Available online at www.sciencedirect.com
SCIENCE DIRECT®
Appl. Comput. Harmon. Anal. ●●● (●●●●) ●●●–●●●
Applied and Computational Harmonic Analysis
www.elsevier.com/locate/acha

Letter to the Editor

Nonlinear wavelet thresholding: A recursive method to determine the optimal denoising threshold

Alexandre Azzalini [a], Marie Farge [a,*], Kai Schneider [b]

[a] *Laboratoire de Météorologie Dynamique, IPSL-CNRS, Ecole Normale Supérieure, Paris, France*
[b] *Laboratoire de Modélisation et de Simulation Numérique en Mécanique, CNRS and Centre de Mathématiques et d'Informatique, Université de Provence, Marseille, France*

Received 21 September 2004; accepted 1 October 2004

Communicated by M.V. Wickerhauser

Abstract

Nonlinear thresholding of wavelet coefficients is an efficient method for denoising signals with isolated singularities. The quasi-optimal value of the threshold depends on the sample size and on the variance of the noise, which is in many situations unknown. We present a recursive algorithm to estimate the variance of the noise, prove its convergence and investigate its mathematical properties. We show that the limit threshold depends on the probability density function (PDF) of the noisy signal and that it is equal to the theoretical threshold provided that the wavelet representation of the signal is sufficiently sparse. Numerical tests confirm these results and show the competitiveness of the algorithm compared to the median absolute deviation method (MAD) in terms of computational cost for strongly noised signals.

Keywords: Wavelets; Denoising; Data analysis; Estimation

* Corresponding author.
E-mail addresses: azzalini@lmd.ens.fr (A. Azzalini), farge@lmd.ens.fr (M. Farge), kschneid@cmi.univ-mrs.fr (K. Schneider).

1063-5203/$ – see front matter
doi:10.1016/j.acha.2004.10.001

1. Introduction

Estimating signals or images from noisy data is a typical problem in data processing with many applications. Many parametric and nonparametric approaches, such as linear kernel estimators, Kalman filters, have been proposed, see, e.g., [5]. Nonlinear thresholding of the empirical wavelet coefficients was originally proposed by Donoho and Johnstone [2] to denoise signals corrupted with Gaussian white noise. It consists in deleting the wavelet coefficients of the noisy signal whose modulus is below a threshold and reconstructing the denoised signal from the remaining coefficients. The threshold depends only on the sample size and on the noise's variance. The method was later generalized to correlated noise and to non-Gaussian situations [6,7]. Wavelet thresholding estimators minimize the maximum L^2-risk in a whole class of finite energy signals including Hölder and Besov spaces without any a priori knowledge of the signal, but the unknown variance of the noise has to be estimated. The median absolute deviation (MAD) is a standard method that estimates the level of the noise by taking the median of the modulus of the smallest scale wavelet coefficients [5]. In the present paper we introduce a new recursive algorithm to estimate the variance of the noise, study its properties regarding convergence, stability and performance, and validate the results with a numerical example.

2. Denoising by nonlinear wavelet thresholding

We consider a discrete signal S of size $N = 2^J$ with vanishing mean, corrupted by a Gaussian white noise of mean zero and variance σ_W^2 resulting in $X_k = S_k + W_k$ for $k = 0, \ldots, N-1$, where X_k and W_k are N samples of the noisy data and the noise, respectively.

We decompose the noisy data X into an orthogonal wavelet series $X = \sum_{\lambda \in \Lambda^J} \tilde{X}_\lambda \psi_\lambda$ where the multi-index $\lambda = (j, i)$ denotes the scale j and the position i of the wavelets. The corresponding index set Λ^J is

$$\Lambda^J = \left\{\lambda = (j, i),\ j = 0, \ldots, J-1,\ i = 0, \ldots, 2^j - 1\right\}.$$

By thresholding the wavelet coefficients $\tilde{X}_\lambda$ and reconstructing the corresponding signal we define a nonlinear operator

$$F_T : X \mapsto F_T(X) = \sum_\lambda \rho_T(\tilde{X}_\lambda)\psi_\lambda \tag{1}$$

with the thresholding function

$$\rho_T(a) = \begin{cases} a & \text{if } |a| > T, \\ 0 & \text{if } |a| \leqslant T, \end{cases}$$

where T denotes the threshold. We denote by Λ_T the index subset of wavelet coefficients $\tilde{X}$ that are selected by the thresholding function ρ_T, such that $\Lambda_T = \{\lambda \in \Lambda^J,\ |\tilde{X}_\lambda| > T\} \subset \Lambda^J$. Donoho and Johnstone [2] showed that the relative quadratic error between the signal S and its estimator $F_T(X)$, defined by

$$\mathcal{E}(T) = \frac{\|S - F_T(X)\|^2}{\|S\|^2}, \tag{2}$$

has its lower bound, $\min_T \mathcal{E}(T)$, close to the minimax error for all signals $S \in \mathcal{H}$ where $\mathcal{H}$ belongs to a wide class of function spaces, including Hölder and Besov spaces. They also showed that the error $\mathcal{E}(T_D)$ corresponding to the threshold

$$T_D = \sigma_W (2 \ln N)^{1/2} \tag{3}$$

is close to the minimum of $\mathcal{E}(T)$. Since T_D depends only on the variance of the noise, it is called universal threshold in contrast to the value $T_{\min}$ that minimizes the error $\mathcal{E}(T)$. However, in many applications σ_W is unknown and has to be estimated from the available noisy data X.

To address the estimation of the noise, we adopt a dual point of view: Instead of considering the denoised part $F_T(X)$ of the noisy signal X, we focus on the residual which was not taken into account in $F_T(X)$; namely,

$$F_T^c(X) = (\mathrm{Id} - F_T)(X) = X - F_T(X) = \sum_{\lambda \in \Lambda^J} \rho_T^c(\tilde{X}_\lambda)\psi_\lambda = \sum_{\lambda \in \Lambda_T^c} \tilde{X}_\lambda \psi_\lambda, \tag{4}$$

where Id denotes the identity. The complementary operator F_T^c uses the complementary thresholding function $\rho_T^c = \mathrm{Id} - \rho_T$ and defines the complementary index set $\Lambda_T^c = \Lambda^J \setminus \Lambda_T$. The residual $F_{T_D}^c(X)$ is a quasi-optimal estimator of the Gaussian white noise W, whose relative error is

$$\mathcal{E}'(T) = \frac{\|X - F_T(X) - W\|^2}{\|W\|^2} = \frac{\|S + W - F_T(X) - W\|^2}{\|W\|^2} = \frac{\|S\|^2}{\|W\|^2}\mathcal{E}(T). \tag{5}$$

3. Recursive algorithm

In [4], we proposed a recursive algorithm for denoising based on the conjecture that, given a threshold T_n, the variance of the noise estimated by $F_{T_n}^c(X)$ yields a threshold T_{n+1} closer to T_D than T_n. In the following, we present the algorithm and check the validity of this conjecture.

Algorithm 1.

Initialization

- Given X_k, $k = 0, \ldots, N-1$. Set $n = 0$ and compute the fast wavelet transform of X to obtain $\tilde{X}_\lambda$.
- Compute the variance σ_0^2 of X as a rough estimate of the variance of W and compute the corresponding threshold $\sigma_0^2 = \frac{1}{N}\sum_{\lambda \in \Lambda^J} |\tilde{X}_\lambda|^2$, $T_0 = (2 \ln N \sigma_0^2)^{1/2}$.
- Set the number of coefficients considered as noise $N_W = \mathrm{Card}(\Lambda^J) = N$.

Main loop

Repeat

- Set $N'_W = N_W$ and count the wavelet coefficients smaller than T_n: $N_W = \mathrm{Card}(\Lambda_{T_n}^c)$.
- Compute the new variance $\sigma_{n+1}^2 = \frac{1}{N}\sum_{\lambda \in \Lambda^J} |\rho_{T_n}^c(\tilde{X}_\lambda)|^2$ and the new threshold $T_{n+1} = (2(\ln N)\sigma_{n+1}^2)^{1/2}$.

- Set $n = n + 1$

until $(N'_W = N_W)$.

Final step

- Compute $F_{T_n}(X)$ from the coefficients $\{\tilde{X}_\lambda\}_{\lambda \in \Lambda_{T_n}}$ using inverse fast wavelet transform and compute $F^c_{T_n}(X) = X - F_{T_n}(X)$.

This algorithm defines a sequence of estimated thresholds $(T_n)_{n\in\mathbb{N}}$ and the corresponding sequence of estimated variances $(\sigma_n^2)_{n\in\mathbb{N}}$. Their convergence depends on their initial value and on the *iteration function*

$$I_{X,N} : \mathbb{R}^+ \mapsto \mathbb{R}^+ \quad \text{such that} \quad T_{n+1} = I_{X,N}(T_n),$$

which is obtained by merging the definitions of σ^2_{n+1} and T_{n+1}:

$$I_{X,N}(T) = \left(\frac{2\ln N}{N} \sum_{\lambda \in \Lambda^J} \left|\rho^c_T(\tilde{X}_\lambda)\right|^2\right)^{1/2} = \left(\frac{2\ln N}{N} \sum_{\lambda \in \Lambda^c_T} |\tilde{X}_\lambda|^2\right)^{1/2}. \tag{6}$$

3.1. Properties of the iteration function

Taking the square of (6), we rewrite the sum as a continuous integral using delta functions:

$$\left(I_{X,N}(T)\right)^2 = \frac{2\ln N}{N} \int_{t=0}^{T} t^2 \sum_{\lambda \in \Lambda^J} \delta\left(|\tilde{X}_\lambda| - t\right) \mathrm{d}t. \tag{7}$$

This expression shows that the function $I_{X,N}(T)$ is piecewise constant with a number of discontinuities smaller than N and is therefore bounded both from below and above. Moreover, the iteration function is monotonically increasing, i.e.,

$$I_{X,N}(T) \leqslant I_{X,N}(T + \Delta T) \quad \forall T, \Delta T \in \mathbb{R}^+. \tag{8}$$

3.2. Convergence

In the following we prove the convergence of the recursive algorithm by applying fixed point arguments to the iteration function $I_{X,N}$. Theorem 1 proves that, if there exists an interval such that the iteration function is above the line $y = x$ at the lower bound and below this line at the upper bound, then the sequence of thresholds converges as soon as it enters this interval. Corollary 1 shows that these particular conditions are always satisfied by the iteration function.

Theorem 1. *We consider an interval $[T_a, T_b] \subset \mathbb{R}^+$ such that $I_{X,N}(T_a) \geqslant T_a$ and $I_{X,N}(T_b) \leqslant T_b$. If there exists a step n_0 such that $T_{n_0} \in [T_a, T_b]$, then $T_n = I_{X,N}(T_{n-1})$ converges to a limit T_ℓ within $[T_a, T_b]$ such that $T_\ell = I_{X,N}(T_\ell)$. The number of iterations n_ℓ is smaller than N.*

S1063-5203(04)00075-2/FLA AID:456 Vol.●●●(●●●) yacha456 [DTD5] P.5 (1-9)
YACHA:m2 v 1.32 Prn:6/01/2005; 10:38 by:EA p. 5

Proof. We suppose that $I_{X,N}(T_{n_0}) < T_{n_0}$. Expression (8) implies that $(I_{X,N} \circ I_{X,N}(T_{n_0})) \leqslant I_{X,N}(T_{n_0})$ and hence

$$T_{n_0+2} = I_{X,N}(T_{n_0+1}) \leqslant T_{n_0+1} = I_{X,N}(T_{n_0}) < T_{n_0}, \tag{9}$$

which shows that the sequence $\{T_n\}_{n\geqslant n_0}$ decreases. As $T_a < T_{n_0}$, expression (8) implies that $I_{X,N}(T_a) \leqslant I_{X,N}(T_{n_0})$. As we assumed $T_a \leqslant I_{X,N}(T_a)$, we find $T_a \leqslant T_{n_0+1}$ and therefore $T_a \leqslant T_n$ for all $n \geqslant n_0$. Hence $\{T_n\}_{n\geqslant n_0}$ decreases, is bounded from below by T_a, and converges to a limit $T_\ell = \inf_{n\geqslant n_0}(T_n)$ between T_a and T_{n_0}. As the iteration function $I_{X,N}$ is piecewise constant with a finite number of discontinuities, its image (including the values taken by the sequence $\{T_n\}_{n>n_0}$) is countable with a cardinality smaller than N. As a consequence, there exists a n_ℓ such that $T_{n_\ell} = T_\ell = \inf_{n\geqslant n_0}(T_n)$. As $\{T_n\}_{n>n_0}$ is decreasing, one has $T_{n_\ell+1} = T_{n_\ell}$, i.e., $T_\ell = I_{X,N}(T_\ell)$. Conversely, if $I_{X,N}(T_{n_0}) > T_{n_0}$ one can show analogously that $\{T_n\}_{n\geqslant n_0}$ is increasing, is bounded from above by T_b, and therefore converges between T_{n_0} and T_b. □

Corollary 1. *One has* $\sup_{T\in\mathbb{R}^+} I_{X,N}(T) = T_0 = (2\ln N)^{1/2}\sigma_0$ *and* $I_{X,N}(0) = 0$. *Therefore, Theorem* 1 *implies that the sequence* $\{T_n\}_{n\in\mathbb{N}}$ *converges to a limit* $T_\ell \in [0, T_0]$.

Proof. When taking the threshold $T = 0$, the residual $F^c_{T=0}(X) = 0$. It naturally follows that $I_{X,N}(0) = 0$. On the other hand, the iteration function is bounded from above by T_0. This maximum value is reached with any threshold T larger than $T_{\max} = \sup_{\lambda\in\Lambda^J} |\tilde{X}_\lambda|$. Therefore, Theorem 1 is valid for any T_b chosen such that $T_b \geqslant \max(T_0, T_{\max})$, with $T_a = 0$ and with $T_{n_0} = T_0$. Hence, the sequence $\{T_n\}_{n\in\mathbb{N}}$ converges to a limit $T_\ell \in [0, T_b]$. As $I_{X,N}(T_b) = T_0$, the limit T_ℓ is actually in $[0, T_0]$. □

An additional point is the stability and self-consistency of the recursive algorithm. Corollary 2 shows that if we apply the recursive algorithm to the already denoised signal, this does not change the result.

Corollary 2. *Let* $\mathcal{A}: X \mapsto F_{T_\ell(X)}(X)$ *be the operator corresponding to the recursive algorithm described above. Then*

$$\mathcal{A} \circ \mathcal{A}(X) = \mathcal{A}(X) \quad \forall X \in \mathcal{H}.$$

This means that $\mathcal{A}$ *is a nonlinear projector.*

Proof. This property can be shown by considering the graph of the iteration function corresponding to $\mathcal{A}(X)$ defined as

$$I_{\mathcal{A}(X),N}(T) = \left(\frac{2\ln N}{N}\sum_{\lambda\in\Lambda^J}\left|\rho^c_T\left(\rho_{T_\ell}(\tilde{X}_\lambda)\right)\right|^2\right)^{1/2}, \tag{10}$$

where $T_\ell > 0$ is the threshold obtained from the recursive algorithm applied once. As expression (10) corresponds to a partial sum of the terms in $I_{X,N}(T)$, one has $I_{\mathcal{A}(X),N}(T) < I_{X,N}(T)\ \forall T \in \mathbb{R}^+$. As Theorem 1 implies that $I_{X,N}(T) \leqslant T$ for all $T \geqslant T_\ell$, we have $I_{\mathcal{A}(X),N}(T) < T$ for $T \geqslant T_\ell$. Hence, there is no fixed point for $I_{\mathcal{A}(X),N}$ in the interval $[T_\ell, +\infty[$. Furthermore, the fact that $\rho^c_T \circ \rho_{T_\ell} = 0\ \forall T < T_\ell$ implies $I_{\mathcal{A}(X),N}(T) = 0\ \forall T < T_\ell$. This means that the only possible fixed point for $I_{\mathcal{A}(X),N}$ is $T = 0$, which is the only possible limit for the sequence of thresholds $\{T_n\}_{n\in\mathbb{N}}$. Finally, the resulting estimation coincides with the first one. □

3.3. Convergence for Gaussian white noise

Stating that the successive estimations of the noise $F^c_{T_n}(X)$ converge close to the best estimation $F^c_{T_{\min}}(X)$ suggests that a Gaussian white noise is invariant with respect to the recursive algorithm. We check this assertion by applying the algorithm to a Gaussian white noise W. As the orthonormality of $\{\psi_\lambda\}_{\lambda\in\Lambda^J}$ implies that $\{\tilde{W}_\lambda\}_{\lambda\in\Lambda^J}$ is also a Gaussian white noise, the analytic expression of the PDF of its wavelet coefficients is known. Berman [3] showed that the probability that the maximum of the modulus of N values of a Gaussian white noise $\tilde{W}$ is inside the interval $[T_D - \sigma_W \ln(\ln N)/\ln N, T_D]$; namely,

$$P(N) = p\left(\max_\lambda\left(|\tilde{W}_\lambda|\right) \in \left[T_D - \frac{\sigma_W \ln(\ln N)}{\ln N}, T_D\right]\right) \tag{11}$$

tends to 1 for large N.

This result shows that for N large enough, the value T_D is a good estimator of the expected maximum modulus of the noise. At the first iteration of the algorithm, we have $T_D = (2\ln N)^{1/2}\sigma_W = (2\ln N)^{1/2}\sigma_0 = T_0$, which yields

$$I_{W,N}(T_0) = I_{W,N}(T_D) = \left(\frac{2\ln N}{N}\sum_{\lambda\in\Lambda^J}\left|\rho_{T_D}(\tilde{W}_\lambda)\right|^2\right)^{1/2} \simeq \left(\frac{2\ln N}{N}\sum_{\lambda\in\Lambda^J}|\tilde{W}_\lambda|^2\right)^{1/2} = T_0 = T_D. \tag{12}$$

This shows that the threshold T_0 obtained at the first iteration of the algorithm is almost a fixed point of the iteration function $I_{W,N}$. In addition, using the analytical expression of the Gaussian PDF of the noise, one can show that the derivative of the iteration function is almost zero around T_D. This forces the threshold T_ℓ to be close to T_D and the algorithm to converge in one iteration.
The remaining question is to determine whether T_ℓ is a correct estimator of T_D, which will be tested using a numerical example.

4. Numerical application

We apply the recursive algorithm to a one-dimensional test signal and illustrate its properties (cf. Fig. 1). We construct a noisy signal X by superposing a Gaussian white noise, with zero mean and variance $\sigma_W^2 = 1$, to a signal S, normalized in such a way that $(\frac{1}{N}\sum_k |S_k|^2)^{1/2} = 10$. The number of samples is $N = 8192$. We first apply the recursive algorithm to the signal S without any noise, then to the noise W only, and finally to the noisy signal X. We study the influence of the iteration functions $I_{S,N}$ and $I_{W,N}$ of the signal or noise alone, and on the iteration function $I_{X,N}$ of the total signal. We compare the results obtained with the threshold T_ℓ computed by the recursive algorithm, the universal threshold T_D computed with the known variance of the noise $\sigma_W^2 = 1$, and the threshold T_m obtained using MAD method [2] which estimates σ_W from the median of the wavelet coefficients of the noisy signal at the smallest scale. The resulting MAD threshold is given by the formula

$$T_m = \frac{(2\ln N)^{1/2}}{0.6745}\operatorname*{med}_{\lambda=(j,i)\in\{(j,i),\, j=J\}}\left(|\tilde{X}_\lambda|\right). \tag{13}$$

The iteration functions for X, S, and W are shown in Fig. 2. One observes that $I_{X,N}$ is superposed on $I_{W,N}$ for small values of T_n, but it follows $I_{S,N}$ for large values of T_n up to the point **C**, corresponding to the first iteration of the algorithm. In Fig. 3, we plot the histograms of the wavelet coefficients of X, S and

ARTICLE IN PRESS

S1063-5203(04)00075-2/FLA AID:456 Vol.●●●(●●●)
YACHA:m2 v 1.32 Prn:6/01/2005; 10:38

yacha456

[DTD5] P.7 (1-9)
by:EA p. 7

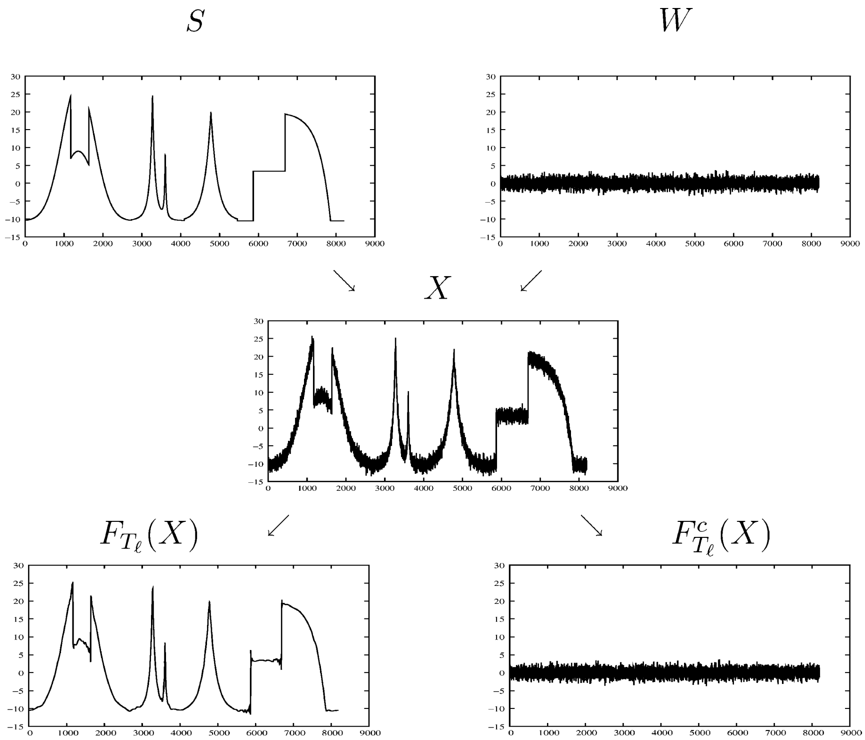

Fig. 1. Construction (top) of a 1D noisy signal $X = S + W$ (middle) and results obtained by the recursive algorithm (bottom).

W which are related to the iteration functions $I_{X,N}$, $I_{S,N}$, and $I_{W,N}$ shown in Fig. 2, since histograms are estimators of PDFs. The sparsity of the wavelet representation of the signal S causes most coefficients $\tilde{S}_\lambda$ to be close to zero, and therefore limits the growth of the corresponding iteration function $I_{S,N}$ which thus remains below the line $y = x$.

One also observes that the histograms of $\tilde{X}_\lambda$ and $\tilde{S}_\lambda$ present the same heavy tails for values larger than the expected maximum magnitude of the noise $T_D = (2\ln N)^{1/2}\sigma_W = 4.24$ (cf. Section 3.3). This agrees with the fact that $I_{X,N}$ is almost identical to $I_{S,N}$ for values larger than T_D, since the heavy tails of the PDF of $\tilde{S}_\lambda$ have a strong weight in the second-order moment of the histogram of the coefficients $\tilde{X}_\lambda$. On the contrary, the coefficients of the noise are concentrated within the range $[-T_D, T_D]$, and their contribution to $I_{X,N}(T)$ for T larger than T_D remains negligible. In contrast, when T is smaller than T_D, most of the coefficients $\tilde{S}_\lambda$ smaller than T are close to zero. Therefore, their contribution to the second-order moment $(I_{X,N}(T))^2$ is dominated by the contribution of the coefficients $\tilde{W}_\lambda$ whose distribution far from zero is wider. Thus the noise W dominates S in $I_{X,N}$ for T smaller than T_D, as soon as S is sparse enough in wavelet space. The consequence is that the intersection **B** of $I_{X,N}$ with $y = x$ remains close to the intersection **A** of $I_{W,N}$ with $y = x$. Therefore, the limit T_ℓ of the recursive algorithm applied to X is close to the limit obtained for the noise alone, which approximates T_D. This is true since no other fixed point is present for values of the threshold larger than T_D, due to the fact that between **B** and **C** the iteration function $I_{X,N}$ is below $y = x$. For this test signal, the algorithm converges to the value $T_\ell = 4.30$, which is close to the universal threshold $T_D = 4.24$. The resulting estimates $F_{T_\ell}(X)$ and $F^c_{T_\ell}(X)$ of S and W,

S1063-5203(04)00075-2/FLA AID:456 Vol.●●●(●●●) yacha456 [DTD5] P.8 (1-9)
YACHA:m2 v 1.32 Prn:6/01/2005; 10:38 by:EA p. 8

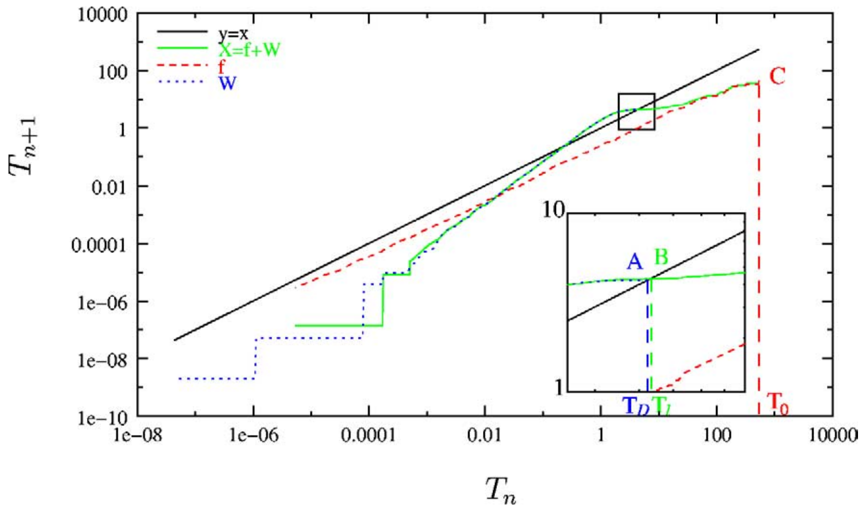

Fig. 2. Iteration functions $I_{W,N}$, $I_{S,N}$, $I_{X,N}$ for W, S, and X, respectively. The points **A** and **B** correspond to the intersections between the graphs of $I_{W,N}$ and $I_{X,N}$ with the line $y = x$, respectively. The point **C** corresponds to the first iteration of the algorithm applied to the noisy signal X and its abscissa is T_0.

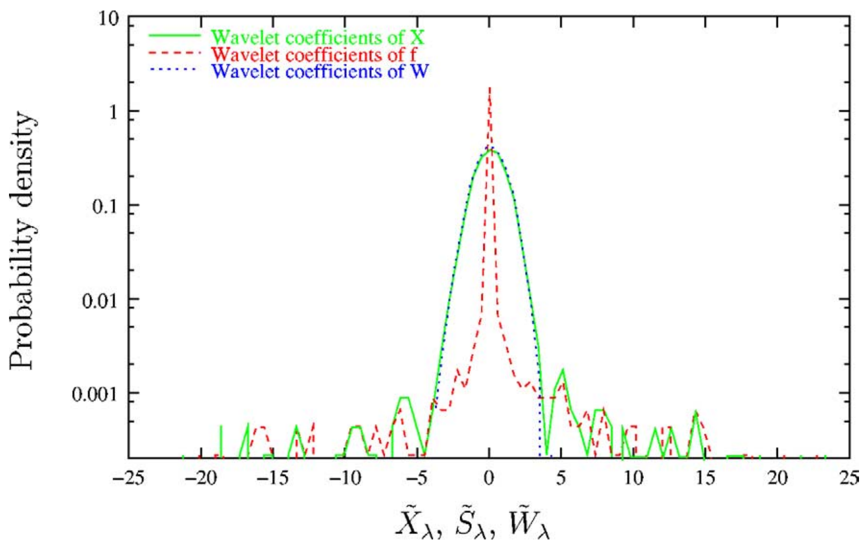

Fig. 3. Histograms of the wavelet coefficients $\tilde{X}_\lambda$, $\tilde{S}_\lambda$, and $\tilde{W}_\lambda$ for the 1D signal.

respectively, are shown in Fig. 1. Table 1 summarizes the values of the threshold T_ℓ, T_m, and T_D, and the resulting mean square errors of the estimations $\mathcal{E}(T_\ell)$, $\mathcal{E}(T_m)$, and $\mathcal{E}(T_D)$ defined in (2).

We observe that, despite the fact that the threshold T_m is closer to T_D than T_ℓ and the error $\mathcal{E}(T_m)$ is smaller than the error $\mathcal{E}(T_\ell)$, the performances of the two methods are of same order. Moreover, the threshold T_D results in a larger error than the thresholds T_m and T_ℓ.

We also observe that the number of iterations n_ℓ is increasing with the signal to noise ratio, i.e., $n_\ell = 1$ for the noise without signal, $n_\ell = 4$ for the noisy signal X, and $n_\ell = 21$ for the signal without noise.

Table 1
Thresholds T_ℓ, T_m, and T_D and the corresponding mean square estimation errors

Signal	n_ℓ	T_ℓ	T_m	T_D	$\mathcal{E}(T_\ell)$	$\mathcal{E}(T_m)$	$\mathcal{E}(T_D)$
X	4	4.34	4.19	4.25	7.28×10^{-3}	7.06×10^{-3}	7.32×10^{-3}
S	21	1.7×10^{-6}	9.9×10^{-7}	0	4.7×10^{-14}	8.9×10^{-16}	0
W	1	4.24	4.19	4.24	$+\infty$	$+\infty$	$+\infty$

We interpret this result by saying that the wavelet coefficients of the noise are responsible for deflecting the graph of $I_{X,N}$ above the line $y = x$. This deflection interrupts the sequence of iterations by forcing the decreasing sequence of thresholds T_n to converge to the intersection point **B**.

The numerical cost of the recursive algorithm is $n_\ell N$ operations, which is, e.g., equal to $4N$ for the case above, since N multiplications and sums are needed at each iteration. The MAD method needs to perform a quick sort on the squared wavelet coefficients, which has a cost of order $N \log N$ plus N multiplications. Both methods additionally require a wavelet transform and its inverse, with order N complexity.

We conclude that the recursive algorithm may run more quickly than the MAD method for weak signal to noise ratios, since the deflection of the iteration function occurs closer to the initial value of the threshold T_0 which speeds up the convergence. Current work [1] also shows that the algorithm yields better results than the MAD method, when applied to signals corrupted with non-Gaussian noise. These additional results are currently being investigated and will be the object of a future publication.

Acknowledgments

We thank Victor Wickerhauser for useful comments and for proposing an alternative algorithm to determine the threshold. By sorting the wavelet coefficients $\tilde{X}_\lambda$ in increasing order of absolute value, the threshold can be determined directly by comparing T_n^2 with $\tilde{X}_n$ for $n = 1, \ldots, N$. The loop terminates at the smallest n for which T_n^2 is at least as big as the n smallest wavelet coefficients and yields the threshold. This algorithm takes $O(N \log N)$ operations to rearrange the coefficients, while our algorithm has a complexity of $O(n_\ell N)$ with $n_\ell \leqslant N$. Therefore, if the number of iterations n_ℓ is larger than $\log N$, a sorting of the wavelet coefficients becomes more efficient. We acknowledge financial support from contract CEA–Euratom No. V. 3258.001.

References

[1] A. Azzalini, M. Farge, K. Schneider, A recursive algorithm for nonlinear wavelet thresholding: Applications to signal and image processing, NAI Note No. 43, Institut Pierre Simon Laplace, March 2004.
[2] D. Donoho, I. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika 81 (1994) 425–455.
[3] S.M. Berman, Sojourners and Extremes of Stochastic Processes, Wadsworth, Reading, MA, 1989.
[4] M. Farge, K. Schneider, N. Kevlahan, Non-Gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptive orthogonal wavelet basis, Phys. Fluids 11 (8) (1999) 2187–2201.
[5] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1998.
[6] R. von Sachs, J. Neumann, Wavelet thresholding: Beyond the Gaussian I.I.D. situation, in: A. Antoniadis, G. Oppenheim (Eds.), Wavelets and Statistics, in: Springer Lecture Notes in Statistics, vol. 103, 1995, pp. 301–329.
[7] R. von Sachs, K. Schneider, Wavelet smoothing of evolutionary spectra by non-linear thresholding, Appl. Comput. Harmon. Anal. 3 (1996) 268–282.