

Asymmetric Prior in Wavelet Shrinkage

Priori asimétrico en contracción de ondículas

ALEX RODRIGO DOS SANTOS SOUSA^a

DEPARTMENT OF STATISTICS, INSTITUTE OF MATHEMATICS AND STATISTICS, UNIVERSITY OF SÃO PAULO, SÃO PAULO, BRAZIL

Abstract

In bayesian wavelet shrinkage, the already proposed priors to wavelet coefficients are assumed to be symmetric around zero. Although this assumption is reasonable in many applications, it is not general. The present paper proposes the use of an asymmetric shrinkage rule based on the discrete mixture of a point mass function at zero and an asymmetric beta distribution as prior to the wavelet coefficients in a non-parametric regression model. Statistical properties such as bias, variance, classical and bayesian risks of the associated asymmetric rule are provided and performances of the proposed rule are obtained in simulation studies involving artificial asymmetric distributed coefficients and the Donoho-Johnstone test functions. Application in a seismic real dataset is also analyzed.

Key words: asymmetric beta distribution; nonparametric regression; wavelet shrinkage.

Resumen

En la contracción de las ondículas bayesianas, se supone que los coeficientes a priori ya propuestos de las ondículas son simétricos alrededor de cero. Aunque esta suposición es razonable en muchas aplicaciones, no es general. El presente artículo propone el uso de una regla de contracción asimétrica basada en la mezcla discreta de una función de masa puntual de ondícula en un modelo de regresión no paramétrico. Se proporcionan propiedades estadísticas tales como sesgo, varianza, riesgos clásicos y bayesianos de la regla asimétrica asociada y se obtienen los rendimientos de la regla propuesta en estudios de simulación que involucran coeficientes distribuidos asimétricos artificiales y las funciones de prueba de Donoho-Johnstone. También se analiza la aplicación en un conjunto de datos sísmicos reales.

Palabras clave: contracción de las ondículas; distribución beta asimétrica; regresión no paramétrico.

^aPh.D. E-mail: alex.sousa89@usp.br

1. Introduction

Wavelet-based methods have been extensively studied and applied in several areas, such as mathematics, signal and image processing, geophysics, genomics and many others. In statistics, applications of wavelets arise mainly in the areas of non-parametric regression, density estimation, functional data analysis and time series analysis. In non-parametric regression, the focus of this work, an unknown function is expanded as linear combination of wavelet basis and the coefficients of this representation are estimated. The use of wavelets representation is attractive in non-parametric regression due their well localized and sparse wavelet coefficients, i.e., the coefficients are typically non-zero or significant on positions where the function has important characteristics to be recovered, as peaks, cusps and discontinuities for example and are zero or very close to zero on smooth regions of the function. These features of wavelets provide computational and analytical advantages. More about wavelet methods in statistics can be seen in [Vidakovic \(1999\)](#).

Due the sparsity property of wavelet coefficients, shrinkage and thresholding methods are generally used to estimate them in the wavelet domain by reducing the magnitude of the observed (empirical) coefficients obtained by application of a discrete wavelet transformation on the original data. There are in fact several shrinkage and thresholding techniques available in the literature. The main works in this area are of [Donoho \(1995a; 1995b\)](#), [Donoho & Johnstone \(1994a; 1994b; 1995\)](#), but also [Donoho et al. \(1995; 1996\)](#), [Vidakovic \(1998\)](#), [Antoniadis et al. \(2001\)](#) and [Johnstone & Silverman \(2005\)](#) can be cited. For more details of shrinkage methods, see [Vidakovic \(1999\)](#) and [Jansen \(2001\)](#).

Bayesian shrinkage methods have also been studied, mainly for the possibility of adding, by means of a prior probabilistic distribution, prior information about the regression, coefficients and other parameters to be estimated. Specifically in the case of wavelets, information about the degree of sparsity of the coefficient vector, the support of these coefficients (if they are limited), among others can be incorporated into the statistical model of study by means of bayesian procedures. In this sense, the choice of the prior distribution of the wavelet coefficients is extremely important to achieve meaningful results. Several bayesian shrinkage procedures have been proposed in the last years in many statistical fields. Some of them are found in [Lian \(2011\)](#), [Beenamol et al. \(2012\)](#), [Karagiannis et al. \(2015\)](#), [Griffin & Brown \(2017\)](#) and [Torkamani & Sadeghzadeh \(2017\)](#). Further, priors models in the wavelet domain were proposed since 1990s, as for example a mixture of gaussian distributions by [Chipman et al. \(1997\)](#), mixtures of a point mass function at zero and a symmetric distribution were considered by [Abramovich et al. \(1998\)](#), [Vidakovic \(1998\)](#) with the use of the t -distribution as the symmetric density in the mixture, [Vidakovic & Ruggeri \(2001\)](#) with the double exponential distribution, [Angelini & Vidakovic \(2004\)](#) with a Γ -Minimax shrinkage rule based on the Bickel distribution, Weibull prior were proposed by [Reményi & Vidakovic \(2015\)](#), Dirichlet-Laplace priors by [Bhattacharya et al. \(2015\)](#), the logistic prior by [Sousa et al. \(2021\)](#) and the symmetric beta distribution by [Sousa et al. \(2021\)](#), among others. The common

use of the mixture prior involving the point mass function at zero is suitable for wavelet coefficients distribution modelling due the sparsity feature of them, i.e, most of the wavelet coefficients are typically zero in practice. In this sense, the point mass function allows to incorporate prior information about sparsity of the wavelet coefficients by the determination of its weight in the prior mixture distribution model.

One feature of the priors already proposed to the wavelet coefficients is that they are symmetric around zero. Although these priors have been well succeeded in many real applications, this symmetry assumption is not a general case, i.e, wavelet coefficients can be asymmetrically distributed. In this sense, the proposition of an asymmetric prior distribution could be welcome for better estimation of asymmetrically distributed wavelet coefficients. Moreover, little attention has been given to bounded priors, which can be important to model bounded energy signals denoising, restricted to the proposition of the uniform and Bickel distributions by [Angelini & Vidakovic \(2004\)](#) and the symmetric beta prior by [Sousa et al. \(2021\)](#), although bounded energy signals occur in practice. Motivated by these reasons, we propose in this work an asymmetric prior distribution to wavelet coefficients based on a discrete mixture of a point mass function at zero and the beta distribution with support on $[-m, m]$. The novelty of this study is, therefore, the application of an asymmetric shrinkage rule associated with this prior model to estimate asymmetrically distributed wavelet coefficients of the unknown function to be recovered in a non-parametric regression model. Thus, the present work can be viewed as an important extension of [Sousa \(2020\)](#), who proposed the use of beta prior only in the symmetric case, restricting its hyperparameters values. As we will see in the simulation studies, the proposed asymmetric shrinkage rule outperforms the symmetric ones when the wavelet coefficients are asymmetric, as expected.

The use of the asymmetric beta prior is interesting by several reasons. First of all, it has a well known shape flexibility obtained by convenient choices of its hyperparameters. Further, its hyperparameters have direct and easy interpretations in terms of asymmetry (left and right asymmetry choices) and shrinkage level, which are very useful for their elicitation by practitioners. Finally, the shrinkage rule under asymmetric beta prior outperformed, in terms of averaged mean squared and absolute errors, the considered shrinkage/thresholding methods in our simulation studies, mainly when the coefficients present high asymmetry.

This paper is organized as follows: Section 2 defines the model and the proposed asymmetric beta prior, Section 3 establishes the associated shrinkage rule, shows statistical properties of the rule, such as variance, bias and risks. Parameter and hyperparameter elicitation are discussed in Section 4. Simulation studies involving artificial asymmetric coefficients and the so called Donoho-Jonhstone test functions to evaluate performances are done in Section 5 and application of the proposed shrinkage rule in a real seismic dataset is done in Section 6. The paper finishes with conclusion and final considerations in Section 7.

2. Statistical Model

Let us consider the non-parametric regression problem involving one dimensional values x_i and y_i of the form

$$y_i = f(x_i) + e_i, \quad i = 1, \dots, n = 2^J, J \in \mathbb{N}, \quad (1)$$

where $x_i \in \mathbb{R}$, $i = 1, \dots, n$, $f \in \mathbb{L}_2(\mathbb{R}) = \{f_0 : \mathbb{R} \rightarrow \mathbb{R} \mid \int f_0^2 < \infty\}$, and e_i , $i = 1, \dots, n$, are zero mean independent normal random variables with unknown variance σ^2 .

The unknown function f can be represented by

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

where $\{\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), j, k \in \mathbb{Z}\}$ is an orthonormal wavelet basis for $\mathbb{L}_2(\mathbb{R})$ 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 . In this context, the data points $(x_1, y_1), \dots, (x_n, y_n)$ can be viewed as an approximation of f at the finest resolution level J with additive and positive noise contamination.

In vector notation, we can rewrite model (1) as

$$\mathbf{y} = \mathbf{f} + \mathbf{e}, \quad (2)$$

where $\mathbf{y} = (y_1, \dots, y_n)'$, $\mathbf{f} = (f(x_1), \dots, f(x_n))'$ and $\mathbf{e} = (e_1, \dots, e_n)'$. To estimate the unknown function f , the standard procedure is to apply a discrete wavelet transform (DWT) on (2), represented by an orthogonal transformation matrix \mathbf{W} of dimension $n \times n$ determined according to the considered wavelet basis, to obtain the following model in the wavelet domain,

$$\mathbf{d} = \boldsymbol{\theta} + \boldsymbol{\epsilon}, \quad (3)$$

where $\mathbf{d} = \mathbf{W}\mathbf{y}$, $\boldsymbol{\theta} = \mathbf{W}\mathbf{f}$ and $\boldsymbol{\epsilon} = \mathbf{W}\mathbf{e}$. For a specific component of the vector \mathbf{d} , we have the simple model $d_k = \theta_k + \epsilon_k$ or, for simplicity,

$$d = \theta + \epsilon, \quad (4)$$

where d is the empirical wavelet coefficient, $\theta \in \mathbb{R}$ is the wavelet coefficient to be estimated and since the orthogonality of wavelet transformation preserves the stochastic structure of gaussian noise, $\epsilon \sim N(0, \sigma^2)$ is normal random error with unknown variance σ^2 . Although \mathbf{W} is used as DWT representation for pedagogical purposes, fast algorithms such as the pyramidal algorithm, which has complexity $\mathcal{O}(n)$ and is a sequence of low and high pass filters, are applied to perform DWT in practice, see Mallat (1998) for more details about DWT. Note that, according to the model (4), $d \mid \theta \sim N(\theta, \sigma^2)$ and then, the problem of estimating a function f becomes a normal location parameter estimation problem in the wavelet domain for each coefficient, with posterior estimation of f by the inverse discrete wavelet transform (IDWT), i.e., $\hat{\mathbf{f}} = \mathbf{W}^T \hat{\boldsymbol{\theta}}$.

One of the main advantages of expanding an unknown function in wavelet basis is the typical sparsity of the vector of wavelet coefficients θ . In fact, the coefficients are nonzero or significant only in localizations of time domain where the function has features to be recovered, as discontinuities or peaks for example and most of the remaining coefficients are zero or very close to zero. In this sense, a good estimator of θ should take this sparsity feature into account. Most of the classical procedures usually apply some kind of thresholding policy on the empirical coefficients \mathbf{d} , i.e, the empirical coefficient d are shrunk to zero if it is less than some threshold value λ , $\lambda > 0$.

For bayesian estimation of θ , it is possible to model the sparsity of θ by proposing a discrete mixture of a point mass function at zero and a continuous distribution, attributing most of the weight on the point mass at zero. We propose in this work the following prior discrete mixture distribution for θ ,

$$\pi(\theta; \alpha, a, b, m) = \alpha \delta_0(\theta) + (1 - \alpha)g(\theta; a, b, m), \quad (5)$$

where $\alpha \in (0, 1)$, $\delta_0(\cdot)$ is the point mass function at zero and $g(\cdot; a, b, m)$ is the beta density function on $[-m, m]$, which is given by

$$g(\theta; a, b, m) = \frac{(\theta + m)^{a-1}(m - \theta)^{b-1}}{(2m)^{a+b-1}B(a, b)} \mathbb{I}_{[-m, m]}(\theta), \quad (6)$$

for $a, b, m > 0$, $a \neq b$ (asymmetric case), $B(\cdot, \cdot)$ the standard beta function and $\mathbb{I}_A(\cdot)$ the usual indicator function on the set A . In fact, the proposed beta density on $[-m, m]$ can be obtained by the transformation $X = 2mY - m$, where Y is a random variable with standard beta distribution on $[0, 1]$. Thus the proposed prior distribution to the wavelet coefficients has α, a, b and m as hyperparameters to be elicited. We will see later that α, a and b values impact directly on the shrinkage level of the estimator. Sousa (2020) proposed the prior model (5) and (6), but restricted to symmetric case $a = b$. A performance comparison between shrinkage rules under symmetric and asymmetric cases are provided in the simulation studies, in Section 5.

According to the prior model given by Equations (5) and (6), we have that the prior expected value $\mathbb{E}_\pi(\theta)$ and variance $Var_\pi(\theta)$ of θ are given respectively by

$$\mathbb{E}_\pi(\theta) = \frac{m(1 - \alpha)(a - b)}{a + b},$$

$$Var_\pi(\theta) = \frac{(1 - \alpha)m^2}{(a + b)^2} \left[\frac{4ab}{a + b + 1} + \alpha(a + b)^2 \right].$$

Moreover, the Pearson skewness coefficient of a random variable X with distribution (6) and the prior Pearson skewness coefficient of θ , which are their third standardized moments, are given respectively by

$$Skew_g(X) = \frac{2(b - a)\sqrt{a + b + 1}}{(a + b + 2)\sqrt{ab}},$$

$$Skew_\pi(\theta) = -\alpha \left[\frac{\mathbb{E}_\pi(\theta)}{SD_\pi(\theta)} \right]^3 + (1 - \alpha) \frac{2(b - a)\sqrt{a + b + 1}}{(a + b + 2)\sqrt{ab}},$$

where $SD_\pi(\theta)$ is the prior standard deviation of θ . First of all, we can note that the proposed prior does not have zero mean wavelet coefficients. Actually, it only occurs when $a = b$, the symmetric context, which is not considered on this work. Further, the Pearson's skewness coefficient tells us that when $a > b$, $Skew_\pi(\theta) < 0$ and the prior is left asymmetric. When $a < b$, $Skew_\pi(\theta) > 0$ and the prior is right asymmetric. Moreover, the asymmetry increases as $|a - b|$ increases. The symmetry occurs when $a = b$ and then, for this work, we consider only the cases $a \neq b$, i.e., the asymmetric setup. Figure 1 shows some beta density functions for $a = 4$, $b \in \{1, 2, 3, 4\}$ and $b = 4$, $a \in \{1, 2, 3\}$, $m = 3$. When $a > b$, beta density is left asymmetric and when $a < b$, it is right asymmetric. For $a = b$, beta density is symmetric around zero. In this work we considered the asymmetric case. Further, note that there is symmetry around $x = 0$ between densities of interchangeable values of parameters a and b . We will see in the next section the impact on the shrinkage rules and their statistical properties according to the skewness of the prior distribution of θ .

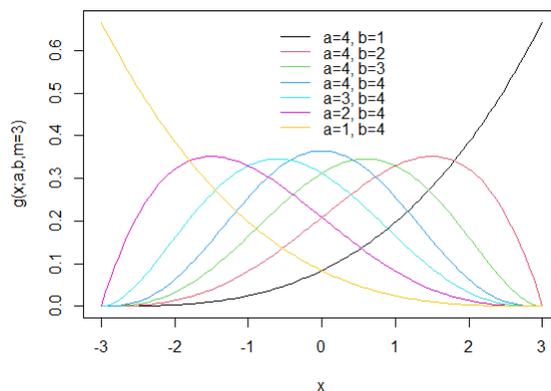


FIGURE 1: Beta density functions (6) for $a = 4$, $b \in \{1, 2, 3, 4\}$ and $b = 4$, $a \in \{1, 2, 3\}$, $m = 3$. When $a > b$, beta density is left asymmetric and when $a < b$, it is right asymmetric. For $a = b$, beta density is symmetric around zero.

3. Statistical Description of the Shrinkage Rule

The model (4), (5) and (6) allows us to obtain the bayesian shrinkage rule $\delta(\cdot)$, which is the Bayes estimator of θ based on the empirical wavelet coefficient d . It is well known that under the quadratic loss function $L(\delta, \theta) = (\delta - \theta)^2$, the Bayes estimator is the posterior expected value of θ , i.e., $\delta(d) = \mathbb{E}_\pi(\theta|d)$. Proposition 1 by Sousa et al. (2021) gives the specific expression of the shrinkage rule under model (4), a general prior distribution of the form $\pi(\theta; \alpha, m, \boldsymbol{\tau}) = \alpha\delta_0(\theta) + (1 - \alpha)g(\theta; \boldsymbol{\tau})$ and for a density function g with support in $[-m, m]$.

Proposition 1. Consider the location parameter θ estimation problem (4). If the prior distribution of θ is of the form $\pi(\theta; \alpha, m, \boldsymbol{\tau}) = \alpha\delta_0(\theta) + (1 - \alpha)g(\theta; \boldsymbol{\tau})$, where g is a density function with support in $[-m, m]$ and parameters $\boldsymbol{\tau}$, then the

shrinkage rule under the quadratic loss function is given by

$$\delta(d) = \frac{(1 - \alpha) \int_{\frac{-m-d}{\sigma}}^{\frac{m-d}{\sigma}} (\sigma u + d) g(\sigma u + d; \boldsymbol{\tau}) \phi(u) du}{\alpha \frac{1}{\sigma} \phi\left(\frac{d}{\sigma}\right) + (1 - \alpha) \int_{\frac{-m-d}{\sigma}}^{\frac{m-d}{\sigma}} g(\sigma u + d; \boldsymbol{\tau}) \phi(u) du} \quad (7)$$

where $\phi(\cdot)$ is the standard normal density function.

Thus we can apply the Proposition 1 to the specific beta density function (6) with $\boldsymbol{\tau} = (a, b)'$ to obtain the shrinkage rule of the proposed model numerically using Monte Carlo methods to calculate integrals in (7). Figure 2 presents the shrinkage rules for $\alpha = 0.9$, $m = 3$, $\sigma = 1$ and (a) left asymmetric case $a > b$, for $a = 7$ and $b \in \{1, 2, 3, 4, 5, 6\}$ and (b) right asymmetric case $a < b$, for $b = 7$ and $a \in \{1, 2, 3, 4, 5, 6\}$. The symmetric shrinkage rule for $a = b = 7$ is also included in both figures for comparison. First of all, we can observe a symmetric behavior of the rules relative to the origin of the cartesian system for interchangeable choices of a and b , i.e., if $\delta_{a,b}(d) = \delta(d)$ for choices of hyperparameters a and b , then $\delta_{a,b}(d) = -\delta_{b,a}(-d)$. For example, $\delta_{7,1}(d) = -\delta_{1,7}(-d)$. This feature is a consequence of the symmetry around zero between the respective densities. Thus, we describe the shrinkage rule and its features just for the left asymmetric case ($a > b$), once the same properties occur symmetrically for the right asymmetric distributional context.

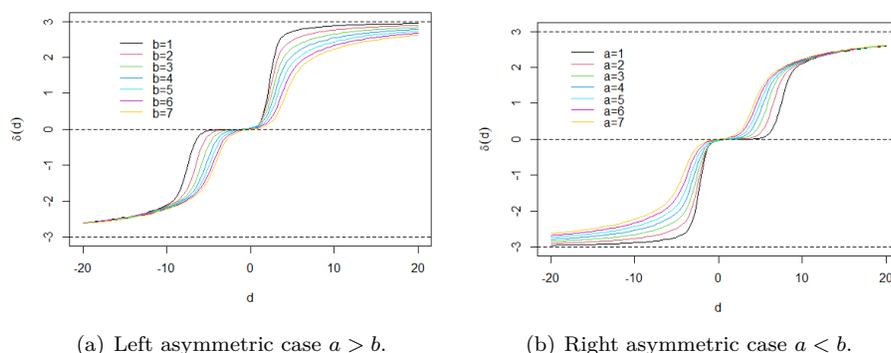
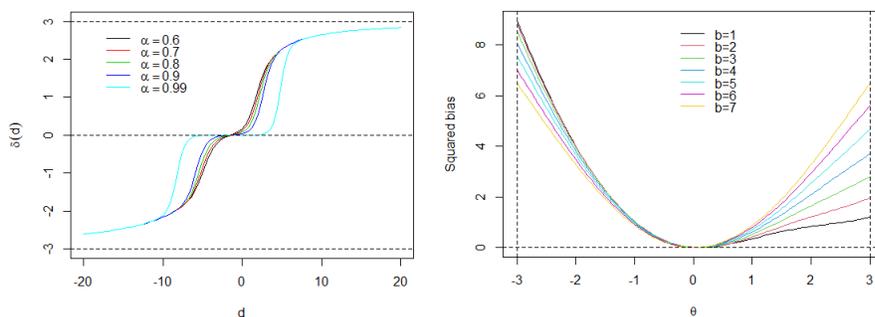


FIGURE 2: Shrinkage rules under beta prior for $\alpha = 0.9$, $m = 3$, $\sigma = 1$ and (a) left asymmetric case ($a > b$), $a = 7$ and (b) right asymmetric case ($a < b$), $b = 7$. The symmetric case $a = b = 7$ is included in both figures for comparison.

Figure 2 (a) shows, as expected, that the shrinkage rules under left asymmetric beta prior perform shrinkage asymmetrically around zero. In fact, negative empirical coefficients are shrunk more than the positive ones. Moreover, the shrinkage level increases as $|a - b|$ increases, once the interval size of d -values that are shrunk to zero is higher. Another property usually taken by bayesian shrinkage rules for bounded wavelet coefficients is that they are also bounded by $[-m, m]$. Since θ is bounded by $[-m, m]$, empirical coefficients d occur with absolute values

higuer than m due noise effect, then they are shrunk to m at most. Figure 3 (a) presents the impact of the hyperparameter α on shrinkage level of the rules, for $\alpha \in \{.6, .7, .8, .9, .99\}$, $a = 7$, $b = 3$ and $m = 3$. As expected, the rule shrinks more as α increases, since this set more weight to the point mass at zero function in the prior model (5). Squared bias, $Bias^2(\theta) = \{\mathbb{E}[\delta(d)|\theta] - \theta\}^2$, and variance



(a) Shrinkage rules for several α values. (b) Squared bias for several b values.

FIGURE 3: Shrinkage rules under beta prior for $a = 7$, $b = 3$ and $m = 3$ for several values of α (a) and squared bias for shrinkage rules for $\alpha = 0.9$, $m = 3$, $\sigma = 1$ and left asymmetric case ($a > b$) with $a = 7$ (b). The symmetric case $a = b = 7$ is included in (b) for comparison.

of the shrinkage rules for left asymmetric case ($a > b$) are provided in Figures 3 (b) and 4 (a) respectively. The estimators are practically unbiased and achieve minimum variance when θ is close to zero (but not for $\theta = 0$). These features also exist when symmetric priors are assumed for θ . However, these properties behave differently for negative and positive θ values. The bias increases faster for negative θ values than for positive ones while the variance increases faster for positive values. Since the shrinkage is stronger for negative values of d , it is reasonable the asymmetrical increase of the bias on the negative values direction with the simultaneously decreasing variance toward it. Figure 4 (b) and Table 1 show, respectively, the classical risks $R_\delta(\theta)$ and Bayes risks r_δ respectively for the same rules considered on the plots of squared bias and variance. In fact, the behavior of the classical risks is the same as the squared bias one, i.e, there is a faster increase of the risk for negative θ values than for positive values, with the minimum risk close to zero. Moreover, we observe that the Bayes risk decreases as the hyperparameter b increases and goes to the symmetric case. Further, Bayes risks for several α values are presented in Table 2. As expected, the bayesian risk decreases as α increases, once this last one implies a higher shrinkage and agree with the prior belief of sparsity of the θ vector.

TABLE 1: Bayes risks of the shrinkage rules under beta prior distribution with hyperparameters $\alpha = 0.9$, $m = 3$, $\sigma = 1$ and $a = 7$.

b	1	2	3	4	5	6
r_δ	0.221	0.182	0.139	0.103	0.078	0.063

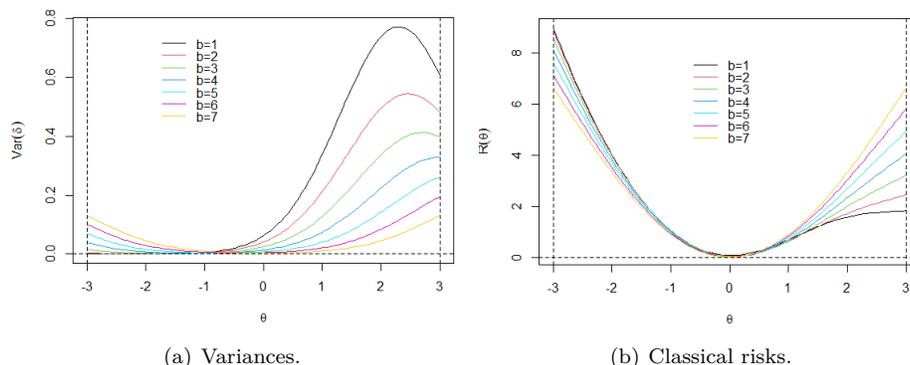


FIGURE 4: Variances and classical risks for shrinkage rules under beta prior for $\alpha = 0.9$, $m = 3$, $\sigma = 1$ and left asymmetric case ($a > b$) with $a = 7$. The symmetric case $a = b = 7$ is included in both figures for comparison.

TABLE 2: Bayes risks of the shrinkage rules under beta prior distribution with hyperparameters $m = 3$, $\sigma = 1$, $a = 7$ and $b = 3$.

α	0.6	0.7	0.8	0.9	0.99
r_δ	0.352	0.299	0.231	0.139	0.019

4. Prior Elicitation

Methods and criteria for determination of the involved parameters and hyperparameters to estimate the coefficients are important in bayesian procedures. In the framework of model (4), (5) and (6), the choices of the σ parameter of the normal random error distribution and the hyperparameters α , m , a and b of the beta prior distribution of the wavelet coefficient are required. We present the methods and criteria already available in the literature for such choices and used in simulation and application studies and some direction on elicitation of the beta shape parameters.

Based on the fact that much of the noise information present in the data can be obtained on the finer resolution scale, for the robust σ estimation, Donoho & Johnstone (1994a) suggest

$$\hat{\sigma} = \frac{\text{median}\{|d_{J-1,k}| : k = 0, \dots, 2^{J-1}\}}{0.6745}. \tag{8}$$

The hyperparameters α and m are the weight of the point mass function at zero of the proposed prior and the upper value of the beta support respectively. Angelini & Vidakovic (2004) suggest the hyperparameters α and m be dependent on the level of resolution j according to the expressions

$$\alpha = \alpha(j) = 1 - \frac{1}{(j - J_0 + 1)\gamma} \tag{9}$$

and

$$m = m(j) = \max_k \{|d_{jk}|\}, \quad (10)$$

where $J_0 \leq j \leq J - 1$, J_0 is the primary resolution level and $\gamma > 0$. They also suggest that in the absence of additional information, $\gamma = 2$ can be adopted.

Finally, the shape hyperparameters a and b should be chosen according to asymmetry and shrinkage levels criteria. Left asymmetry imposes $a > b$ and right one says $a < b$. As $|a - b|$ increases, the asymmetry level and shrinkage in the direction of this asymmetry increase. In practice, assuming symmetric around zero noise, as the considered model (4), the observed asymmetry of the empirical coefficients d can provide some information about wavelet coefficients asymmetry criteria and be a starting point to elicit it. Based on our experiments, we suggest $a, b \in \{1, 2, 3, 4, 5, 6, 7\}$ as default possible values.

5. Simulation Studies

Two simulation studies were done to evaluate the performance of the proposed shrinkage rule and to compare it with well known shrinkage/thresholding methods. The first one (simulation study 1) had empirical coefficients vector \mathbf{d} artificially generated according to the models (4), (5) and (6) and the second one (simulation study 2) involved Donoho-Johnstone test functions, which are usually applied in the literature to compare wavelet-based methods.

In both simulation studies, the performances of our proposed shrinkage rule (denoted by ASYM BETA in Tables and Figures) were compared with soft thresholding with threshold parameter chosen according to the following policies: universal thresholding (UNIV) proposed by Donoho & Johnstone (1994b), false discovery rate (FDR) proposed by Abramovich & Benjamini (1996), cross validation (CV) of Nason (1996) and Stein unbiased risk estimator (SURE) of Donoho & Johnstone (1995). Moreover, we also compared with bayesian shrinkage methods: bayesian adaptive multiresolution shrinker (BAMS) of Vidakovic & Ruggeri (2001), large posterior mode (LPM) of Cutillo et al. (2008) and a symmetric beta shrinkage rule (SYM BETA) proposed by Sousa et al. (2021).

We used the mean squared error (MSE), $MSE = \frac{1}{n} \sum_{i=1}^n [\hat{f}(x_i) - f(x_i)]^2$ and the mean absolute error (MAE) $MAE = \frac{1}{n} \sum_{i=1}^n |\hat{f}(x_i) - f(x_i)|$ as performance measures of the shrinkage rules on each run of the simulation. For each function, the simulation was repeated M times and the comparison measures, the average of the obtained MSEs and MAEs, $AMSE = \frac{1}{M} \sum_{j=1}^M MSE_j$ and $AMAE = \frac{1}{M} \sum_{j=1}^M MAE_j$, were respectively calculated. Thus, the best method in terms of averaged mean squared error and averaged mean absolute error is the one with the smallest values of AMSE and AMAE respectively.

In both simulation studies, normal random noise vectors were generated according to three signal to noise ratio values (SNR), 3, 6 and 9, two sample sizes were considered, $n = 512$ and 2048 and wavelet basis Daubechies with eight null moments (Daub8) was applied.

5.1. Simulation Study 1

To evaluate the performance of the proposed shrinkage rule in asymmetric distributed wavelet coefficients, we generated the wavelet coefficients θ according to the models (5) and (6) in two sceneries of asymmetry and one symmetric scenery controlled by the choice of a and b . The symmetric scenery was considered to evaluate the performance of the asymmetric shrinkage rule in a symmetric distributed coefficients. The hyperparameters choices for the first asymmetric scenery were $a = 3$, $b = 7$, $m = 10$ and $\alpha = 0.9$, which give $\mathbb{E}_\pi(\theta) = -0.40$, $Var_\pi(\theta) = 2.20$ and $Skew_\pi(\theta) = 0.07$ ($Skew_g(X) = 0.48$). Thus, this scenery has a moderate asymmetry and a large amount of null coefficients due the impact of the weight α of the point mass function at zero. On the second scenery one had more extreme hyperparameters choices, $a = 1$, $b = 20$, $m = 30$ and $\alpha = 0.6$, which give $\mathbb{E}_\pi(\theta) = 10.86$, $Var_\pi(\theta) = 179.78$ and $Skew_\pi(\theta) = 0.16$ ($Skew_g(X) = 1.73$). This last one introduces more asymmetry and less null coefficients. Finally, for the symmetric scenery, we considered $a = 5$, $b = 5$, $m = 30$ and $\alpha = 0.9$, corresponding to $\mathbb{E}_\pi(\theta) = 0$, $Var_\pi(\theta) = 89.18$ and $Skew_\pi(\theta) = 0$ ($Skew_g(X) = 0$).

Tables 3, 4 and 5 show the AMSEs and AMAEs for the first, second and third sceneries respectively obtained for $M = 1000$ simulation runs in each scenery of sample size and SNR. When performing this simulation study, sceneries with the right asymmetric case $a < b$ were originally considered, but as mentioned in Section 3, the dynamic of the proposed asymmetric shrinkage rule is the same as for $a > b$, thus we just analyze in this first simulation study the left asymmetry context for simplicity.

TABLE 3: AMSE and AMAE of the shrinkage/thresholding rules in the simulation study 1 for the empirical wavelet coefficients vector artificially generated according to the models (4), (5) and (6) for $\alpha = 0.9$, $a = 3$, $b = 7$ and $m = 10$ -Scenery 1.

n	Method	SNR = 3	SNR = 6	SNR = 9
		AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
512	UNIV	0.439 (0.367)	0.119 (0.191)	0.054 (0.129)
	FDR	0.445 (0.362)	0.129 (0.202)	0.067 (0.144)
	CV	0.634 (0.429)	0.177 (0.226)	0.082 (0.153)
	SURE	0.122 (0.204)	0.031 (0.103)	0.014 (0.069)
	BAMS	0.402 (0.381)	0.375 (0.373)	0.371 (0.373)
	LPM	0.288 (0.391)	0.072 (0.196)	0.032 (0.130)
	SYM BETA	0.041 (0.132)	0.060 (0.063)	0.005 (0.042)
	ASYM BETA	0.035 (0.120)	0.012 (0.056)	0.004 (0.035)
2048	UNIV	0.310 (0.356)	0.087 (0.186)	0.040 (0.126)
	FDR	0.237 (0.303)	0.065 (0.158)	0.030 (0.107)
	CV	0.404 (0.407)	0.116 (0.217)	0.054 (0.147)
	SURE	0.078 (0.182)	0.020 (0.092)	0.009 (0.062)
	BAMS	0.237 (0.327)	0.218 (0.317)	0.215 (0.315)
	LPM	0.240 (0.384)	0.060 (0.192)	0.026 (0.128)
	SYM BETA	0.040 (0.124)	0.124 (0.059)	0.010 (0.040)
	ASYM BETA	0.037 (0.111)	0.008 (0.051)	0.003 (0.036)

TABLE 4: AMSE and AMAE of the shrinkage/thresholding rules in the simulation study 1 for the empirical wavelet coefficients vector artificially generated according to the models (4), (5) and (6) for $\alpha = 0.6$, $a = 1$, $b = 20$ and $m = 30$ -Scenery 2.

n	Method	SNR = 3	SNR = 6	SNR = 9
		AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
512	UNIV	8.316 (1.780)	2.082 (0.888)	0.922 (0.593)
	FDR	14.350 (2.879)	27.230 (3.152)	26.385 (3.092)
	CV	11.456 (2.076)	2.854 (1.039)	1.261 (0.694)
	SURE	2.345 (1.022)	0.589 (0.511)	0.260 (0.340)
	BAMS	2.097 (1.064)	0.401 (0.414)	0.256 (0.312)
	LPM	6.933 (2.116)	1.733 (1.059)	0.770 (0.706)
	SYM BETA	3.971 (1.426)	1.836 (0.901)	1.651 (0.819)
	ASYM BETA	0.307 (0.269)	0.111 (0.162)	0.057 (0.115)
2048	UNIV	12.204 (2.219)	3.058 (1.109)	0.922 (0.740)
	FDR	9.333 (1.996)	2.838 (1.280)	26.385 (1.010)
	CV	17.047 (2.607)	4.279 (1.304)	1.261 (0.870)
	SURE	2.680 (1.051)	0.672 (0.526)	0.260 (0.350)
	BAMS	2.425 (1.075)	0.449 (0.439)	0.256 (0.340)
	LPM	7.464 (2.108)	1.867 (1.055)	0.770 (0.703)
	SYM BETA	3.790 (1.425)	1.706 (0.882)	1.419 (0.764)
	ASYM BETA	0.351 (0.330)	0.120 (0.198)	0.057 (0.141)

TABLE 5: AMSE and AMAE of the shrinkage/thresholding rules in the simulation study 1 for the empirical wavelet coefficients vector artificially generated according to the models (4), (5) and (6) for $\alpha = 0.9$, $a = 5$, $b = 5$ and $m = 30$ -Scenery 3.

n	Method	SNR = 3	SNR = 6	SNR = 9
		AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
512	UNIV	0.089 (0.188)	0.024 (0.099)	0.011 (0.067)
	FDR	0.089 (0.187)	0.027 (0.103)	0.012 (0.071)
	CV	0.115 (0.213)	0.031 (0.112)	0.014 (0.077)
	SURE	0.028 (0.113)	0.007 (0.058)	0.003 (0.039)
	BAMS	0.335 (0.359)	0.334 (0.358)	0.334 (0.357)
	LPM	0.093 (0.244)	0.023 (0.121)	0.010 (0.081)
	SYM BETA	0.012 (0.065)	0.002 (0.030)	0.001 (0.019)
	ASYM BETA	0.018 (0.072)	0.008 (0.043)	0.007 (0.035)
2048	UNIV	0.163 (0.265)	0.0468 (0.143)	0.021 (0.097)
	FDR	0.126 (0.232)	0.035 (0.123)	0.016 (0.084)
	CV	0.214 (0.304)	0.063 (0.167)	0.029 (0.115)
	SURE	0.041 (0.136)	0.010 (0.069)	0.004 (0.047)
	BAMS	0.373 (0.392)	0.371 (0.391)	0.371 (0.391)
	LPM	0.116 (0.272)	0.029 (0.136)	0.012 (0.090)
	SYM BETA	0.022 (0.091)	0.004 (0.041)	0.001 (0.026)
	ASYM BETA	0.028 (0.100)	0.010 (0.054)	0.007 (0.041)

We can observe an excellent performance of the shrinkage rule under asymmetric prior in both asymmetric sceneries. In fact, our rule had the best performance in

terms of AMSE and AMAE in all the considered sceneries of sample size and SNR and both the contexts of hyperparameters choice. We emphasize the second sceneries, with strong asymmetry and less sparsity degree of the coefficients, where the difference in performance of the rule with the others are significant. It suggests that when the wavelet coefficients are asymmetrically distributed, our proposed rule should be considered as the shrinker to be applied.

Another interesting observed feature is the good performance of the rule for low SNR values. When $\text{SNR} = 3$, the beta rule had significant difference against the comparison methods, i.e., the shrinker can be well succeeded even for presence with high noise level in the data, which is a desirable for shrinkers.

Moreover, the performance of the asymmetric rule in the symmetric scenery was close to the symmetric rule one, the best method in this context as expected, in both measures.

Boxplots of MSEs and MAEs obtained for the shrinkage/thresholding rules in both contexts for $n = 512$ and $\text{SNR} = 3$ are presented in Figures 5 (a) and (b).

5.2. Simulation Study 2

To conclude our simulation studies, we evaluated the performance of the proposed shrinkage rule in the four Donoho-Jonhstone test functions called Bumps, Blocks, Doppler and Heavisine. Figure 6 (a) shows these popular functions in statistical wavelet research, once each of them has interesting features such as discontinuities, spikes and oscillations that are important to be captured in curve estimation procedures by wavelet modelling.

In fact, the four functions present asymmetrically distributed wavelet coefficients with different levels. Bumps and Doppler functions have the highest levels of wavelet coefficients asymmetry and the Heavisine, the shortest one, i.e., its coefficients are almost symmetric around zero. These previous knowledge allows us to explain the different performances of our asymmetric shrinker among the functions.

In each function, data were generated according to the addition of normal random noise with the same scenarios of sample size and SNR of simulation study 1. The AMSEs and AMAEs obtained for each rule are in Tables 6 and 7, for $M = 500$ simulation runs for each scenario.

We observe that the proposed shrinkage rule had great performance in practically all the scenarios. It was the best one for the performance measures in Bumps and Doppler functions and beat the comparison rules in some scenarios for Blocks function. Even for Heavisine function, that has the least asymmetric coefficients distribution, which was dominated by CV method, our asymmetric beta rule had reasonable performance, with AMSE and AMAE very close to the best methods on almost all the scenarios for this function. This feature should be emphasized: although the asymmetric shrinker had the best performance when applied in empirical wavelet coefficients of functions with significant asymmetrical wavelet coefficients distributions, which was already expected, it also had good results when the wavelet coefficients of the signal are practically symmetric. This

provides some sort of flexibility of the proposed shrinker, which is essential for real data applications. Finally, as observed in simulation study 1, we also have good results for low SNR values, such as 3.

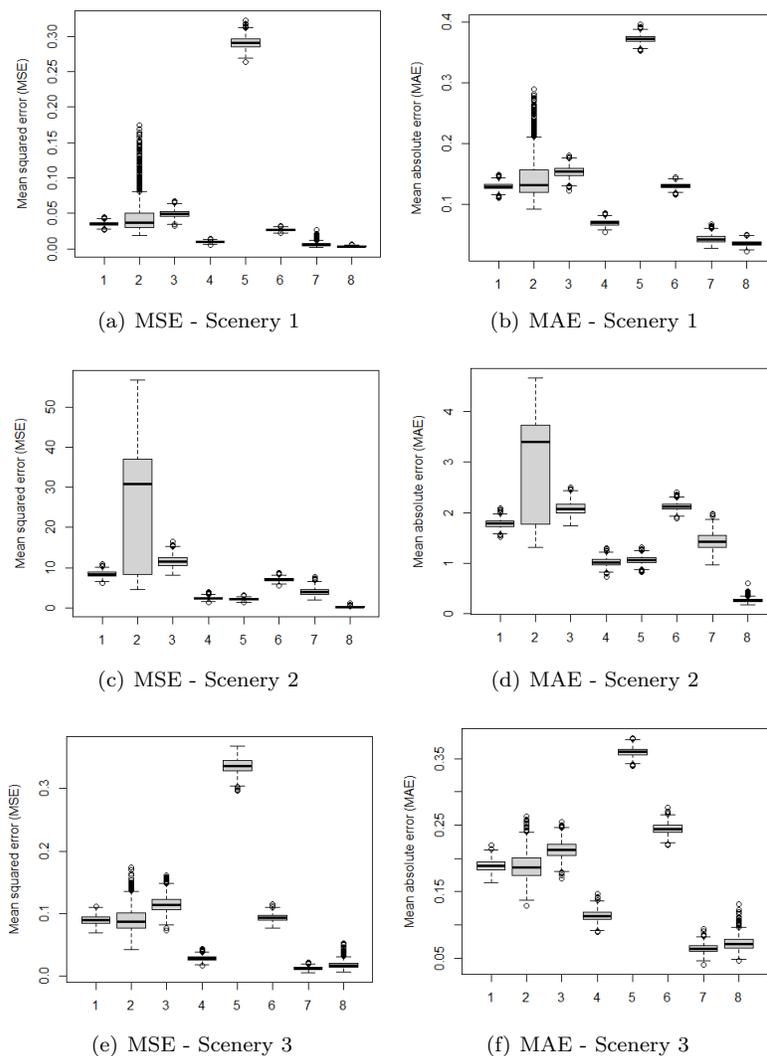
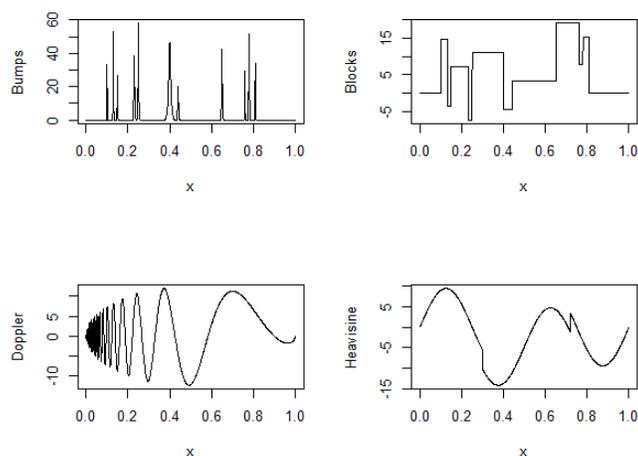
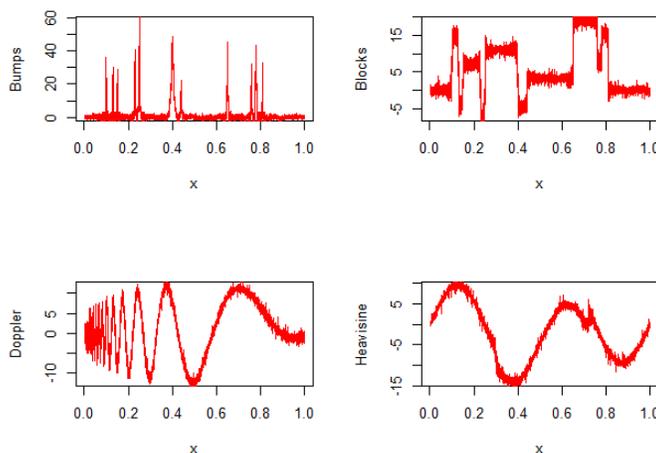


FIGURE 5: Boxplots of MSEs and MAEs of shrinkage/thresholding rules in simulation study 1 involving artificial asymmetric distributed wavelet coefficients for $n = 512$ and $\text{SNR}=3$ according to the models (4), (5) and (6) for $\alpha = 0.9$, $a = 3$, $b = 7$ and $m = 10$ (Scenery 1) and $\alpha = 0.6$, $a = 1$, $b = 20$ and $m = 30$ (Scenery 2) and $\alpha = 0.9$, $a = 5$, $b = 5$ and $m = 30$ (Scenery 3). The associated rules are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM and 7-SYM BETA and 8- ASYM BETA.



(a) Donoho-Johnstone test functions.



(b) Fitted curves.

FIGURE 6: Donoho-Johnstone test functions used as underlying signals in simulation study 2 (a) and fitted curves obtained after application of asymmetric shrinkage rule under beta prior for $n = 512$ and $SNR = 6$.

Figures 6 (b), 7 and 8 show fitted curves obtained after denoising by asymmetric shrinkage rule and boxplots of the MSEs and MAEs of the shrinkage/thresholding rules respectively for $n = 512$ and $SNR = 6$. One can note that the fitted curves recover the main features of each signal, as jumps and spikes. Moreover, the MSEs of the proposed rule (rule number 8 at the boxplots) had low variation.

TABLE 6: AMSE and AMAE of the shrinkage/thresholding rules in the simulation study for the empirical wavelet coefficients vector artificially generated with Donoho Jonhstone test functions Bumps and Blocks as underlying signals and additive normal random noise.

Signal	n	Method	SNR = 3	SNR = 6	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
Bumps	512	UNIV	11.122 (1.773)	3.882 (1.146)	2.014 (0.862)
		FDR	9.313 (1.664)	3.324 (1.073)	1.774 (0.814)
		CV	11.444 (1.792)	7.665 (1.517)	4.533 (1.219)
		SURE	3.656 (1.229)	1.159 (0.713)	0.578 (0.511)
		BAMS	2.833 (1.240)	1.355 (0.827)	1.160 (0.772)
		LPM	5.441 (1.862)	1.363(0.930)	0.606 (0.620)
		ASYM BETA	2.763 (1.205)	1.348 (0.757)	0.684 (0.612)
	ASYM BETA	2.968 (1.250)	1.075 (0.729)	0.548 (0.528)	
	2048	UNIV	5.050 (1.223)	1.769 (0.761)	0.935 (0.569)
		FDR	3.582 (1.052)	1.173 (0.635)	0.602 (0.465)
		CV	1.610 (0.816)	0.578 (0.477)	0.381 (0.378)
		SURE	1.651 (0.818)	0.510 (0.466)	0.250 (0.328)
		BAMS	1.635 (0.943)	0.573 (0.531)	0.482 (0.478)
		LPM	5.453 (1.861)	1.359 (0.931)	0.604 (0.620)
ASYM BETA		1.555 (0.822)	0.734 (0.546)	0.550 (0.451)	
ASYM BETA	1.450 (0.764)	0.465 (0.432)	0.298 (0.310)		
Signal	n	Method	SNR = 3	SNR = 6	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
Blocks	512	UNIV	6.940 (1.852)	2.846 (1.151)	1.524 (0.840)
		FDR	5.911 (1.693)	2.244 (1.013)	1.164 (0.736)
		CV	2.575 (1.146)	1.000 (0.694)	0.673 (0.561)
		SURE	2.827 (1.181)	0.904 (0.667)	0.448 (0.475)
		BAMS	2.469 (1.172)	1.163 (0.771)	1.018 (0.719)
		LPM	5.469 (1.861)	1.363 (0.930)	0.605 (0.622)
		ASYM BETA	2.811 (1.494)	1.945 (1.192)	1.376 (1.036)
	ASYM BETA	2.834 (1.227)	0.950 (0.706)	0.416 (0.470)	
	2048	UNIV	3.417 (1.241)	1.376 (0.772)	0.757 (0.567)
		FDR	2.688 (1.093)	0.967 (0.645)	0.513 (0.464)
		CV	1.307 (0.788)	0.444 (0.452)	0.248 (0.333)
		SURE	1.359 (0.796)	0.441(0.453)	0.220 (0.320)
		BAMS	1.502 (0.901)	0.499 (0.491)	0.418 (0.441)
		LPM	5.450 (1.860)	1.361 (0.931)	0.605 (0.620)
ASYM BETA		2.245 (1.234)	0.993 (0.870)	0.497 (0.591)	
ASYM BETA	1.306 (0.769)	0.445 (0.450)	0.231 (0.321)		

TABLE 7: AMSE and AMAE of the shrinkage/thresholding rules in the simulation study for the empirical wavelet coefficients vector artificially generated with Donoho Jonhstone test functions Doppler and Heavisine as underlying signals and additive normal random noise.

Signal	n	Method	SNR = 3	SNR = 6	SNR = 9
			AMSE (AMAE)	AMSE (AMAE)	AMSE (AMAE)
Doppler	512	UNIV	2.645 (1.143)	1.102 (0.752)	0.608 (0.566)
		FDR	2.540 (1.119)	0.967 (0.699)	0.514 (0.520)
		CV	1.269 (0.821)	0.518 (0.520)	0.374 (0.438)
		SURE	1.318 (0.835)	0.435 (0.480)	0.216 (0.337)
		BAMS	1.527 (0.933)	0.537 (0.549)	0.461 (0.512)
		LPM	5.448 (1.861)	1.365 (0.930)	0.603 (0.620)
		SYM BETA	1.656 (1.030)	0.987 (0.812)	0.751 (0.722)
		ASYM BETA	1.303 (0.781)	0.469 (0.485)	0.212 (0.334)
	2048	UNIV	1.155 (0.724)	0.457 (0.463)	0.257 (0.345)
		FDR	1.040 (0.691)	0.377 (0.419)	0.199 (0.301)
		CV	0.557 (0.520)	0.191 (0.300)	0.097 (0.212)
		SURE	0.573 (0.527)	0.194 (0.302)	0.097 (0.213)
		BAMS	1.084 (0.770)	0.211 (0.335)	0.160 (0.283)
		LPM	5.439 (1.864)	1.358 (0.930)	0.604 (0.620)
Heavisine	512	SYM BETA	1.033 (0.808)	0.705 (0.663)	0.378 (0.497)
		ASYM BETA	0.468 (0.458)	0.174 (0.280)	0.103 (0.204)
		UNIV	0.570 (0.529)	0.348 (0.400)	0.237 (0.334)
		FDR	0.594 (0.543)	0.368 (0.410)	0.228 (0.328)
		CV	0.511 (0.503)	0.219 (0.329)	0.122 (0.246)
		SURE	0.573 (0.531)	0.360 (0.405)	0.249 (0.341)
		BAMS	1.142 (0.797)	0.258 (0.380)	0.204 (0.333)
		LPM	5.429 (1.863)	1.365 (0.931)	0.604 (0.618)
	2048	SYM BETA	1.112 (0.790)	0.788 (0.660)	0.703 (0.633)
		ASYM BETA	0.623 (0.753)	0.267 (0.608)	0.136 (0.563)
		UNIV	0.358 (0.394)	0.193 (0.290)	0.123 (0.229)
		FDR	0.389 (0.410)	0.186 (0.284)	0.111 (0.216)
		CV	0.264 (0.349)	0.109 (0.222)	0.060 (0.163)
		SURE	0.360 (0.395)	0.197 (0.292)	0.111 (0.217)
	BAMS	0.982 (0.725)	0.143 (0.278)	0.100 (0.226)	
	LPM	5.444 (1.861)	1.359 (0.930)	0.604 (0.620)	
	SYM BETA	0.926 (0.732)	0.724 (0.657)	0.429 (0.505)	
	ASYM BETA	0.548 (0.649)	0.795 (0.551)	0.326 (0.392)	

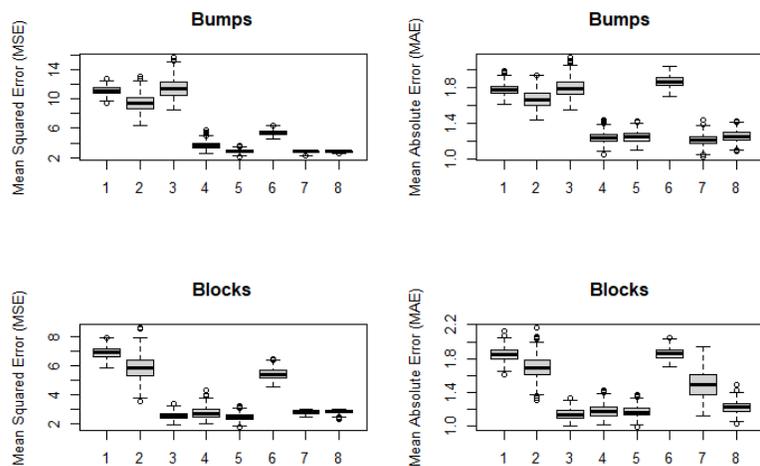


FIGURE 7: Boxplots of MSEs and MAEs of shrinkage/thresholding rules in simulation study 2 involving Donoho-Johnstone test functions Bumps and Blocks for $n = 512$ and $\text{SNR} = 6$. The associated rules are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-SYM BETA and 8-ASYM BETA.

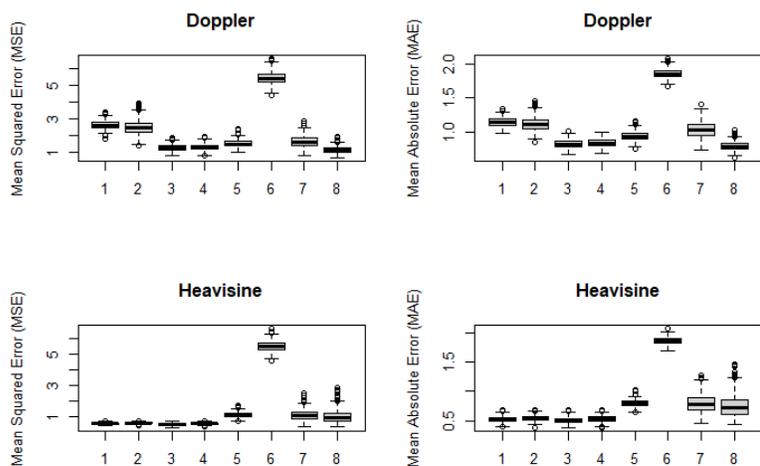


FIGURE 8: Boxplots of MSEs and MAEs of shrinkage/thresholding rules in simulation study 2 involving Donoho-Johnstone test functions Doppler and Heavisine for $n = 512$ and $\text{SNR} = 6$. The associated rules are: 1-UNIV, 2-FDR, 3-CV, 4-SURE, 5-BAMS, 6-LPM, 7-SYM BETA and 8-ASYM BETA.

6. Application: Seismic Dataset

Coso, California, is a geothermal area, with geological structure tectonically active and its geological and geophysics properties have been studied along the last decades by experts of the area. Once with a seismogram at hand, the experts classify the subsets of the data as primary waves (P-waves), secondary waves (S-waves) and body waves, according to some features of the signal, as amplitude, velocity and other geophysical parameters. The precision of the classification could be extremely important to predict events, such earthquakes. In this sense, denoising of seismogram with statistical methods is crucial for precision of such wave type classifications and naturally, wavelet based statistical methods are some of the most attractive and proposed methods for this purpose. Chik et al. (2009), To et al. (2009), Ansari et al. (2010), Beenamol et al. (2012, 2016), Mousavi et al. (2016) and Vargas and Veiga (2017) are some of relevant works related to wavelet based methods applied in denoising seismic data.

We applied the proposed shrinkage rule in denoising seismic amplitudes dataset collected from Coso, California and available in RSEIS R package (Lees et al., 2020). The available dataset has 726 seismic amplitudes measured in a short time interval. We considered $J = \lfloor \log_2 726 \rfloor = 9$ resolution levels, thus $n = 2^9 = 512$ data points. The considered seismogram is shown in Figure 9 (a). For more details of Coso geological studies and the dataset, the reader is addressed to Lees (2004).

After application of a DWT using Daub10 basis, we observed an empirical right asymmetry of the coefficients, with $Skew(d) = 0,25$ and $\hat{\sigma} = 298,38$. Then, our asymmetric shrinkage rule was applied for denoising the empirical coefficients, with hyperparameters $a = 2$, $b = 3$, $\alpha = \alpha(j)$ and $m = m(j)$ according to (9) and (10) respectively. The denoised seismogram is shown in Figure 9 (b). From comparison of figures in (9), one can observe considered noise reduction, mainly on the final period of the seismogram, when the seismic activity becomes to decrease.

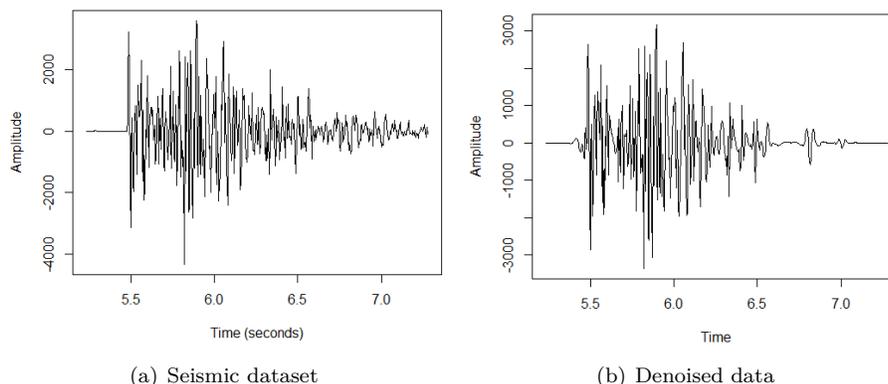


FIGURE 9: Considered seismogram from Coso, California (a) and its denoised version (b) after application of the proposed asymmetric shrinkage rule under beta prior.

Noise reduction can also be observed directly from the wavelet coefficients of the considered seismogram and its denoised version, shown in Figures 10 (a) and (b) respectively. In fact, denoising occurs mainly at high resolution levels, where most of non-zero magnitudes of the coefficients are typically attributed to noise, as already mentioned in Section 5.

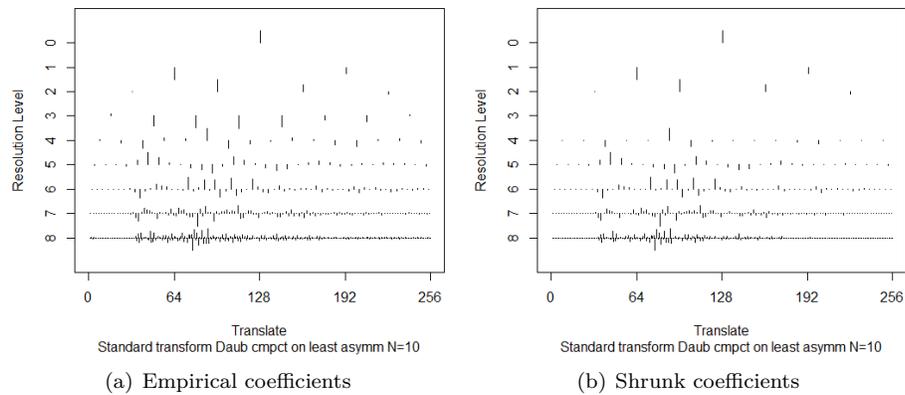


FIGURE 10: Empirical wavelet coefficients of the considered seismogram dataset (a) and their shrunk versions after denoising by the proposed asymmetric shrinkage rule under beta prior (b).

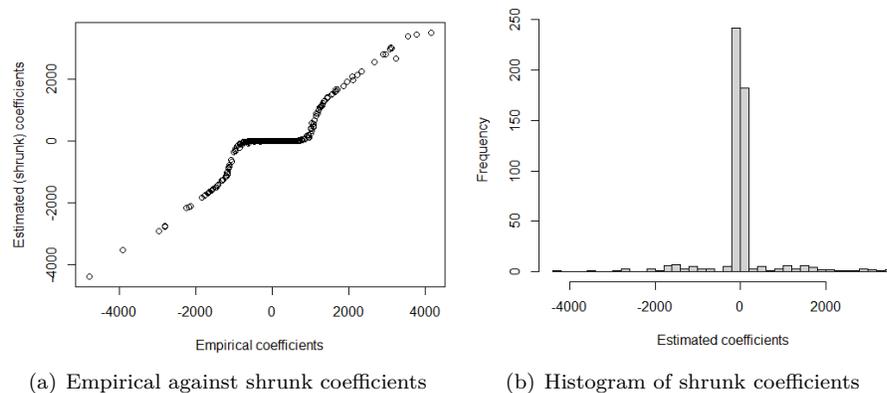


FIGURE 11: Empirical coefficients against their shrunk versions (a) and histogram of shrunk coefficients obtained by application of asymmetric shrinkage rule under beta prior.

Finally, we present the asymmetric shrinkage process in Figures 11 (a) and (b). The first one shows us the plot of empirical coefficients against their shrunk versions ones. There is a weak right asymmetry around zero, i.e, empirical coefficients greater than zero had a little bit stronger shrinkage than the smaller ones, which is the effect of the chosen hyperparameters of the beta, $a = 2$ and

$b = 3$. Although the asymmetry is not high, the adaptive asymmetric prior assigned according to this weak deviation of symmetry improved the denoising performance of the shrinkage rule. Figure 11 (b) presents the histogram of shrunk coefficients, which emphasizes the sparsity of the estimated coefficients vector, with more than a half of the 512 coefficients shrunk to zero or very close to it.

7. Final Considerations

We propose in this work the use of asymmetric prior based on the beta distribution to the wavelet coefficients, which is a novelty in the wavelet shrinkage, since all the already proposed shrinkage methods are typically symmetric around zero. Moreover, few studies are concerned to bounded energy signals, which imply in bounded wavelet coefficients. In this sense, a bounded prior proposition, as the beta distribution, can be an alternative for applications in this context.

The easy interpretation of the beta hyperparameters a and b in terms of asymmetry and shrinkage level and the well known flexibility of this distribution allow elicitation of the hyperparameters and adaptivity in modelling the coefficients, which are very attractive in bayesian setup. Further, the associated shrinkage rule had great performance in simulation studies and outperformed the considered shrinkage/thresholding methods in most of the scenarios, mainly when coefficients are highly asymmetric. Even when the coefficients are close to symmetry, the asymmetric shrinkage rule showed satisfactory results. These features allow the asymmetric beta to be considered by practitioners as a candidate to bayesian modelling of wavelet coefficients.

The proposed shrinkage rule was obtained under squared loss function, which is symmetric around zero. Due the considered asymmetry aspect of the wavelet coefficients, asymmetric loss functions can naturally be considered instead of the squared function one. Further, the impact of wavelet basis choice, the proposition of other asymmetric distributions to wavelet coefficients, the evaluation of the proposed beta shrinkage rule in other performance measures are trivial possible extensions of this work.

Acknowledgements

This work was supported by CAPES¹ fellowship.

[Received: December 2020 — Accepted: August 2021]

References

Abramovich, F. & Benjamini, Y. (1996), 'Adaptive thresholding of wavelet coefficients', *Computational Statistics and Data Analysis* **22**(4), 351–361.

¹Coordination of Superior Level Staff Improvement, Brazil

- Abramovich, F., Sapatinas, T. & Silverman, B. (1998), 'Wavelet thresholding via a bayesian approach', *Royal Statistical Society* pp. 725–749.
- Angelini, C. & Vidakovic, B. (2004), 'Gama-minimax wavelet shrinkage: a robust incorporation of information about energy of a signal in denoising applications', *Statistica Sinica* (14), 103–125.
- Antoniadis, A., Bigot, J. & Sapatinas, T. (2001), 'Wavelet estimators in nonparametric regression: a comparative simulation study', *Journal of Statistical Software* (6), 1–83.
- Beenamol, M., Prabavathy, S. & Mohanalin, J. (2012), 'Wavelet based seismic signal de-noising using shannon and tsallis entropy', *Computers and Mathematics with Applications* (64), 3580–3593.
- Bhattacharya, A., Pati, D., Pillai, N. & Dunson, D. (2015), 'Dirichlet-laplace priors for optimal shrinkage', *Journal of the American Statistical Association* (110), 1479–1490.
- Chipman, H., Kolaczyk, E. & McCulloch, R. (1997), 'Adaptive bayesian wavelet shrinkage', *Journal of the American Statistical Association* (92), 1413–1421.
- Cuttillo, L., Jung, Y., Ruggeri, F. & Vidakovic, B. (2008), 'Larger posterior mode wavelet thresholding and applications', *Journal of Statistical Planning and Inference* (138), 3758–3773.
- Donoho, D. L. (1995a), 'De-noising by soft-thresholding', *IEEE Transactions on Information Theory* (41), 6136–6142.
- Donoho, D. L. (1995b), 'Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition', *Applied and Computational Harmonic Analysis* (2), 101–126.
- Donoho, D. L. & Johnstone, I. M. (1994a), 'Ideal denoising in an orthonormal basis chosen from a library of bases', *Comptes Rendus de l'Académie des Sciences* (319), 1317–1322.
- Donoho, D. L. & Johnstone, I. M. (1994b), 'Ideal spatial adaptation by wavelet shrinkage', *Biometrika* (81), 425–455.
- Donoho, D. L. & Johnstone, I. M. (1995), 'Adapting to unknown smoothness via wavelet shrinkage', *Journal of the American Statistical Association* (90), 1200–1224.
- Donoho, D. L., Johnstone, I. M., Kerkycharian, G. & Picard, D. (1995), 'Wavelet shrinkage: Asymptopia? (with discussion)', *Royal Statistical Society B* (57), 301–369.
- Donoho, D. L., Johnstone, I. M., Kerkycharian, G. & Picard, D. (1996), 'Density estimation by wavelet thresholding', *Annals of Statistics* **24**, 508–539.

- Griffin, J. & Brown, P. (2017), ‘Hierarchical shrinkage priors for regression models’, *Bayesian Analysis* **12**(1), 135–159.
- Jansen, M. (2001), *Noise reduction by wavelet thresholding*, Springer, New York.
- Johnstone, L. & Silverman, B. (2005), ‘Empirical bayes selection of wavelet thresholds’, *The Annals of Statistics* **33**, 1700–1752.
- Karagiannis, G., Konomi, B. & Lin, G. (2015), ‘A bayesian mixed shrinkage prior procedure for spatial-stochastic basis selection and evaluation of gpc expansion: applications to elliptic spdes’, *Journal of Computational Physics* **284**, 528–546.
- Lian, H. (2011), ‘On posterior distribution of bayesian wavelet thresholding’, *Journal of Statistical Planning and Inference* **141**, 318–324.
- Mallat, S. G. (1998), *A Wavelet Tour of Signal Processing*, Academic Press, San Diego.
- Nason, G. P. (1996), ‘Wavelet shrinkage using cross-validation’, *Journal of the Royal Statistical Society B* **58**, 463–479.
- Reményi, N. & Vidakovic, B. (2015), ‘Wavelet shrinkage with double weibull prior’, *Communications in Statistics: Simulation and Computation* **44**(1), 88–104.
- Sousa, A. (2020), ‘Bayesian wavelet shrinkage with logistic prior’, *Communications in Statistics: Simulation and Computation* .
- Sousa, A., Garcia, N. & Vidakovic, B. (2021), ‘Bayesian wavelet shrinkage with beta prior’, *Computational Statistics* **36**(2), 1341–1363.
- Torkamani, R. & Sadeghzadeh, R. (2017), ‘Bayesian compressive sensing using wavelet based markov random fields’, *Signal Processing: Image Communication* **58**, 65–72.
- Vidakovic, B. (1998), ‘Nonlinear wavelet shrinkage with bayes rules and bayes factors’, *Journal of the American Statistical Association* **93**, 173–179.
- Vidakovic, B. (1999), *Statistical Modeling by Wavelets*, Wiley, New York.
- Vidakovic, B. & Ruggeri, F. (2001), ‘Bams method: Theory and simulations’, *Sankhya: The Indian Journal of Statistics* **63**, 234–249.