 4 on the 9th level and so on.

For brevity, I only reported in detail the results for the Bumps function using Symmlet 8, $j_0 = 3$ and AMSE. Different combinations of test functions, wavelets, primary resolution levels and goodness-of-5t measures yield basically similar results. Fig. 1 contains the results of the simulation study. As observed in the 5gure, the Block Post Mean, Block Post Med and Block BF estimators produced estimates with smaller AMSE than the Block JS, Neigh Block and Neigh Coeff estimators, uniformly across the various combinations of signal morphology and RSNR. Amongst the non-Bayesian

block wavelet thresholding estimators, the Neigh Coeff estimator always produced the smallest AMSE whilst the Block JS estimator always produced the largest. (This is consistent with the conclusions drawn by Cai and Silverman, ffi001.) I noted that each of the Block Post Mean, Block Post Med and Block BF estimators produced less variable estimates than the Block JS, Neigh Block and Neigh Coeff estimators in every case. Quantitatively, the Block Post Med and Block BF estimators are almost identical and both appear to be slightly outperformed by the Block Post-Mean estimator in every case.

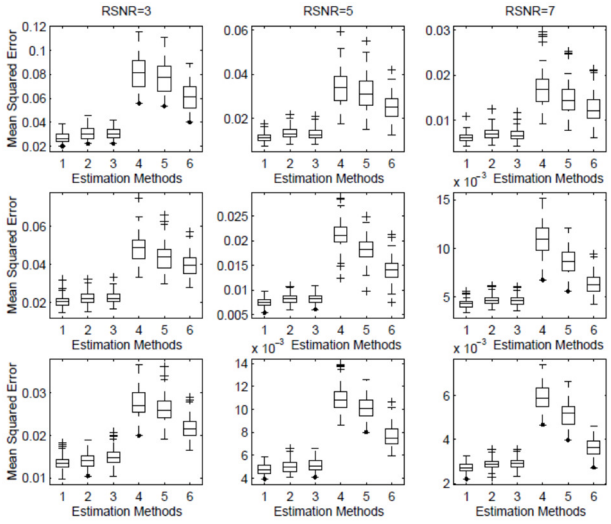

Fig. 1. Boxplots of 100 simulation results for the Bumps function for all nine combinations of signal points (top: 256; middle: 512; bottom: 1024) and RSNRs (left: 3; middle: 5; right: 7). In each panel, there are six box plots indicating the mean squared error, from left to right, for the estimates produced by Block Post Mean, (2) Block Post Med, (3) Block BF, (4) Block JS, (5) Neigh Block and (6) Neigh Coeff.

Table 1 shows the mean and standard deviation of the CPU times involved in computing 100 estimates for the Bumps function by each wavelet estimate or and sample size. As anticipated, the Block JS, Neigh Block and Neigh Coeff estimators are superior in terms of CPU time to the Block Post Mean, Block Post Med and Block BF estimators. This is due to the maximization step described in Section 3.3 that can be computationally expensive, depending on the number of iterations and the convergence criterion used. However, with the fast computing environments that exist nowadays, computing time is not a major issue. To illustrate that the proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators are appealing visually as well as quantitatively, we present in Fig. 2 a noisy Heavi Sine function sampled at n=1024 equally spaced points on [0; 1] with RSNR =7. Fig. 3 displays the reconstructions obtained from one simulation using the six wavelet estimators. All the wavelet estimators recover the underlying function reasonably well. The Block Post Mean, Block Post Med and Block BF estimators give a better reconstruction of the discontinuities in the Heavi Sine function than the Block JS, Neigh Block and Neigh Coeff estimators. At the same time, although they appear to be slightly noisier over the regions where the underlying function is smooth, they are not noisy enough to be visually unpleasant.

Method	CPU time ($n=256$)	CPU time ($n=512$)	CPU time ($n=1024$)
BlockJS	0.0181 (0.0042)	0.0272 (0.0035)	0.0345 (0.0026)
NeighBlock	0.0309 (0.0051)	0.0619 (0.0032)	0.1238 (0.0022)
NeighCoeff	0.0289 (0.0048)	0.0608 (0.0033)	0.1221 (0.0024)
BlockPostMean	4.7554 (0.7123)	6.4543 (0.7256)	8.658 (0.7434)
BlockPostMed	4.8763 (0.7134)	6.5324 (0.7265)	8.7324 (0.7543)
BlockBF	4.6563 (0.7012)	6.3787 (0.7123)	8.5472 (0.7234)

Table 1-Mean and standard deviation (in brackets) of the CPU times involved in computing 100 estimates for the Bumps function by the six wavelet estimators (Block JS, Neigh Block, Neigh Coeff, Block Post Mean, Block Post Med and Block BF) and the three sample sizes (n=26, 512 and 1024)

B. Inductance plethysmography data: I further illustrated the performance of the above wavelet estimators using a data set from anesthesiology collected by inductance plethysmography. The

recordings were made by the Department of Anaesthesia at the Bristol Royal Informary and measure the flow of air during breathing. The same data set has been analyzed in Nason (1996) and Abramovich et al. (1998). I referred to these papers for more details. Fig. 4 shows a section of plethysmograph recording lasting approximately 80 s (n=4096 signal points). The two main sets of regular oscillations correspond to normal breathing. The disturbed behavior in the centre of the plot, where the normal breathing pattern disappears, corresponds to the patient vomiting. Fig. 5 shows the various reconstructions of the plethysmograph recording. The recommended block length $l_j = 2^{[\log_2(j/2)]}$ for the proposed Block Post Mean, Block Post Med and Block BF estimators was used.

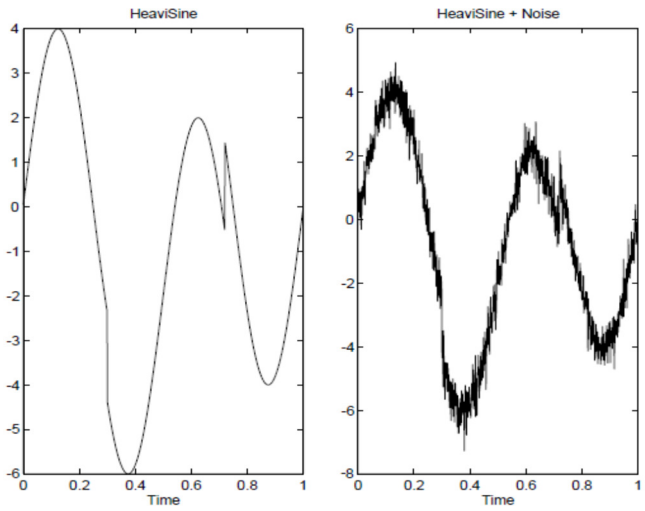

Fig. 2. (Left): the Heavi Sine function based on 10ffi4 signal points; (Right): a noisy version of the Heavi Sine function with RSNR =7.

Plethysmography data	0.8472
BlockPostMean	0.8433
BlockPostMed	0.8433
BlockBF	0.8433
BlockJS	0.8268
NeighCoeff	0.8249
NeighBlock	0.8214

Table 2: Highest peak value (5rst peak) in the data and the corresponding function estimates shown in Fig. 5 for the inductance plethysmography recording of Fig. 4

As observed in the figure, the Block JS and Neigh Block estimators remove the noise but tend to attenuate the peaks, whilst the Neigh Coeff estimator contains some high frequency effects in the region near time 0.8 and also attenuates the peaks. On the other hand, although the Block Post Mean, Block Post Med and Block BF estimators do not eliminate completely the large fluctuation in the region near time 0.8, they produce a satisfying smooth fit without the attenuation. This is apparent, for example, from the heights of the first peak tabulated in Table 2. It is interesting to note that the heights of the first peak value yielded by the Block-Post Mean, Block Post Med and Block BF estimators are identical. This common value is also equal to two decimal places to the value given by the Bayes Thresh method of Abramovich et al. (1998), in which adjustment of one of the hyper parameters of their prior distribution was necessary (see Eq. (9) and Table 1 in this paper form or e-details). The main point here is that the Block Post Mean, Block Post Med and Block BF estimators automatically produce estimates similar to that obtained by Bayes Thresh after ad hoc tuning of the hyper parameters.

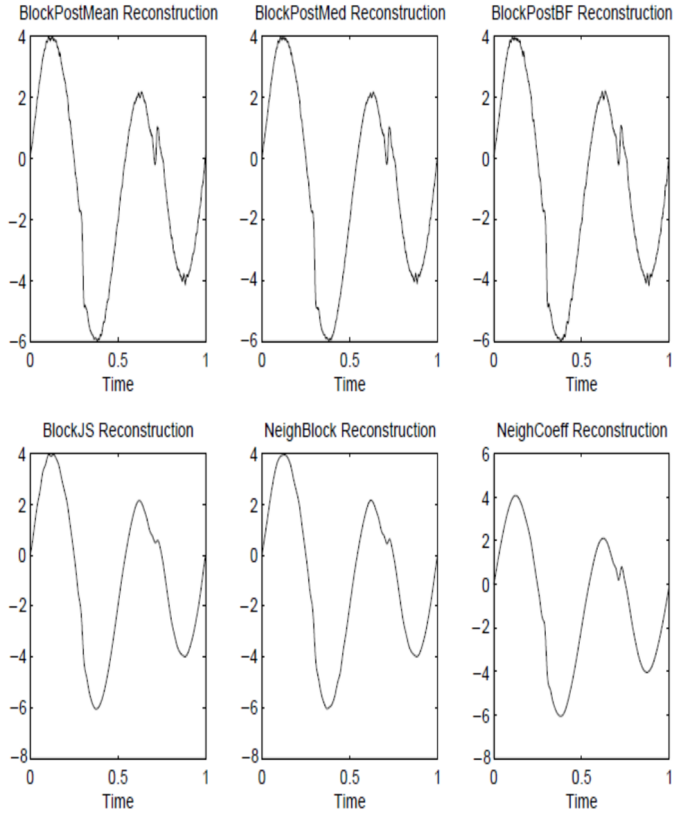

Fig. 3. Reconstructions of the Heavi Sine function from one simulation based on 1024 signal points and RSNR =7 for the estimates produced by, from left to right, (top) Block Post Mean, Block Post Med and Block BF, (bottom) Block JS, Neigh Block and Neigh Coeff.

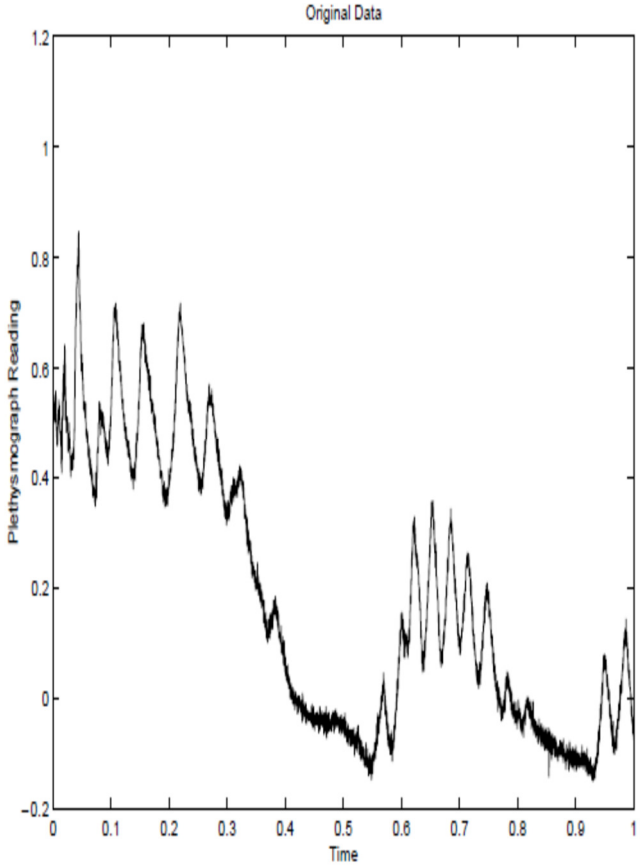

Fig. 4. Section of an inductance plethysmography recording lasting approximately 80 s.

It has been demonstrated that, with an adequate choice of block length at each resolution, the empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators are superior to several existing non-Bayesian block wavelet thresholding estimators in 5nite sample situations. However, due to the maximization inherent in their methodology described in Section 3.3, they are computationally more expensive. The proposed empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators could have been presented as a possibly useful addition to the growing range of wavelet-based function estimation tools.

V. CONCLUSIONS

An empirical Bayes approach to incorporating information on neighboring empirical wavelet coefficients into function estimation that results in block wavelet shrinkage and block wavelet thresholding estimators has been discussed. A simulation study has been conducted to illustrate the resulting wavelet estimators, and these estimators have been compared with several existing non-Bayesian block wavelet thresholding estimators. An application to a data set collected in an anaesthesiological study has also been presented.

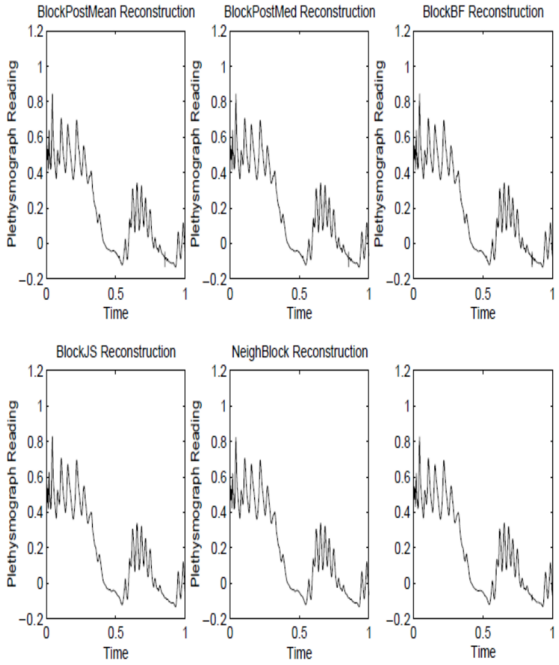

Fig. 5. Smooth estimates of the inductance plethysmography recording shown in Fig. 4 obtained by, from left to right, (top) Block Post Mean, Block Post Med and Block BF (bottom) BlockJS, Neigh Block and NeighCoeff.

The position of the blocks has a significant effect on the performances of the proposed estimators. Although not presented here, by averaging over different block centers, the resulting empirical Bayes block wavelet shrinkage and block wavelet thresholding estimators have been found to have superior numerical performances, at the cost of higher computational complexity. This technique was also used for the classical block wavelet thresholding estimators discussed in Hall et al. (1997), Cai (1999) and Cai and Silverman (ffi001) to show improvements in mean squared error over the standard position of the blocks. However, we believe that, by first choosing the positions and the lengths of the blocks adaptively and then applying our empirical Bayes wavelet estimators to the resulting blocks will produce better estimates of the unknown response function. This is an interesting topic for further research we hope to address elsewhere.

REFERENCES

[1] Abramovich, F., Bailey, T.C., Sapatinas, T., 2000. Wavelet analysis and its statistical applications. The Statistician 49, 1–29.
[2] Chipman, H.A., Kolaczyk, E.D., McCulloch, R.E., 1997. Adaptive Bayesian wavelet shrinkage. J. Am. Statist. Assoc. 92, 1413–1421.
[3] Daubechies, I., 1992. Ten Lectures on Wavelets. SIAM, Philadelphia.
[4] Donoho, D.L., Johnstone, I.M., 1995. Adapting to unknown smoothness via wavelet shrinkage. J. Amer. Statist. Assoc. 90, 1200–1224.
[5] Everitt, B.S., 1987. Introduction to Optimization Methods and their Application in Statistics. Chapman & Hall, London.
[6] Hall, P., Kerkyacharian, G., Picard, D., 1999. On the minimax optimality of block thresholded wavelet estimators. Statist. Sinica 9, 33–50.
[7] Hall, P., Penev, S., Kerkyacharian, G., Picard, D., 1997. Numerical performance of block thresholded wavelet estimators. Statist. Comput. 7, 115–124.
[8] Johnstone, I.M., Silverman, B.W., 1998. Empirical Bayes approaches to mixture problems and wavelet regression. Technical Report, School of Mathematics, University of Bristol, UK.
[9] Mallat, S.G., 1989. A theory for multi-resolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. 11, 674–693.
[10] MEuller, P., Vidakovic, B. (Eds.), 1999. Bayesian Inference in Wavelet Based Models. Lecture Notes in Statistics, Vol. 141. Springer, New York.
[11] Nason, G.P., 1996. Wavelet shrinkage using cross-validation. J. Roy. Statist. Soc. B 58, 463–479.
[12] Ogden, R.T., 1997. Essential Wavelets for Statistical Applications and Data Analysis. BirkhEauser, Boston.
[13] Vannucci, M., Corradi, F., 1999a. Covariance structure of wavelet coefficients: theory and models in a Bayesian perspective. J. Roy. Statist. Soc. B 61, 971–986.
[14] Vannucci, M., Corradi, F., 1999b. Modelling dependence in the wavelet domain. In: MEuller, P.,
[15] Vidakovic, B. (Eds.), Bayesian Inference in Wavelet Based Models, Lecture Notes in Statistics, Vol. 141. Springer, New York, pp. 173–186.
[16] Vidakovic, B., 1998a. Non-linear wavelet shrinkage with Bayes rules and Bayes factors. J. Amer. Statist. Assoc. 93, 173–179.
[17] Vidakovic, B., 1998b. Wavelet based nonparametric Bayes methods. In: Dey, D.D., MEuller, P., Sinha, D. (Eds.), Practical Nonparametric and Semi parametric Bayesian Statistics, Lecture Notes in Statistics, Vol. 133. Springer, New York, pp. 133–155.
[18] Vidakovic, B., 1999. Statistical Modeling by Wavelets. Wiley, New York.
[19] Vidakovic, B., MEuller, P., 1995. Wavelet shrinkage with affine Bayes rules with applications. Technical Report, Institute of Statistics and Decision Sciences, Duke University, USA.

