Journal of Statistics: Advances in Theory and Applications
Volume 26, Number 1, 2021, Pages 13-44
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jsata_7100122240

A COMPARISON SIMULATION STUDY OF STANDARD WAVELET SHRINKAGE METHODS IN NONPARAMETRIC REGRESSION MODELS WITH POSITIVE NOISE

ALEX RODRIGO DOS S. SOUSA

Department of Statistics
University of São Paulo
Brazil
e-mail: alex.sousa89@usp.br

Abstract

Shrinkage estimators are usually applied to estimate wavelet coefficients by reducing the magnitudes of wavelet empirical coefficients in nonparametric regression modelling. There are several well succeeded available wavelet shrinkage estimators in the literature, but most of them work under the assumption of Gaussian noise in the original data. Although Gaussian noise might be observed in practice and allows several good estimation properties, it is not a general case. One might have data with additive non-gaussian noise and, specifically for this work, strictly positive noise. This paper evaluates the performance of standard wavelet shrinkage estimators in denoising data under

2020 Mathematics Subject Classification: 62G08.
Keywords and phrases: positive noise, wavelet shrinkage, nonparametric regression.
Received October 27, 2021; Revised November 9, 2021

positive noise by the conduction of simulation studies involving the so called Donoho and Johnstone test functions. It was observed good performances of bayesian shrinkage rules in terms of averaged mean squared and averaged median absolute errors measures.

1. Introduction

Statistical curve estimation techniques have been applied in several fields, such as engineering, economy, geology, among others. One of the most used techniques considers a nonparametric regression model, where points are observed from an unknown function with an additive random noise. The function can be estimated by its representation in terms of a linear combination of basis functions, such as polynomials, splines, Fourier, wavelets, and others, in such a way that the problem of estimating the function becomes a problem of estimating a finite number of basis coefficients of the expansion. The choice of the convenient basis to be considered depends on the function space that the unknown function belongs to, and also on the features of the function, if they are known.

The focus of this work is wavelets basis expansion, which is very suitable for estimating functions with important local features, as discontinuities, peaks and oscillations, once wavelets are well localized both in time and frequency, i.e., the nonzero wavelet coefficients are usually localized at features to be recovered of the function to be estimated, and most of the remaining coefficients are zero or very close to zero, at smooth regions of the function. Thus, wavelets typically represent a function in a sparse way. See Vidakovic [23] for more details about wavelet-based methods in Statistics. Further, see also Daubechies [5] and Mallat [17] for theoretical descriptions about wavelets.

Although most of the wavelet coefficients of a function representation are zero, noisy wavelet coefficients, that are usually called empirical wavelet coefficients, are obtained in practice after the application of a discrete wavelet transformation on the data. In this sense, several shrinkage and thresholding estimators have been proposed to estimate the wavelet coefficients by reducing the magnitudes of their empirical counterparts, or even estimating them as zero if the empirical coefficients are smaller than a certain threshold value. The seminal papers of Donoho [6-9] and Donoho and Johnstone [10-12] brought the spirit of these ideas, with the proposition of the so called soft and hard thresholding estimators and their properties, among other important results. Since these series of papers, shrinkage and thresholding rules have been developed and can be seen in Vidakovic [23], Jansen [14], and Nason [20].

One common feature of standard shrinkage and thresholding methods for wavelet coefficients estimation is the Gaussian noise assumption on the data. Although this context might occur in practice, it is not a general case. Few works concerning to wavelet shrinkage in nongaussian noise data can be found in the literature. Neumann and von Sachs [21], Leporini and Pesquet [16], Antoniadis et al. [2], and Averkamp and Houdré [3] are some of relevant works in this sense, but no one of them considers data under strictly positive noise. Recently, Sousa and Garcia [22] proposed bayesian wavelet shrinkage rules based on the logistic and beta priors to the wavelet coefficients of original data under exponential and log-normal noises, and they had good performances in the simulation studies against standard techniques. However, the proposed estimators are computational expensive, infeasible for large sample sizes, once they require Markov Chain Monte Carlo methods implementation. In this sense, it is easy to formulate a simple question: how do standard shrinkage and thresholding methods perform when dealing with data with positive noise? This work addresses this question.

The goal of this work is to evaluate the performances of standard wavelet shrinkage and thresholding methods in nonparametric regression models with positive noise. Simulation studies were conducted by considering eight standard and well known methods: Universal thresholding, that was proposed by Donoho and Johnstone [10], False Discovery Rate (FDR) by Abramovich and Benjamini [1], Cross Validation by Nason [19], SUREshrink by Donoho and Johnstone [12], Bayesian Adaptive Multiresolution Smoother (BAMS) by Vidakovic and Ruggeri [24], Large Posterior Mode (LPM) by Cutillo et al. [4], Amplitude-Scale

Invariant Bayes Estimator (ABE) by Figueiredo and Nowak [13], and Empirical Bayes Thresholding by Johnstone and Silverman [15]. Further, we considered positive random noises under uniform, exponential and lognormal distributions.

This paper is organized as follows: the considered statistical models and some background about discrete wavelet transformation and wavelet shrinkage are provided in Section 2. The considered shrinkage and thresholding methods in the simulation studies are described in Section 3. Simulation studies and results are analyzed in Section 4. We conclude the paper with some considerations in Section 5.

2. Statistical Models and Wavelet Analysis

We consider $n = 2^J$, $J \in \mathbb{N}$, points $(x_1, y_1), \cdots, (x_n, y_n)$ from the nonparametric regression model

$$y_i = f(x_i) + e_i, \quad i = 1, \cdots, n, \tag{2.1}$$

where x_i, $i = 1, \cdots, n$, are equispaced scalars, $f \in L^2(\mathbb{R}) = \{f : \int f^2 < \infty\}$ is an unknown function and e_i are independent and identically distributed (iid) random noises such that $e_i > 0$, $i = 1, \cdots, n$. The goal is to estimate f without assumptions about its functional structure, i.e., the estimation procedure will take only the data points into account. To do so, we represent f in wavelet basis expansion as

$$f(x) = \sum_{j, k \in \mathbb{Z}} \theta_{j,k}\, \psi_{j,k}(x), \tag{2.2}$$

where $\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k),\ j, k \in \mathbb{Z}\}$ is an orthonormal wavelet basis for $L^2(\mathbb{R})$ that is constructed by dilations j and translations k of a function ψ called wavelet or mother wavelet and $\theta_{j,k}$ are wavelet coefficients that describe features of f at spatial location $2^{-j}k$ and scale 2^j or resolution level j.

We consider here random noises under uniform, exponential and lognormal distributions. Their densities are given by

• Uniform noise: $e_i \sim \mathrm{U}(0, b)$

$$h(e_i;\, b) = \frac{1}{b}\, \mathbb{I}_{(0,b)}(e_i), \quad b > 0; \tag{2.3}$$

• Exponential noise: $e_i \sim \mathrm{Exp}(\beta)$

$$h(e_i;\, \beta) = \frac{1}{\beta} \exp\left\{-\frac{e_i}{\beta}\right\} \mathbb{I}_{(0,\infty)}(e_i), \quad \beta > 0; \tag{2.4}$$

• Lognormal noise: $e_i \sim \mathrm{LN}(0, \tau)$

$$h(e_i;\, \tau) = \frac{1}{e_i \tau \sqrt{2\pi}} \exp\left\{-\frac{\log^2(e_i)}{2\tau^2}\right\} \mathbb{I}_{(0,\,\infty)}(e_i), \quad \tau > 0, \tag{2.5}$$

where $\mathbb{I}_A(\cdot)$ is the usual indicator function on the set A and $\log(\cdot)$ is the natural logarithm.

The estimation process of f starts by estimation of the wavelet coefficients. In vector notation, model (2.1) can be written as

$$\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{e}, \tag{2.6}$$

where $\boldsymbol{y} = [y_1, \cdots, y_n]'$, $\boldsymbol{f} = [f(x_1), \cdots, f(x_n)]'$ and $\boldsymbol{e} = [e_1, \cdots, e_n]'$. A discrete wavelet transform (DWT), which is typically represented by an orthonormal transformation matrix $\boldsymbol{W}_{n\times n} = (w_{ij})_{1\le i,\, j\le n}$, is applied on both sides of (2.6), obtaining the following model in wavelet domain:

$$\boldsymbol{d} = \theta + \varepsilon, \tag{2.7}$$

where $\boldsymbol{d} = \boldsymbol{Wy}$ is called empirical coefficients vector, $\theta = \boldsymbol{Wf}$ is the wavelet coefficients vector, and $\varepsilon = \boldsymbol{We}$ is the random noise vector. It should be observed that although the random noises e_i in the original nonparametric model are positive, the random noises ε_i in the wavelet domain are not necessarily positive. The estimation of $\boldsymbol{\theta}$ is done by the application of a shrinkage estimator $\boldsymbol{\delta}$ on the empirical coefficients $\boldsymbol{d}$, which acts coefficient by coefficient due to the decorrelation property of DWT, i.e., $\boldsymbol{\delta}(\boldsymbol{d}) = [\delta(d_1), \cdots, \delta(d_n)]'$. Thus, the estimator $\hat{\theta}$ of a single wavelet coefficient θ is

$$\hat{\theta} = \delta(d),$$

where d is the respective single empirical coefficient at the same vector position of θ. Finally, $\boldsymbol{f}$ is then estimated by the inverse discrete wavelet transform (IDWT),

$$\hat{\boldsymbol{f}} = \boldsymbol{W}^t \hat{\boldsymbol{\theta}}.$$

The assumption of zero mean Gaussian noises in the original regression model (2.1) brings several estimation advantages. The most important one is that random noises in the wavelet domain remain zero mean gaussian with the same scale parameter as in the original model. Thus, the wavelet coefficient estimation problem becomes a normal location parameter estimation issue. In this sense, most of the standard shrinkage and thresholding methods were developed under Gaussian noise assumption.

In this work, we are interested in analyzing the performance of some of the most applied standard techniques in estimating wavelet coefficients under positive noise in the original model. It is clear that most of the statistical properties of the estimators are lost under this setup, but it would be interesting if these methods could still be applied in these contexts.

3. Comparison Methods

In this section, we describe the considered methods for empirical wavelet coefficients denoising in the simulation studies. These methods may be shrinkage or thresholding. A shrinkage rule reduces the magnitude of an empirical coefficient but does not necessarily assign it to zero. On the other hand, a thresholding rule shrinks a small coefficient to exactly zero. Moreover, a thresholding rule may have different criteria for choosing the threshold value. It will be seen that Universal, FDR, Cross Validation, and SUREshrink methods provide criteria for eliciting the threshold value for one considered thresholding policy, the soft thresholding. These four methods are considered here as classical thresholding methods. Further, we also considered four bayesian methods, such that one of them is a shrinker, BAMS, and the three remaining ones are thresholding, LPM, ABE, and EBAYES.

The eight considered methods were developed under the model (2.1), but with the assumption that the random noises e_i are i.i.d. normal with zero mean and variance σ^2. The noise standard deviation σ can be estimated by the Median Absolute Deviation (MAD) of the empirical coefficients of the finest resolution level, according Donoho and Johnstone [10],

$$\hat{\sigma} = \frac{\text{median}\left\{\left|d_{J-1,k}\right| : k = 0, \dots, 2^{J-1}\right\}}{0.6745}. \tag{3.1}$$

Although the techniques were originally proposed under gaussian noise assumption, we are interested in evaluate if they can be applied for denoising data under positive noise. A direct impact of this assumption is that the random noises do not have zero mean, but a mean greater than zero. This is an important structural change, but the above mentioned fact that the random noises in the wavelet domain (2.7) are not necessarily positive must be in mind. The shrinkage and thresholding methods will act on empirical coefficients with real valued random noises, although original data are positive noisy in (2.1). Then, we will keep (3.1) as a suitable estimate of σ for the standard methods applications, once the empirical coefficients at finest resolution level remain the best source of information about random noises in the wavelet domain.

Next, we give a brief description about the considered methods. Observe that the chosen methods were developed according to a variety of properties and approaches. For example, some of them are based on nonparametric methods (Cross Validation), on hypothesis testing (FDR), on probabilistic approach (Universal) and on a risk measure (SUREshrink). Further, the Bayesian rules are built based on the posterior mean (BAMS and ABE), posterior mode (LPM), and posterior median (EBAYES) of the wavelet coefficient distributions. Again, the fact that noises in wavelet domain are not positive assures that these estimators remain suitable even under positive noise in the original data.

3.1. Standard thresholding techniques

The thresholding techniques in the simulation studies are based on the soft thresholding policy introduced by Donoho and Johnstone [11] and given by

$$\delta(d) = \begin{cases} 0, & \text{if } |d| \le \lambda, \\ \operatorname{sgn}(d)(|d| - \lambda), & \text{if } |d| > \lambda, \end{cases} \tag{3.2}$$

where $\lambda > 0$ is the threshold value and $\operatorname{sgn}(d)$ represents the sign of d. In fact, the thresholding rule (3.2) shrinks empirical wavelet coefficients with magnitudes in $[-\lambda, \lambda]$ to zero and reduces their magnitudes in λ unities for those outside this interval. Larger values of λ imply in higher degrees of thresholding performed by the rule, resulting in a more sparse estimated coefficients vector $\hat{\boldsymbol{\theta}}$. Figure 1 shows an example of a soft thresholding rule for $\lambda = 5$ and $d \in [-20, 20]$.

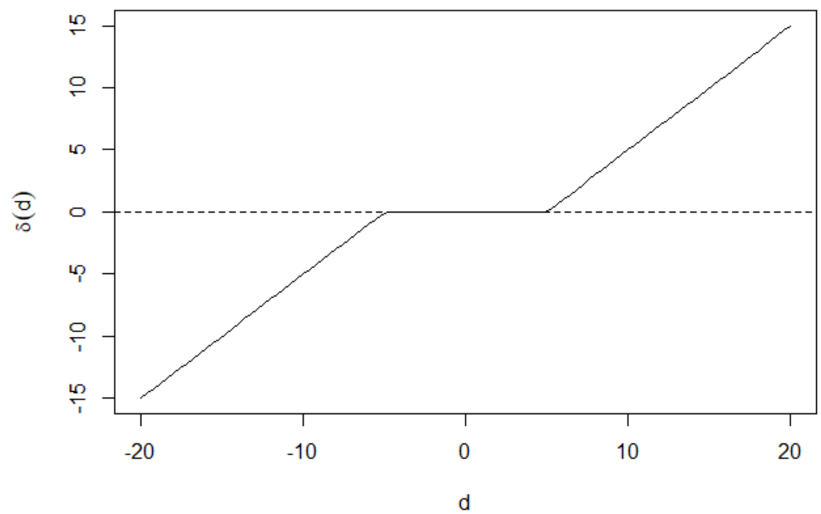

Figure 1. Soft thresholding rule (3.2) for $\lambda = 5$.

There are several methods for choosing λ in the literature, see Vidakovic [23]. For the simulation studies of this work, we considered four of them, Universal, False Discovery Rate, Cross Validation and SUREshrink techniques, that will be described in the next subsections.

3.1.1. Universal (UNIV)

Donoho and Johnstone [11] proposed to choose λ depending only on the sample size n and the noise standard deviation σ as

$$\lambda_{\text{UNIV}} = \sigma\sqrt{2\log(n)}. \tag{3.3}$$

The universal threshold rule has important asymptotic minimax properties and its threshold value $\sigma\sqrt{2\log(n)}$ can be viewed by a probabilistic approach, once for independent standard normal random variables $\{X_i\}_{i=1}^{n}$, $\mathbb{P}\left(\max\{|X_i|\} > \sqrt{2\log(n)}\right) \to 0$, as $n \to \infty$. Thus universal thresholding has a high probability of Gaussian noise removal from the empirical coefficients, see Vidakovic [23] for more details.

3.1.2. False discovery rate (FDR)

False discovery rate method was proposed by Abramovich and Benjamini [1] and selects a threshold value based on hypothesis testing involving the wavelet coefficients. It works by the following steps, that were obtained by Nason [20]:

- For each d_i, calculate the associated p-value p_i for testing $H_0 : \theta_i = 0$ against $H_1 : \theta_i \neq 0$, given by $p_i = 2((1 - \Phi(|d_i|/\sigma))$, $i = 1, \cdots, n$.

- Order the obtained p-values: $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}$.

- Let i_0 be the largest i such that $p_{(i)} \leq (i/n)q$ for some level q.

- The threshold value is given by

$$\lambda_{\text{FDR}} = \sigma\Phi^{-1}(1 - p_{i_0}/2), \tag{3.4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution.

3.1.3. Cross validation (CV)

Cross validation methods are commonly used to estimate predictive errors. Their main idea is to separate the original sample in two parts, one to adjust the statistical model by estimating parameters (trainning set) and the other one to evaluate the quality of adjustment (test set) by applying it on the estimated model and comparing their predictions against the observed values. Nason [19] proposed the threshold to be chosen according to the following cross validation criterion, that was obtained by Morettin [18]:

- Remove observations y_i with odd index i.

- Estimate f by $\hat{f}^{(\text{even})}$ based on the remaining observations y_i with even indexes and a particular threshold value λ.

• Obtain $\bar{y}_i^{(\text{odd})} = \begin{cases} (y_{2i-1} + y_{2i+1})/2, & \text{for } i = 1, \cdots, n/2 - 1, \\ (y_1 + y_{n-1})/2, & \text{for } i = n/2. \end{cases}$

Obtain $\hat{f}^{(\text{odd})}$ and $\bar{y}_i^{(\text{even})}$ similarly.

• The optimal threshold value is given by

$$\lambda_{\text{CV}} = \arg\min_{\lambda} \hat{M}(\lambda), \tag{3.5}$$

where

$$\hat{M}(\lambda) = \sum_i [(\hat{f}^{(\text{even})}(x_i) - \bar{y}_i^{(\text{odd})})^2 + (\hat{f}^{(\text{odd})}(x_i) - \bar{y}_i^{(\text{even})})^2]. \tag{3.6}$$

3.1.4. SUREshrink (SURE)

Donoho and Johnstone [12] proposed a threshold choice based on the Stein's unbiased risk estimation (SURE) given by

$$\text{SURE}(\lambda; \boldsymbol{d}) = n - 2\sum_{i=1}^{n} \mathbb{I}\{|d_i| \le \lambda\sigma\} + \sum_{i=1}^{n} \min\{(d_i/\sigma)^2, \lambda^2\}, \tag{3.7}$$

that estimates the risk $\mathbb{E}[(\boldsymbol{\delta}(\boldsymbol{d}) - \boldsymbol{d})^2]$ for $\boldsymbol{\delta}(\boldsymbol{d})$ being the soft thresholding rule (3.2). Thus, the optimal threshold value λ_{SURE} is the argument that minimizes (3.7),

$$\lambda_{\text{SURE}} = \arg\min_{0 \le \lambda \le \sqrt{2\log(n)}} \text{SURE}(\lambda; \boldsymbol{d}). \tag{3.8}$$

3.2. Bayesian techniques

3.2.1. Bayesian adaptive multiresolution smoother (BAMS)

BAMS was proposed by Vidakovic and Ruggeri [24] and is obtained by considering $\sigma^2 \sim \text{Exp}(1/\mu)$, $\mu > 0$, and the following prior distribution π to a single wavelet coefficient:

$$\pi(\theta) = \alpha\delta_0(\theta) + (1 - \alpha)\text{DE}(0, \tau), \tag{3.9}$$

where $\alpha \in (0, 1)$, $\delta_0(\cdot)$ is a point mass function at zero and $\mathrm{DE}(0, \tau)$ is the double exponential density with location parameter equals to zero and scale parameter τ, $\tau > 0$, given by $g(\theta; 0, \tau) = \frac{1}{2\tau} \exp\left\{-\frac{|\theta|}{\tau}\right\}$, $\theta \in \mathbb{R}$. The associated shrinkage rule under squared error loss, $\delta_{\mathrm{BAMS}}(d) = \mathbb{E}(\theta|d)$ is

$$\delta_{\mathrm{BAMS}}(d) = \frac{(1-\alpha)m(d)\delta(d)}{(1-\alpha)m(d) + \alpha \mathrm{DE}\left(0, \frac{1}{\sqrt{2\mu}}\right)}, \tag{3.10}$$

where $m(\cdot)$ and $\delta(\cdot)$ are the predictive distribution of d and the shrinkage rule respectively under assumption that $\theta \sim \mathrm{DE}(0, \tau)$ and given by

$$m(d) = \frac{\tau \exp\left\{-\frac{|d|}{\tau}\right\} - \frac{1}{\sqrt{2\mu}} \exp\left\{-\sqrt{2\mu}|d|\right\}}{2\tau^2 - \frac{1}{\mu}},$$

and

$$\delta(d) = \frac{\tau\left(\tau^2 - \frac{1}{2\mu}\right) d \exp\{-|d|/\tau\} + \frac{\tau^2}{\mu}\left(\exp\left\{-|d|/\sqrt{2\mu}\right\} - \exp\{-|d|/\tau\}\right)}{\left(\tau^2 - \frac{1}{2\mu}\right)\left(\tau \exp\{-|d|/\tau\} - \frac{1}{\sqrt{2\mu}} \exp\left\{-|d|/\sqrt{2\mu}\right\}\right)}.$$

3.2.2. Large posterior mode (LPM)

Cutillo et al. [4] proposed a bayesian thresholding rule that is based on the Maximum a Posteriori (MAP) principle. Under the model (2.1) with Gaussian noise, it is assumed a normal prior for θ, i.e., $\theta|t^2 \sim N(0, t^2)$, where $t^2 \sim (t^2)^{-k}$, $k > 1/2$. The LPM thresholding rule picks the larger mode in absolute value of the posterior distribution of $\theta|d$. Further, there is an interesting feature of the posterior that allows the Bayes rule to be thresholding. The posterior can be unimodal at zero or bimodal trivially at zero and at another local mode. For the unimodal case, the empirical coefficient is then shrunk to zero. For the second one, it is shrunk slightly to the local mode. The closed form of the rule is

$$\delta_{\text{LMP}}(d) = \frac{d + \text{sgn}(d)\sqrt{d^2 - 4\sigma^2(2k-1)}}{2}\,\mathbb{I}_{[\lambda_{\text{LPM}},+\infty)}(|d|), \tag{3.11}$$

where $\lambda_{\text{LPM}} = 2\sigma\sqrt{2k-1}$.

3.2.3. Amplitude-scale invariant Bayes estimator (ABE)

Figueiredo and Nowak [13] proposed a bayesian thresholding rule that does not depend on prior hyperparameters, which must be elicited for other bayesian shrinkage/thresholding methods. It assumes an amplitude-scale invariant (noninformative) prior $\pi(\theta) \propto |\theta|^{-1}$ for θ from the model (2.7). The Bayes rule is thresholding, given by

$$\delta_{\text{ABE}}(d) = \frac{(d^2 - 3\sigma^2)_+}{d}, \tag{3.12}$$

where $(x)_+ = \max\{0, x\}$. Note that the rule depends only on the noise variance parameter σ^2.

Figure 2 shows BAMS, LPM and ABE rules. It is clear the shrinkage behaviour of BAMS, i.e., although it shrinks the empirical coefficient magnitude, the shrunk coefficient is not necessarily zero. On the other hand, LPM and ABE rules are thresholding, once they shrink a sufficiently small coefficient to zero.

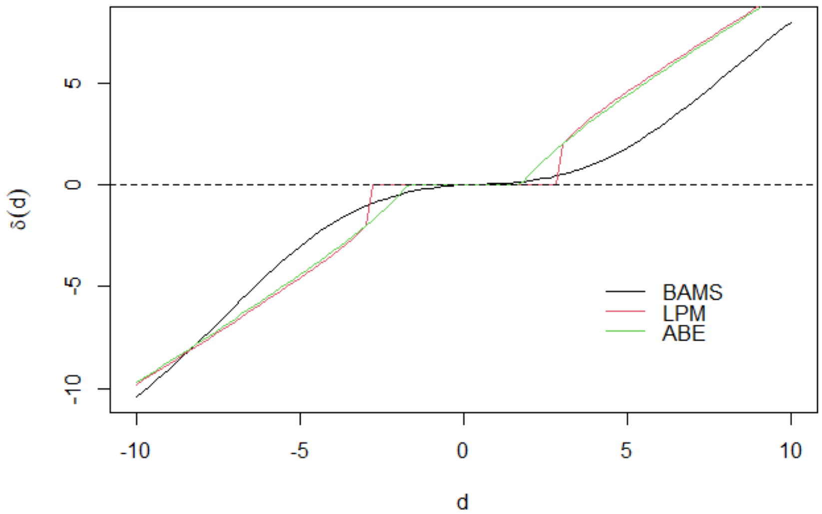

Figure 2. BAMS, LPM and ABE bayesian rules. The first one is a shrinkage rule while LPM and ABE are thresholding rules.

3.2.4. Empirical Bayes thresholding (EBAYES)

This method was proposed by Johnstone and Silverman [15] and provides a thresholding rule based on the posterior median of $\theta|d$ (although the method is also developed for the posterior mean in the original paper). The proposed prior for θ is

$$\pi(\theta) = \alpha\delta_0(\theta) + (1-\alpha)g(\theta; \boldsymbol{a}), \tag{3.13}$$

where $\alpha \in (0, 1)$ and g is a heavy-tailed unimodal and symmetric distribution, such as double exponential or a quasi-Cauchy distributions. The hyperparameters α and $\boldsymbol{a}$ are estimated by the marginal maximum likelihood (MML) of the empirical coefficient d, i.e., they are obtained from the data, which is typical for empirical Bayes procedures. If $\tilde{F}(\theta|d) = \int_\theta^\infty \pi(u|d)du$ and $\hat{\alpha}(d)$ is the estimate of α by the MML, then

$$\delta_{\text{EBAYES}}(d) = \begin{cases} 0, & \text{if } \hat{\alpha}(d)\tilde{F}(0|d) \leq 1/2, \\ \tilde{F}^{-1}\left(\frac{1}{2\hat{\alpha}(d)}\Big|d\right), & \text{otherwise.} \end{cases} \tag{3.14}$$

4. Simulation Studies and Results

The performances of the rules in denoising artificial data and estimating curves were obtained in simulation studies. The considered underlying functions in model (2.1) were the Donoho-Johnstone (D-J) test functions called Bumps, Blocks, Doppler, and Heavisine, which are extremely applied in wavelet-based methods research. These functions have interesting features such as peaks, discontinuities, oscillations and constant parts that usually appear in practice to be recovered. Table 1 shows D-J functions definitions for $x \in [0, 1]$ and Figure 3 presents their respective plots.

Table 1. Donoho-Johnstone test functions definitions for $x \in [0, 1]$ used as underlying functions in the simulation studies

Donoho-Johnstone test functions
BUMPS
$f(x) = \sum_{l=1}^{11} h_l K\left(\frac{x - x_l}{w_l}\right)$
where
$K(x) = (1 + \|x\|)^{-4}$
$(x_l)_{l=1}^{11} = (0.1, 0.13, 0.15, 0.23, 0.25, 0.40, 0.44, 0.65, 0.76, 0.78, 0.81)$
$(h_l)_{l=1}^{11} = (4, 5, 3, 4, 5, 4.2, 2.1, 4.3, 3.1, 5.1, 4.2)$
$(w_l)_{l=1}^{11} = (0.005, 0.005, 0.006, 0.01, 0.01, 0.03, 0.01, 0.01, 0.005, 0.008, 0.005)$
BLOCKS
$f(x) = \sum_{l=1}^{11} h_l K(x - x_l)$
where
$K(x) = (1 + \text{sgn}(x))/2$
$(x_l)_{l=1}^{11} = (0.1, 0.13, 0.15, 0.23, 0.25, 0.40, 0.44, 0.65, 0.76, 0.78, 0.81)$
$(h_l)_{l=1}^{11} = (4, -5, 3, -4, 5, -4.2, 2.1, 4.3, -3.1, 2.1, -4.2)$
DOPPLER
$f(x) = \sqrt{x(1-x)} \sin\left(\frac{2.1\pi}{x + 0.05}\right)$
HEAVISINE
$f(x) = 4 \sin(4\pi x) - \text{sgn}(x - 0.3) - \text{sgn}(0.72 - x)$

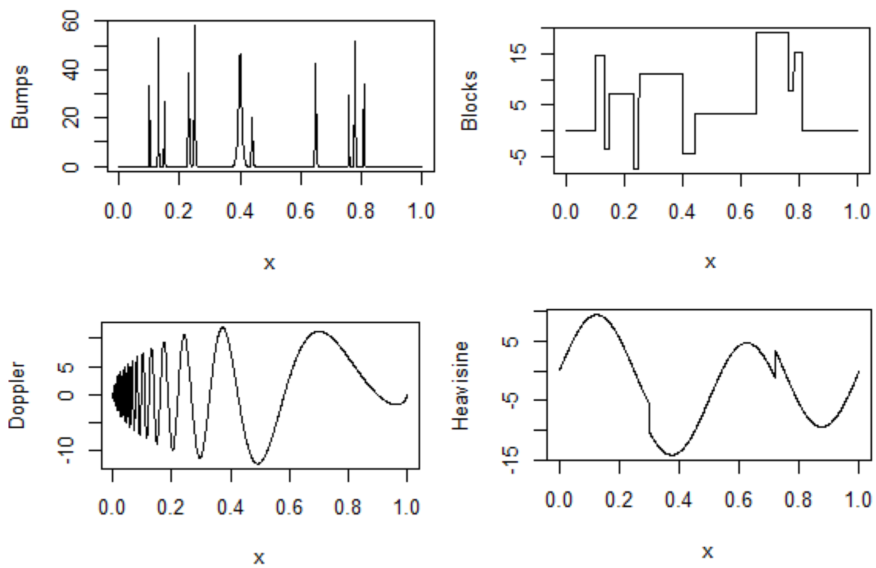

Figure 3. Donoho-Johnstone test functions.

For each underlying function, datasets were generated by adding random noise according to model (2.1) with uniform, exponential and lognormal noise distributions (2.3), (2.4), and (2.5), respectively. For this data generation process, we considered two sample sizes, $n = 32$ and 2048, and the parameters of the noise distributions were determined according to two considered signal to noise ratio (SNR) values, $\text{SNR} = 3$ and 9.

We adopted $M = 200$ replications for each scenario of underlying function, noise distribution, sample size and signal to noise ratio value and, for each shrinkage/thresholding rule, two performance measures were calculated, the averaged mean squared error (AMSE) given by

$$\text{AMSE} = \frac{1}{Mn} \sum_{m=1}^{M} \sum_{i=1}^{n} [\hat{f}^{(m)}(x_i) - f(x_i)]^2,$$

where $\hat{f}^{(m)}(\cdot)$ is the estimate of the function at a particular point in the m-th replication and the averaged median absolute error (AMAE),

$$\text{AMAE} = \frac{1}{M}\sum_{m=1}^{M} \text{median}\{|\hat{f}^{(m)}(x_i) - f(x_i)| : i = 1, \cdots, n\}.$$

Tables 2, 3, and 4 show the AMSEs and AMAEs of the methods in the simulation studies under uniform, exponential and lognormal noises respectively, with the best method performance in bold for each scenario. For the three noise distributions, the methods had performances closed to each other for both measures, but mainly for AMAE. In general, the methods performed better for denoising data under exponential noise than uniform and lognormal noises and the worst general performances were under uniform noise. Further, the methods worked better for scenarios with $n = 2048$ and $\text{SNR} = 9$, as expected. The increasing performances occured mainly for $\text{SNR} = 9$ against $\text{SNR} = 3$, i.e., the SNR value had more impact on the performance measures than the sample size, although better works were observed for $n = 2048$ than for $n = 32$.

Table 2. AMSE and AMAE of the shrinkage and thresholding rules in the simulation study for DJ-test functions under **uniform** noise

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Bumps	32	UNIV	46.583 (5.380)	8.224 (2.249)
		FDR	43.597 (5.229)	15.167 (2.398)
		CV	50.158 (5.438)	24.662 (2.638)
		SURE	44.402 (5.159)	8.993 (1.874)
		BAMS	**20.650 (3.906)**	4.380 (1.599)
		LPM	21.865 (4.038)	**2.414 (1.310)**
		ABE	31.470 (5.159)	3.454 (1.378)
		EBAYES	27.789 (4.517)	3.575 (1.538)
	2048	UNIV	28.694 (4.676)	4.516 (1.508)
		FDR	20.280 (4.216)	2.497 (1.413)
		CV	17.822 (4.122)	2.187 (1.392)
		SURE	18.075 (4.138)	2.082 (1.382)
		BAMS	17.579 **(4.022)**	2.216 (1.364)
		LPM	21.734 (4.035)	2.417 **(1.348)**
		ABE	18.515 (4.143)	2.128 (1.381)
		EBAYES	**17.427** (4.092)	**1.987** (1.367)
Doppler	32	UNIV	26.793 (4.107)	4.995 (1.440)
		FDR	29.564 (4.079)	6.659 (1.473)
		CV	23.956 (4.054)	5.117 (1.445)
		SURE	22.764 (4.053)	2.522 (1.348)
		BAMS	**18.358 (3.766)**	2.696 **(1.261)**
		LPM	21.629 (4.020)	**2.367** (1.347)
		ABE	23.618 (3.941)	2.630 (1.315)
		EBAYES	21.634 (4.130)	3.152 (1.440)
	2048	UNIV	19.064 (4.065)	2.475 (1.387)
		FDR	17.491 (4.087)	2.042 (1.366)
		CV	16.813 (4.056)	1.905 (1.354)
		SURE	16.899 (4.065)	1.916 (1.357)
		BAMS	17.034 **(4.024)**	1.894 **(1.304)**
		LPM	21.750 (4.051)	2.419 (1.345)
		ABE	16.823 (4.090)	1.917 (1.358)
		EBAYES	**16.625** (4.049)	**1.877** (1.349)

Table 2. (Continued) right part

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Blocks	32	UNIV	33.111 **(3.618)**	9.488 (1.760)
		FDR	32.943 (3.634)	17.459 (2.010)
		CV	32.759 (3.628)	16.679 (1.922)
		SURE	33.051 (3.627)	17.467 (2.004)
		BAMS	**21.450** (3.834)	5.298 (1.542)
		LPM	21.930 (4.131)	**2.440 (1.345)**
		ABE	31.607 (4.383)	4.298 (1.530)
		EBAYES	29.154 (4.231)	14.268 (2.263)
	2048	UNIV	23.886 **(3.751)**	3.837 (1.365)
		FDR	19.316 (3.956)	2.386 (1.349)
		CV	17.588 (4.027)	2.052 (1.342)
		SURE	17.816 (4.017)	2.049 (1.341)
		BAMS	17.508 (4.018)	2.154 **(1.328)**
		LPM	21.809 (4.042)	2.419 (1.345)
		ABE	18.148 (4.078)	2.110 (1.362)
		EBAYES	**17.410** (4.047)	**1.964** (1.350)
Heavisine	32	UNIV	17.888 (4.029)	2.516 (1.317)
		FDR	17.887 (4.031)	2.491 (1.316)
		CV	17.923 (4.029)	2.459 (1.312)
		SURE	17.895 (4.030)	2.474 (1.314)
		BAMS	16.024 (3.680)	4.001 (1.562)
		LPM	21.457 (4.058)	2.482 (1.335)
		ABE	**13.996 (3.010)**	3.893 (1.633)
		EBAYES	18.419 (4.022)	**2.423 (1.311)**
	2048	UNIV	16.824 (4.045)	2.063 (1.333)
		FDR	16.730 (4.038)	1.937 (1.346)
		CV	16.551 (4.042)	1.867 (1.346)
		SURE	16.683 (4.037)	1.944 (1.345)
		BAMS	16.941 (4.030)	**1.831 (1.318)**
		LPM	21.758 (4.048)	2.414 (1.349)
		ABE	**16.464 (4.025)**	1.895 (1.357)
		EBAYES	16.483 (4.055)	1.851 (1.350)

Table 3. AMSE and AMAE of the shrinkage and thresholding rules in the simulation study for DJ-test functions under **exponential** noise

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Bumps	32	UNIV	1.528 (0.887)	10.277 (2.265)
		FDR	9.427 (1.634)	19.113 (2.410)
		CV	20.712 (1.798)	28.800 (2.639)
		SURE	2.754 (1.088)	16.869 (2.100)
		BAMS	2.911 (1.175)	5.020 (1.580)
		LPM	**0.357 (0.304)**	**3.398 (0.9072)**
		ABE	0.512 (0.453)	4.756 (1.483)
		EBAYES	1.307 (0.698)	5.203 (1.576)
	2048	UNIV	0.801 (0.546)	5.108 (1.528)
		FDR	0.334 (0.461)	2.638 (1.369)
		CV	0.490 (0.493)	2.397 (1.333)
		SURE	0.270 (0.412)	2.308 (12.441)
		BAMS	0.602 (0.527)	2.225 (1.267)
		LPM	0.369 **(0.297)**	3.313 **(0.895)**
		ABE	**0.251** (0.427)	**2.207** (1.312)
		EBAYES	0.263 (0.398)	2.290 (1.208)
Doppler	32	UNIV	0.910 (0.628)	5.741 (1.443)
		FDR	2.090 (0.685)	10.102 (1.501)
		CV	1.895 (0.722)	6.426 (1.461)
		SURE	0.474 (0.506)	3.509 (1.268)
		BAMS	1.394 (0.736)	**3.016** (1.213
		LPM	**0.353 (0.310)**	3.311 **(0.911)**
		ABE	0.377 (0.462)	3.261 (1.292
		EBAYES	0.521 (0.560)	4.426 (1.476)
	2048	UNIV	0.358 (0.470)	2.464 (1.336)
		FDR	0.232 (0.438)	1.962 (1.295)
		CV	0.228 (0.435)	1.930 (1.283)
		SURE	0.232 (0.405)	2.046 (1.213)
		BAMS	0.300 (0.402)	**1.842** (1.206)
		LPM	0.367 **(0.297)**	3.299 (0.887)
		ABE	**0.210** (0.426)	1.881 (1.280)
		EBAYES	0.234 (0.406)	2.091 (1.225)

Table 3. (Continued) right part

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Blocks	32	UNIV	1.840 (0.881)	10.865 (1.891)
		FDR	15.734 (2.449)	17.479 (2.133)
		CV	14.712 (2.387)	16.872 (2.073)
		SURE	15.734 (2.449)	17.486 (2.130)
		BAMS	3.877 (1.286)	5.816 (1.532)
		LPM	**0.378 (0.307)**	**3.378 (0.939)**
		ABE	0.548 (0.446)	5.821 (1.607)
		EBAYES	12.255 (2.023)	14.001 (2.271)
	2048	UNIV	0.693 (0.491)	4.160 (1.312)
		FDR	0.317 (0.434)	2.469 (1.281)
		CV	0.334 (0.437)	2.255 (1.255)
		SURE	0.264 (0.401)	2.252 (1.198)
		BAMS	0.545 (0.467)	**2.151** (1.230)
		LPM	0.368 **(0.298)**	3.325 **(0.890)**
		ABE	**0.249** (0.427)	2.174 (1.290)
		EBAYES	0.258 (0.397)	2.251 (1.201)
Heavisine	32	UNIV	0.573 (0.486)	2.457 (1.257)
		FDR	0.559 (0.472)	**2.468** (1.242)
		CV	0.609 (0.492)	2.437 (1.237)
		SURE	0.559 (0.476)	2.461 (1.234)
		BAMS	1.148 (0.627)	3.766 (1.479)
		LPM	**0.371 (0.308)**	3.097 **(0.926)**
		ABE	0.432 (0.451)	3.701 (1.581)
		EBAYES	0.375 (0.406)	2.629 (1.176)
	2048	UNIV	0.273 (0.436)	1.941 (1.274)
		FDR	0.216 (0.424)	1.817 (1.278)
		CV	0.212 (0.421)	1.815 (1.277)
		SURE	0.226 (0.404)	1.882 (1.264)
		BAMS	0.244 (0.402)	**1.771** (1.226)
		LPM	0.366 **(0.296)**	3.297 **(0.893)**
		ABE	**0.207** (0.425)	1.842 (1.289)
		EBAYES	0.229 (0.411)	2.030 (1.245)

Table 4. AMSE and AMAE of the shrinkage and thresholding rules in the simulation study for DJ-test functions under **lognormal** noise

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Bumps	32	UNIV	18.345 (2.837)	6.779 (2.065)
		FDR	23.773 (2.884)	12.772 (2.313)
		CV	33.447 (3.109)	23.781 (2.490)
		SURE	24.057 (2.665)	7.610 (1.792)
		BAMS	8.768 (1.669)	4.030 (1.550)
		LPM	**8.217 (1.031)**	**2.014 (1.029)**
		ABE	10.424 (2.144)	2.813 (1.192)
		EBAYES	9.900 (1.950)	2.996 (1.389)
	2048	UNIV	9.993 (2.116)	3.782 (1.347)
		FDR	5.620 (1.707)	1.856 (1.238)
		CV	5.635 (1.730)	1.843 (1.236)
		SURE	6.124 (1.470)	1.688 (1.168)
		BAMS	5.897 (1.524)	1.840 (1.207)
		LPM	8.470 **(1.000)**	2.026 **(1.002)**
		ABE	**5.570** (1.692)	1.715 (1.219)
		EBAYES	6.696 (1.388)	**1.673** (1.147)
Doppler	32	UNIV	10.195 (1.805)	4.253 (1.320)
		FDR	13.568 (1.857)	5.663 (1.340)
		CV	10.681 (1.797)	4.360 (1.322)
		SURE	9.040 (1.575)	2.123 (1.186)
		BAMS	8.783 (1.385)	2.388 (1.178)
		LPM	9.924 **(1.017)**	**2.010** (1.020)
		ABE	**8.156** (1.750)	2.182 (1.205)
		EBAYES	9.610 (1.766)	2.544 (1.316)
	2048	UNIV	4.529 (1.758)	2.014 (1.235)
		FDR	4.384 (1.638)	1.573 (1.205)
		CV	**4.084** (1.699)	1.554 (1.197)
		SURE	5.476 (1.443)	1.586 (1.156)
		BAMS	5.040 (1.497)	**1.521** (1.145)
		LPM	8.122 (1.000)	2.033 **(1.000)**
		ABE	4.480 (1.662)	1.535 (1.201)
		EBAYES	6.132 (1.410)	1.596 (1.158)

Table 4. (Continued) right part

Signal	n	Method	SNR = 3	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)
Blocks	32	UNIV	16.109 (2.103)	8.000 (1.695)
		FDR	19.834 (2.222)	17.067 (2.027)
		CV	19.660 (2.179)	16.252 (1.923)
		SURE	20.165 (2.222)	17.051 (2.026)
		BAMS	10.385 (1.653)	4.897 (1.516)
		LPM	**9.165 (1.025)**	**2.047 (1.006)**
		ABE	12.096 (2.134)	3.382 (1.335)
		EBAYES	17.141 (2.386)	13.769 (2.200)
	2048	UNIV	7.556 (1.688)	3.226 (1.234)
		FDR	5.266 (1.614)	1.796 (1.185)
		CV	**5.242** (1.632)	1.712 (1.177)
		SURE	5.976 (1.437)	1.668 (1.145)
		BAMS	5.704 (1.499)	1.778 (1.171)
		LPM	8.425 **(1.003)**	2.029 **(1.000)**
		ABE	5.380 (1.674)	1.699 (1.202)
		EBAYES	6.555 (1.391)	**1.657** (1.142)
Heavisine	32	UNIV	**5.078** (1.546)	2.052 (1.168)
		FDR	5.916 (1.504)	2.013 (1.159)
		CV	5.820 (1.506)	1.993 (1.163)
		SURE	6.539 (1.471)	2.001 (1.160)
		BAMS	8.032 (1.589)	3.408 (1.375)
		LPM	9.125 **(1.012)**	2.018 **(1.008)**
		ABE	7.543 (1.886)	2.899 (1.386)
		EBAYES	7.578 (1.387)	**1.952** (1.134)
	2048	UNIV	**3.521** (1.684)	1.657 (1.177)
		FDR	4.319 (1.631)	1.512 (1.185)
		CV	3.798 (1.677)	1.508 (1.181)
		SURE	5.582 (1.442)	1.556 (1.152)
		BAMS	5.184 (1.520)	**1.457** (1.153)
		LPM	8.339 **(1.001)**	2.026 **(0.997)**
		ABE	4.670 (1.709)	1.512 (1.192)
		EBAYES	6.255 (1.430)	1.573 (1.162)

Table 5 summarizes the number of scenarios in terms of underlying function, sample size and signal to noise ratio that each method had the best performance, according to uniform, exponential and lognormal noises. In general, we observed that bayesian methods performed better than the classical ones. Under uniform noise, EBAYES was the best for AMSE measure and BAMS for AMAE. LPM and ABE also were the best in some scenarios. In the considered classical methods set, only universal thresholding won some scenarios. Under exponential noise, LPM had great superiority, being the best considering AMSE in six of the sixteen scenarios and also for AMAE, which was the best in all the contexts. This behaviour also occured under lognormal noise, where LPM was the best in all scenarios in terms of AMAE and in five of the sixteen contexts in terms of AMSE. Thus, for AMAE measure, LPM was the best one for all scenarios under exponential and lognormal noises. It should be noted that FDR and SURE methods did not win any scenario for both measures. Moreover, UNIV and CV won only a few number of scenarios, if compared with bayesian methods performances.

Table 5. Number of scenarios that each method was the best in terms of AMSE and AMAE according to noise distribution in the simulation study

	Noise Distribution					
	Uniform		**Exponential**		**Lognormal**	
Method	**AMSE**	**AMAE**	**AMSE**	**AMAE**	**AMSE**	**AMAE**
UNIV	0	2	0	0	2	0
FDR	0	0	0	0	0	0
CV	0	0	1	0	2	0
SURE	0	0	0	0	0	0
BAMS	4	8	4	0	2	0
LPM	3	3	6	16	5	16
ABE	2	2	5	0	2	0
EBAYES	7	1	0	0	3	0

Finally, Figures 4 and 5 show boxplots of the mean squared error (MSE) of the methods in the replications under $n = 2048$ and SNR = 9 for each underlying function and Figures 6 and 7 present the boxplots of the median absolute error (MAE) of them. In general, both the MSEs and MAEs of the methods had small variation in their replications, which means precision in the obtained values. Moreover, it is possible to visualize that there were not meaningful differences in MSEs among the methods in most of the scenarios but LPM was predominant in terms of MAEs.

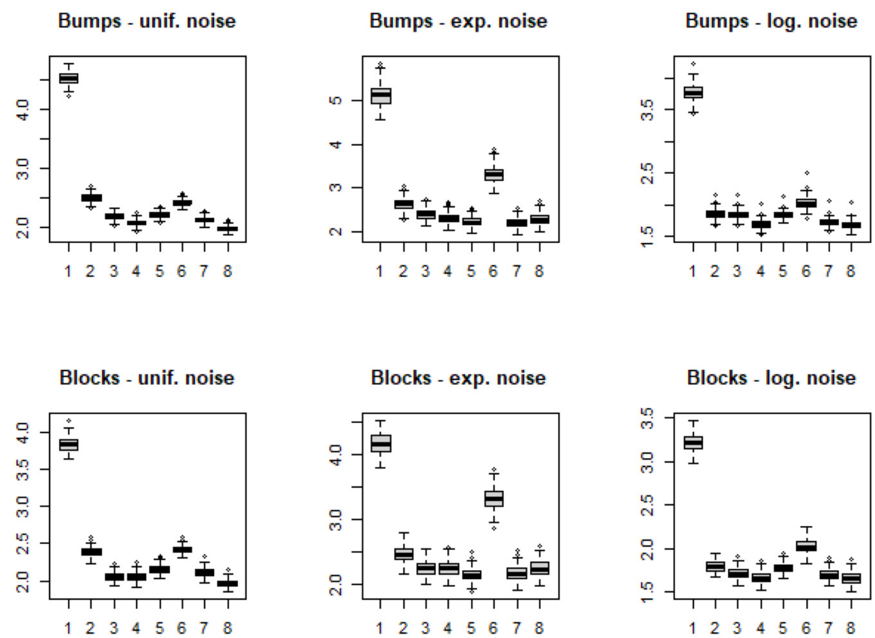

Figure 4. Boxplots of MSE of the methods in simulation studies for Bumps and Blocks functions, $n = 2048$ and SNR = 9. The associated methods are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-ABE, and 8-EBAYES.

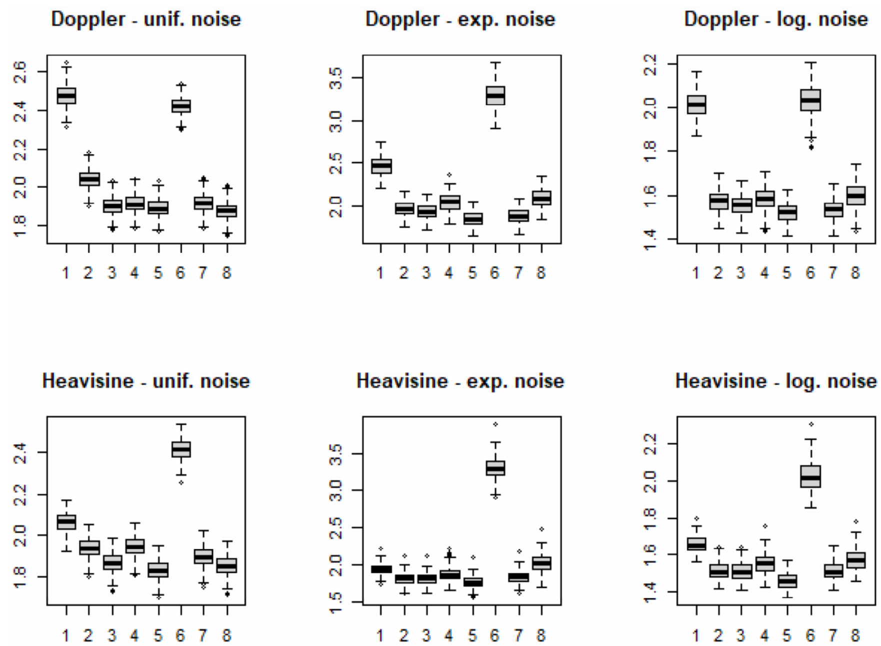

Figure 5. Boxplots of MSE of the methods in simulation studies for Doppler and Heavisine functions, $n = 2048$ and SNR = 9. The associated methods are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-ABE, and 8-EBAYES.

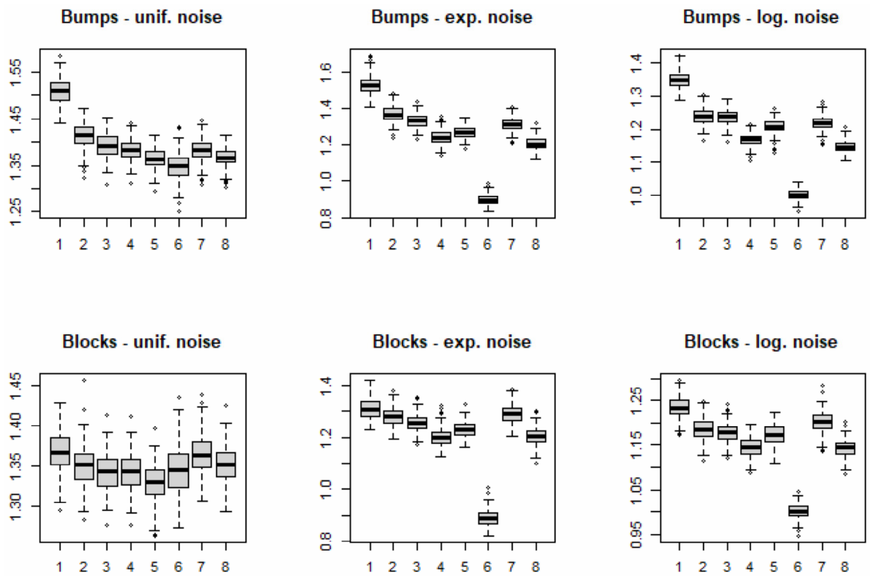

Figure 6. Boxplots of MAE of the methods in simulation studies for Bumps and Blocks functions, $n = 2048$ and SNR = 9. The associated methods are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-ABE, and 8-EBAYES.

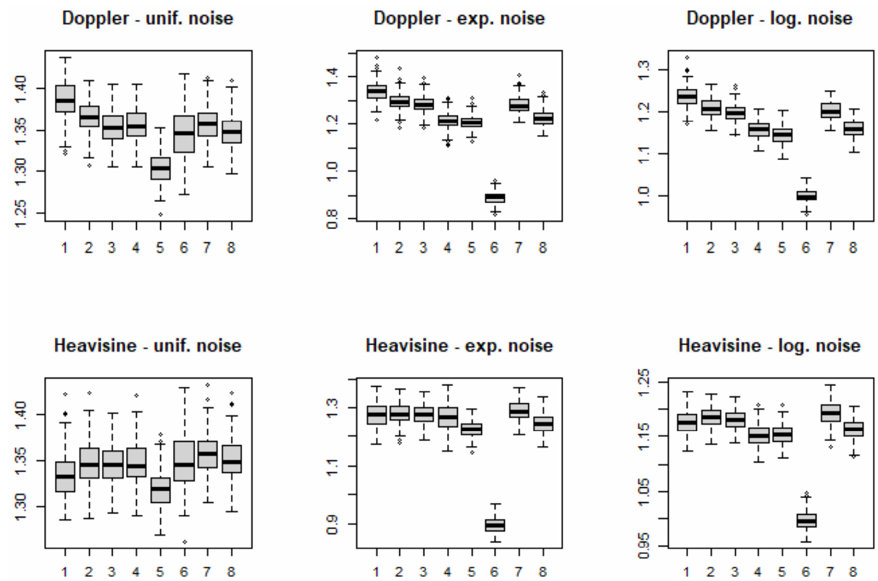

Figure 7. Boxplots of MAE of the methods in simulation studies for Doppler and Heavisine functions, $n = 2048$ and SNR = 9. The associated methods are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-ABE, and 8-EBAYES.

5. Final Considerations

This work proposed the evaluation of the performances of standard wavelet shrinkage and thresholding methods in estimating an unknown function in a nonparametric regression model with strictly positive noise. It was considered the uniform, exponential and lognormal distributions to the random noise in the nonparametric regression model and eight shrinkage and thresholding estimators were considered in the simulation studies.

In general, the bayesian shrinkage estimators performed better in terms of averaged mean squared error (AMSE) and averaged median absolute error (AMAE). Under AMAE measure, Large Posterior Mode (LPM) method had the best results in all the scenarios under exponential and lognormal noises. Moreover, False Discovery Rate (FDR) and SUREshrink (SURE) methods were not the best in any considered scenario. It should be emphasized that the methods were developed under gaussian noise assumption, thus the simulation results suggest some kind of flexibility in terms of noise distribution for the bayesian shrinkage methods, i.e., they can even be applied successfully in denoising data under strictly positive noise.

The performances of the methods under the assumption of other noise distributions with positive support or even a theoretical generalization in this sense, the consideration of other shrinkage and thresholding estimators and the impact of the chosen wavelet basis in the estimation process are some welcome future works in this area.

Acknowledgement

This work was supported by a CAPES[1] fellowship.

References

[1] F. Abramovich and Y. Benjamini, Adaptive thresholding of wavelet coefficients, Computational Statistics & Data Analysis 22(4) (1996), 351-361.

DOI: https://doi.org/10.1016/0167-9473(96)00003-5

[2] A. Antoniadis, D. Leporini and J. C. Pesquet, Wavelet thresholding for some classes of non-Gaussian noise, Statistica Neerlandica 56(4) (2002), 434-453.

DOI: https://doi.org/10.1111/1467-9574.00211

[3] R. Averkamp and C. Houdré, Wavelet thresholding for non-necessarily Gaussian noise: Idealism, The Annals of Statistics 31(1) (2003), 110-151.

DOI: https://doi.org/10.1214/aos/1046294459

[4] L. Cutillo, Y. Y. Jung, F. Ruggeri and B. Vidakovic, Larger posterior mode wavelet thresholding and applications, Journal of Statistical Planning and Inference 138(12) (2008), 3758-3773.

DOI: https://doi.org/10.1016/j.jspi.2007.12.015

---

[1]Coordination of Superior Level Staff Improvement, Brazil.

[5] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

DOI: https://doi.org/10.1137/1.9781611970104

[6] D. L. Donoho, Nonlinear Wavelet Methods of Recovery for Signals, Densities, and Spectra from Indirect and Noisy Data, Proceedings of Symposia in Applied Mathematics, Volume 47, American Mathematical Society, Providence: RI, 1993.

[7] D. L. Donoho, Unconditional bases are optimal bases for data compression and statistical estimation, Applied and Computational Harmonic Analysis 1(1) (1993), 100-115.

DOI: https://doi.org/10.1006/acha.1993.1008

[8] D. L. Donoho, De-noising by soft-thresholding, IEEE Transactions on Information Theory 41(3) (1995), 613-627.

DOI: https://doi.org/10.1109/18.382009

[9] D. L. Donoho, Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition, Applied and Computational Harmonic Analysis 2(2) (1995), 101-126.

DOI: https://doi.org/10.1006/acha.1995.1008

[10] D. L. Donoho and I. M. Johnstone, Ideal denoising in an orthonormal basis chosen from a library of bases, Comptes Rendus de l'Académie des Sciences: Paris A 319 (1994), 1317-1322.

[11] D. L. Donoho and I. M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81(3) (1994), 425-455.

DOI: https://doi.org/10.2307/2337118

[12] D. L. Donoho and I. M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, Journal of the American Statistical Association 90(432) (1995), 1200-1224.

[13] M. A. T. Figueiredo and R. D. Nowak, Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior, IEEE Transactions on Image Processing 10(9) (2001), 1322-1331.

DOI: https://doi.org/10.1109/83.941856

[14] M. Jansen, Noise Reduction by Wavelet Thresholding, Springer, New York, 2001.

[15] I. M. Johnstone and B. W. Silverman, Empirical Bayes selection of wavelet thresholds, The Annals of Statistics 33(4) (2005), 1700-1752.

DOI: https://doi.org/10.1214/009053605000000345

[16] D. Leporini and J. C. Pesquet, Bayesian wavelet denoising: Besov priors and non-Gaussian noises, Signal Processing 81(1) (2001), 55-67.

DOI: https://doi.org/10.1016/S0165-1684(00)00190-0

[17] S. G. Mallat, A Wavelet Tour of Signal Processing, Academic Press, San Diego, 1998.

[18] P. A. de Morettin, Ondas e Ondaletas: Da Análise de Fourier á Análise de Ondaletas de Séries Temporais, 2nd Edition, São Paulo, University of São Paulo Press, 2014.

[19] G. P. Nason, Wavelet shrinkage using cross-validation, Journal of the Royal Statistical Society: Series B (Methodological) 58(2) (1996), 463-479.

DOI: https://doi.org/10.1111/j.2517-6161.1996.tb02094.x

[20] G. P. Nason, Wavelet Methods in Statistics with R, Springer, 2008.

[21] M. H. Neumann and R. von Sachs, Wavelet Thresholding: Beyond the Gaussian I. I. D. Situation, Wavelets and Statistics, Springer, 1995.

[22] A. R. S. dos Sousa and N. L. Garcia, Wavelet Shrinkage in Nonparametric Regression Models with Positive Noise (2021), arXiv:2109.06102.

[23] B. Vidakovic, Statistical Modeling by Wavelets, Wiley, New York, 1999.

DOI: https://doi.org/10.1002/9780470317020

[24] B. Vidakovic and F. Ruggeri, BAMS method: Theory and simulations, Sankhya: The Indian Journal of Statistics: Series B 63(2) (2001), 234-249.

■