16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

WAVELET DENOISING BASED ON LOCAL REGULARITY INFORMATION

Antoine Echelard and Jacques Lévy-Véhel

Projet APIS INRIA Saclay - Île-de-France
Parc Orsay Université, 4 rue J. Monod, 91893 Orsay Cedex, France
antoine.echelard@gmail.com, jacques.levy-vehel@inria.fr

ABSTRACT

We present a denoising method that is well fitted to the processing of extremely irregular signals such as (multi)fractal ones. Such signals are often encountered in practice, *e.g.*, in biomedical applications. The basic idea is to estimate the regularity of the original data from the observed noisy ones using the large scale information, and then to extrapolate this information to the small scales. We present theoretical results describing the precise properties of the method. Numerical experiments show that this denoising scheme indeed performs well on irregular signals.

1. INTRODUCTION AND MOTIVATIONS

Signal/image denoising is an important task in many applications including biology, medicine, astronomy, geophysics, and many more. For such applications and others, it is important to denoise the observed data in such a way that the features of interest to the practitioner are preserved. In several cases, local irregularity has been shown to be such a feature of interest. A particular example is provided by the study of RR intervals from ECG. It has been shown *e.g.* in [7, 8] that such signals are multifractal, *i.e.* everywhere irregular with an irregularity that varies rapidly in time. More importantly, the degree of multifractality is strongly correlated with the condition of the heart. For this and other applications, denoising should clearly be performed in such a way that the local regularity is controlled through the process.

A popular set of denoising methods is based on decomposing the corrupted signal in a wavelet basis, processing the wavelet coefficients, and then going back to the time domain. In the case of additive white noise, this is justified by two fundamental facts: first, many real-world signals have a sparse structure in the wavelet domain, *i.e.* a few coefficients are significant, and most are small or zero. Second, an orthonormal wavelet transform turns white noise into white noise, so that all wavelet coefficients of a white noise are statistically equal. Denoising in the wavelet domain thus allows to separate in an easy way "large", significant coefficients, from "small" ones due mainly to noise.

The first and simplest methods for denoising based on the above principles are the so-called hard and soft-thresholding [2, 3]. Since the time these methods were introduced, a huge number of improvements have been proposed, ranging from block thresholding [11] to Bayesian approaches [12] and many more.

A limitation of most of these methods is that they are not well adapted to highly textured or everywhere irregular signals, in particular (multi)fractal or multifractional ones with possibly rapidly varying local regularity, as are RR intervals alluded to above. It is in particular well-known that, when the original signal is itself irregular, most wavelet-based denoising methods will typically produce an oversmoothed signal and/or so-called "ringing" effects. Indeed, as recalled above, the basic idea behind wavelet thresholding is that many real-world signals have a sparse wavelet representation, with few large wavelet coefficients. Putting small coefficients to 0 in the noisy signal will then in general do no harm, since these are mainly due to noise. Everywhere irregular signals, on the other hand, have significant coefficients scattered all other the time-frequency plane. At high frequencies, these significant but relatively small coefficients in the signal crucially determine the local irregularity. Zeroing small coefficients will thus typically destroy the regularity information. As a consequence, it is no surprise that a specific method has to be designed for such signals.

To be more precise, let us consider a simple example where the kind of effect we are talking about occurs. We shall consider a noised version of a Weierstrass function. Recall that the Weierstrass function is defined as:

$$W(x) = \sum_{n=0}^{\infty} \lambda^{-nh} \sin(\lambda^n x),$$

where $\lambda \geq 2, h \in (0,1)$. This function is often considered in fractal analysis, and it has also been used in various applications, *e.g.* as a model for the sea surface [13]. It has Hölder exponent h at each point (see below for definitions), so that, the smaller h is, the more irregular the graph of W looks like. On figure 1 is displayed a Weierstrass function sampled on 2^{18} points and with $h = 1/2$. This function is then corrupted with additive white Gaussian noise (not shown here). The noisy version is processed with hard-thresholding of the wavelet coefficients, where the threshold is chosen so as to minimize the L^2 error with the (known) original signal. On the bottom of the figure is displayed the denoised function. The wavelet coefficients of the original, noisy, and denoised signals are shown on figure 2.

One sees on these graphs that the signal is oversmoothed. This is easily explained if one looks at the wavelet coefficients: at each scale $j = 1 \ldots n$, the wavelet coefficients of W are of the order of $2^{-j(h+0.5)}$ or smaller. When j gets large, *i.e.* at small scales, they thus become negligible, on average, with respect the ones of the noise. More precisely, there exists a scale j_0, depending on the standard deviation of the noise, such that, for $j > j_0$, the structure of the signal at these scales is lost because of the noise (see below for a mathematical statement). Optimizing the threshold for best L^2 reconstruction will typically lead to putting the coefficients at scales $j > j_0$ to zero. This is apparent in the bottom parts of the figures. The denoised signal will thus have no high-frequency structure, in contrast to the original one. But it is precisely the small-scale coefficients that produce

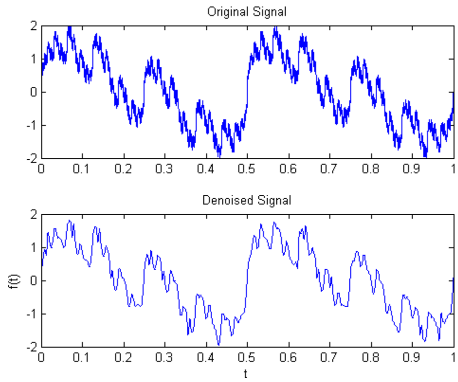

Figure 1: Weierstrass function with $h = 0.5$ sampled on 2^{18} points (top) and denoised version using hard-thresholding (bottom). Here and below we use Daubechies 4.

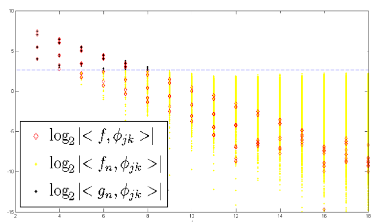

Figure 2: Logarithms of wavelet coefficients for the Weierstrass function: original(f), noisy (f_n), and denoised (g_n) signals (each column corresponds to a scale). Coefficients equal to 0 are not represented. Threshold is indicated by the horizontal line.

the impression of texture, or irregularity, as is exemplified on figure 3, where we have represented graphs of functions whose wavelet coefficients differ on specific ranges of scale.

Our approach to tackle this problem is based on a recent paradigm where one tries to use *a priori* information on the signal to enhance the result of denoising through thresholding. A clear statement of this paradigm is given in [1]:

"Given an imperfectly described signal, it is often the case that a few of its parameters are given with good precision whereas other parameters are known only vaguely or are *a priori* essentially unknown. If one believes, however, that the unknown parameters are somehow correlated with the known ones, then it is reasonable to try to extrapolate the unknown parameters from the available ones. This involves exploitation of additional, external principles of our choice, the ones that are believed to express relations between the two groups of parameters."

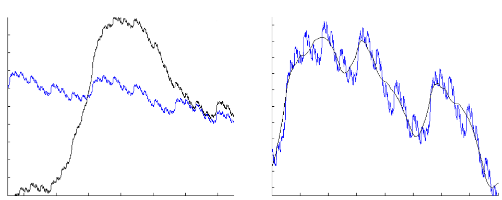

Figure 3: Left: two functions with wavelet coefficients differing only at large scales. These two functions display the same aspect in term of regularity. Right: Two functions with wavelet coefficients differing only at small scales. These two functions display different aspects in term of regularity.

When a signal is corrupted by white noise, the information carried by a wavelet coefficient of the original signal whose absolute value is smaller than the standard deviation of the noise is essentially lost. Instead of setting to zero "small" coefficients in the noisy signal as is done in hard-thresholding, one tries to extrapolate their value from the robust "large" ones. Several authors have put this paradigm to use. For instance, in [4], denoising is performed in two steps: a) Apply a classical denoising method, such as hard thresholding. The non-thresholded coefficients are stored in a set M. b) Estimate the original coefficients from the thresholded noisy ones by minimizing the total variation under the constraint that the coefficients in M are left unchanged.

In other words, instead of putting the small coefficients to 0, the method adjusts them in such a way that the total variation is minimized. This approach does suppress most of the "ringing", but it tends to be slow and to oversmooth textured parts.

Coifman et Sowa [1] have proposed a general formulation that applies not only to thresholding, but also to other kinds of processing such as quantification often present in compression applications. It may be described as follows: Let J be the set of indices for which the wavelet coefficients are smaller than the threshold, and let I be the complementary set. The resulting signal $\tilde{f}_n$ will verify the following global constraints: the coefficients of $\tilde{f}_n$ whose indices belong to I will remain unchanged. The other ones are chosen in such a way that they remain smaller in absolute value than a given constant (typically, the threshold), and that an "energy" $\Phi(\tilde{f}_n)$ is minimized. The functional Φ must verify certain properties, and depends on the specific application.

In contrast to the two methods just described (and to most other approaches), our implementation of the above paradigm consists in extrapolating the unknown, small, coefficients by imposing a *local* constraint rather that a *global* one. More precisely, we set the small coefficients in such a way that the local regularity at each point of the denoised signal matches the one of the original one. Of course, since the original signal is unknown, so is its regularity. We will thus show how to estimate the local regularity of the original signal from the noisy observations. Although local in time (*i.e.* specific to each point), this information has some robustness, as it is global in scale (it is computed from coefficients at several scales). It is precisely through these regularity estimates that one uses the information present in large coefficients to extrapolate the small ones.

The remainder of this paper is organized as follows. In the next section, we recall some basic facts concerning the connection between local regularity and wavelet coefficients. In section 3 we describe our denoising scheme. For this method to work, we need to estimate the local regularity of the original signal from the noisy observations. Section 4 explains how to do this. Finally, numerical experiments are displayed in section 5.

2. WAVELETS AND LOCAL REGULARITY

We shall measure the local regularity in terms of *pointwise Hölder exponents*. This exponent is defined at each point of a locally bounded function f as:

$$\alpha_f(t) = \sup\left\{\alpha, \limsup_{h\to 0} \frac{|f(t+h)-f(t)|}{|h|^\alpha} = 0\right\}.$$

(This definition is valid only for $\alpha \in (0,1)$. For α larger than 1, one has to replace the term $f(t)$ by a polynomial. We shall ignore this complication for now, as the wavelet-based characterization we will use below takes care of it in a transparent manner.)

When there is no risk of confusion, we shall write $\alpha(t)$ in place of $\alpha_f(t)$. Let us explain the geometrical meaning of α. Roughly speaking, saying that a function f has exponent α at t_0 means that, around t_0, the graph of f "looks like" the curve $t \mapsto f(t_0) + c|t-t_0|^\alpha$ in the following sense: for any positive ε, there exists a neighbourhood of t_0 such that the path of f inside this neighbourhood is included in the envelope defined by the two curves $t \mapsto f(t_0) + c|t-t_0|^{\alpha-\varepsilon}$ and $t \mapsto f(t_0) - c|t-t_0|^{\alpha-\varepsilon}$, while this property is no longer true for any negative ε (see figure 4). A "large" α means that f is smooth at t_0, while an irregular behaviour of f at t_0 translates into α close to 0.

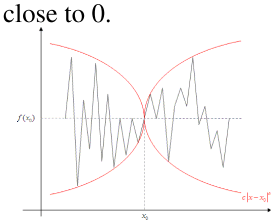

Figure 4: Graphical interpretation of the Hölder exponent.

The pointwise Hölder exponent is not the only way to measure the local regularity of a function. Many other exponents exist, that give complementary information. In the following, we shall assume that the considered signal satisfy a technical assumption to the effect that its pointwise exponent is equal to its *local Hölder exponent*, defined as:

$$\alpha_f^l(t) = \sup\left\{\alpha, \limsup_{h\to 0} \sup_{(x,y):|t-x|<h,|t-y|<h} \frac{|f(x)-f(y)|}{|x-y|^\alpha} = 0\right\}.$$

Such an assumption is for instance verified by the Weierstrass function and by many fractal, multifractal and multifractional signals (see [10]). In addition, our method may be generalized to functions that do not verify it. This however requires lengthy developments which cannot be described properly here. See [5] for details.

When $\alpha_f = \alpha_f^l$, their common value may be estimated in a simple way with the help of wavelet analysis, as we recall now. In the following, we shall assume for simplification that all the considered functions are compactly supported on $[0,1]$. We shall always consider a wavelet ψ such that the set of functions $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, for $(j,k) \in \mathbf{Z}$, form an orthonormal basis of $L^2(\mathbf{R})$. Finally, ψ is supposed to have the necessary number of vanishing moments and required decay properties.

Recall that a function f is said to belong to the global Hölder space $C^\beta(\mathbb{R})$ if there exists a constant C such that, for all x, y, $|f(x)-f(y)| \le C|x-y|^\beta$. The supremum of those β such that f belongs to $C^\beta(\mathbb{R})$ is called the global Hölder exponent of f, and is denoted $\alpha_g(f)$.

The most basic results relating global regularity and wavelet coefficients is:

Proposition 2.1. *[9] Assume $f \in C^\varepsilon(\mathbb{R})$ for some $\varepsilon > 0$ has bounded support. Then:*

$$1/2 + \alpha_g(f) = \liminf_{j\to\infty} \min_{k\in\mathbb{Z}} \frac{\log|\langle f, \psi_{jk}\rangle|}{-j}$$

When $\alpha_f^l(t) = \alpha_f(t)$, the above result may be "localized" by considering only those coefficients lying "above" t: one then gets that the pointwise Hölder exponent of f at t is given by a liminf similar to the one in (2.1), except that only relevant coefficients are considered. See [5, 10] for more details.

3. WAVELET BASED DENOISING WITH CONTROL OF THE LOCAL REGULARITY

3.1 Local regularity of hard-thresholded signals

In the first section, we saw that hard-thresholding wavelet coefficients produce the visual impression of oversmoothing irregular signals (figure1). In this section, we shall give a precise mathematical meaning to this fact.

A difficulty arises since we are working on discrete signals: indeed, the very definition of Hölder exponents requires to let the scale tend to 0, which cannot be done here. We thus need a modified definition of α that both makes sense at finite resolution and allows to capture the visual impression of regularity on sampled signals, such as on figure 3 . In view of the fact that the perceived regularity depends on the considered range of scales, it seems natural to define an "exponent between two scales" as follows:

$$\alpha_g(j_1, j_2, f) = \min_{j\in[j_1,j_2]} \min_{k\in\mathbb{Z}} \frac{\log|\langle f, \psi_{jk}\rangle|}{-j} - 1/2$$

In addition to reflecting the regularity of a function at the considered scales, this definition has adequate asymptotic properties: it is not hard to prove that, for a given function f, $\alpha_g(h(n), n, f)$ tends to $\alpha_g(f)$ when n tends to infinity, provided the function h tends sufficiently slowly to infinity. In addition, this property still holds (with minor additional constraints) if one replaces f with its sampled approximation at scale n [5]. The exponent between two scales thus approximates the genuine regularity for large enough n. Furthermore, it behaves in an intuitive way when computed on a noisy signal:

Theorem 3.1. *Let f belong to $C^{\varepsilon_0}(\mathbb{R})$ for some $\varepsilon_0 > 0$. Let g_n be a sampling on 2^n points of $f + B$, where B is a white Gaussian noise. Then, for any increasing function $h(n) \le n$ tending to $+\infty$, $\alpha_g(h(n), n, g_n)$ tends almost surely to 0.*

In other words, for any f with minimal regularity, the exponent between scales $h(n)$ and n will tend to 0 for any sensible starting scale h. This result and all subsequent ones are proved in [5]. The following theorem gives a precise meaning to the statement that "hard thresholding tends to oversmooth signals":

Theorem 3.2. *Let f belong to $C^{\varepsilon_0}(\mathbb{R})$ for some $\varepsilon_0 > 0$. Let g_n be the signal obtained by sampling at scale n a version of f corrupted with standard white Gaussian noise. Let $\tilde{f}_n$ be the signal denoised by hard-thresholding using the universal threshold $\lambda_n = 2^{-n/2}\sqrt{2n\ln 2}$, i.e.:*

$$<\tilde{f}_n> = \begin{cases} \langle g_n, \psi_{jk}\rangle & \text{if } \left|\langle g_n, \psi_{jk}\rangle\right| > \lambda_n \\ 0 & \text{if } \left|\langle g_n, \psi_{jk}\rangle\right| \le \lambda_n \end{cases}$$

Let θ be an increasing function such that $\alpha_g(f) = \lim_{n\to\infty} \alpha_g(\theta(n), n, f)$.

1. *Assume the function h verifies:*

$$\exists \varepsilon > 0, \forall n \in \mathbb{N}, \exists i \in \mathbb{N} : \theta(i) \in \left[h(n) \ldots \frac{n}{1+2\alpha_g(f)}(1-\varepsilon) \right]$$

then $\alpha_g(h(n), n, \tilde{f}_n)$ tends in probability to $\alpha_g(f)$.

2. *Assume h verifies:*

$$\exists \varepsilon > 0 : \forall n \in \mathbb{N} : h(n) \geq \frac{n}{1+2\alpha_g(f)}(1+\varepsilon)$$

then $\alpha_g(h(n), n, g_n)$ tends in probability to $+\infty$.

Intuitively, this results means that, when one observes the signal from "far away", (*i.e.* for $h(n)$ smaller than $\frac{n}{1+2\alpha_g(f)}$), the denoised signal has the same perceived regularity as the one of the original signal, while when one looks at fine details (case 2), one sees an infinitely smooth signal. This is exactly what figure 1 suggests.

3.2 A denoising scheme that maintains the local regularity

Our aim is to modify the hard-thresholding scheme so that the regularity of the denoised signal, as measured by the exponent between two scales, matches the one of the original signal. As we have seen in the case of the Weierstrass function, the problem is that, below a critical scale c, which is such that $2^{-c(\alpha+0.5)}$ is of the order of the standard deviation of the noise σ, hard-thresholding will put to 0 informative coefficients. While this is not a serious problem for signals with a sparse wavelet representation (for which the method was originally conceived), it becomes so for "fractal" signals. We thus want to maintain some information at small scale. This cannot be done using only the observed coefficients, since, for $j > c$, in most cases, their value will be strongly influenced by the noise. Following the paradigm explained in the introduction, we will rather extrapolate those coefficients from the large, robust, ones at larger scales. More precisely, we follow the steps below;

- Estimate the critical scale c_n, defined as the one where the coefficients of the white noise become predominant as compared to the ones of the signal.
- Estimate the regularity s_n of the original signal at the considered point, using coefficients at scales larger than c_n.
- Assign to the small scale coefficients a value that is "coherent" with the ones of the coefficients at larger scales:

$$\langle \tilde{f}_n, \psi_{jk} \rangle = \min\left(\left| \langle g_n, \psi_{jk} \rangle \right|, 2^{K_n - j(s_n+1/2)} \right) \operatorname{sgn}(\langle g_n, \psi_{jk} \rangle)) \quad (1)$$

for $j > c_n$, where K_n and s_n are estimated from the wavelet coefficients at scales $j < c_n$ (see section 4).

This means that, at small scales, we do not accept too large coefficients, that is ones which would not be compatible with the estimated Hölder regularity of the signals (statistically, there will always be such coefficients, since the noise has no regularity: its coefficients do not decrease with scale). On the other hand, "small" coefficients (ones not exceeding $2^{K_n - j(s_n+1/2)}$) are left unchanged. Note that both the estimated regularity s_n and the critical scale c_n depend on the considered point. Remark also that this procedure may be seen as a location-dependent shrinkage of the coefficients.

One can prove the following property, which essentially says that the above method does a good job in recovering the regularity of the original signal, provided that one is able to estimate with good accuracy its Hölder exponent at any given point t:

Proposition 3.3. *Let f belong to $C^{\varepsilon_0}(\mathbb{R})$ for some $\varepsilon_0 > 0$, and let α denote its Hölder exponent at point t.*

Let $(s_n)_n$ be a sequence of real numbers tending almost surely (resp. in probability) to α. Let $c(n) = \frac{n}{1+2s_n}$Let $\tilde{f}_n$ be defined as above.

Then, for any function h tending to infinity with $h(n) \leq n$, $\alpha_g(h(n), n, \tilde{f}_n)$ tends almost surely (resp. in probability) to α.

For our method to be put to practical use, there thus remains to estimate the critical scale and Hölder exponent from the noisy observations. This is the topic of the next section.

4. ESTIMATING THE LOCAL REGULARITY OF A SIGNAL FROM NOISY OBSERVATIONS

Our main result concerning the estimation of the critical scale is the following one.

Theorem 4.1. *Let $(x_i)_{i\in\mathbb{N}}$ denote the wavelet coefficients of $f \in C^{\varepsilon_0}(\mathbb{R})$ "above" (i.e such that $k = [2^{-j}]$) a point t where the local and pointwise Hölder exponent of f coincide. Let $\beta = \liminf_{i\to\infty} \frac{-\log|x_i|}{i}$. Assume that there exists a decreasing sequence (ε_n) such that $\varepsilon_n = o\left(\frac{1}{n}\right)$ when $n \to \infty$ and $\frac{-\log|x_i|}{i} \geq \beta - \varepsilon_i$, for all i. Let (y_i) denote the noisy coefficients corresponding to the x_i.*

Let:

$$\mathscr{L}_n(p) = \frac{1}{(n-p+1)^2} \sum_{i=p}^{n} y_i^2,$$

and denote $p^* = p^*(n)$ an integer such that:

$$\mathscr{L}_n(p^*) = \min_{p:1\leq p\leq n-b\log(n)} \mathscr{L}_n(p),$$

where $b > 1$ is a fixed number. Let finally $q(n) = \frac{n}{2(\beta-\frac{1}{n})}$. Then, almost surely:

$$\forall a > 1, \quad p^*(n) \leq q(n) + a\log(n), \quad n \to \infty \quad (2)$$

In addition, if the sequence (x_i) verifies the following condition: there exists a sequence of positive integers (θ_n) such that, for all n large enough and all $\theta \geq \theta_n$:

$$\frac{1}{\theta} \sum_{i=q-\theta}^{q-1} x_i^2 > b\sigma_n^2 \frac{1-\frac{\delta_*}{\beta}}{(1-\frac{\delta^*}{\beta})^2},$$

where $\delta_* \in (0, \frac{1}{2})$ and $\delta^* \in (\frac{1}{2}, \beta)$.

Then, almost surely:

$$\forall a > 1, \quad p^*(n) \geq q(n) - \max(a\log(n), \theta_n), \quad n \to \infty \quad (3)$$

In other words, when the conditions of the theorem are met, any minimizer of $\mathscr{L}$ is, within an error of $O(\log(n))$, approximately equal to the searched for critical scale. This

16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

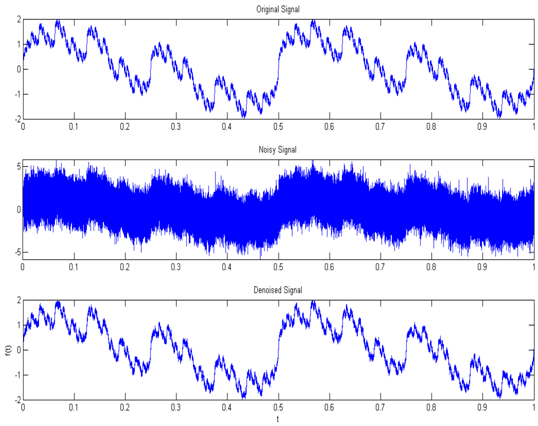

Figure 5: Top: original Weierstrass function. Middle: noisy version. Bottom: signal obtained with the regularity preserving method.

allows in turn estimation of the Hölder exponent through the next corollary:

Corollary With the same notations and assumptions as in the theorem above, with the additional condition that θ_n is not larger than $b\log(n)$ for all sufficiently large n, define:

$$\hat{\beta}(n) = \frac{n}{2p^*(n)} + \frac{1}{n}.$$

Then the following inequality holds almost surely for all large enough n:

$$|\hat{\beta}(n) - \beta| \leq 2b\beta^2 \frac{\log(n)}{n}.$$

We thus set $s_n = \hat{\beta}(n) + \frac{1}{2}$ in (1). K_n is estimated as the offset in the linear least square regression of the logarithm of the absolute value of the wavelet coefficients with respect to scale, at scales larger than $p^*(n)$.

5. NUMERICAL EXPERIMENTS

We show the result of the denoising scheme described above on the Weierstrass function considered in the first section. Figure 5 shows the original, noisy and denoised versions.

As a further evaluation of the behaviour of the method, we have estimated the regularization dimension of the various signals involved. The regularization dimension is a kind of fractal dimension that is well adapted to signal processing, and which is equal to the more familiar box dimension for the (noise-free) Weierstrass function. Here are the results:

- Original function: 1.50 (theoretical dimension: 1.5).
- Noised version: 2.47 (theoretical dimension: 2.5).
- Denoising using hard-thresholding with threshold optimized for best L^2 reconstruction: 1.08.
- Denoising using the proposed method: 1.57.

Finally, we plot on figure 6 the estimated Hölder exponents of the same functions, using an estimator based on oscillations. One clearly sees that the proposed denoisng scheme has superior performances in this respect. Note that neither the regularization dimension nor the pointwise Hölder exponents were estimated through procedures involving wavelets.

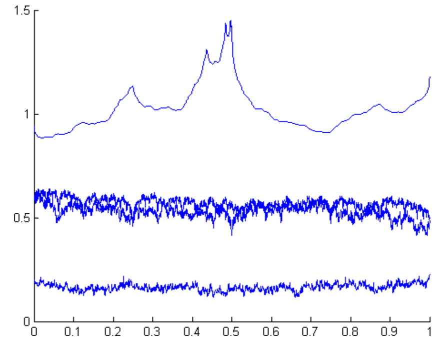

Figure 6: Estimated Hölder exponents. Top curve: signal obtained through hard-thresholding. Middle curves: original Weierstrass function and version denoised using the proposed scheme. Bottom curve: Noisy signal.

Codes for replicating all the experiments displayed in this work are freely available in the *FracLab* toolbox [6].

As a final note, we remark that our method works relatively well on signals with sufficiently many samples: more precisely, a reliable estimation of the Hölder exponent requires that the number of scales n be significantly larger than its logarithm. While sample sizes of 2^{15} or more are common in applications such as biomedicine, finance or Internet traffic analysis, it is not the case for image processing. Though our method has a straightforward extension in higher dimensions, practical implementation would only give good results on extremely high resolution images.

REFERENCES

[1] R. R. Coifman and A. Sowa, "Combining variational calculus and wavelets for image enhancement", *Appl. Comput. Harmon. Anal.*, vol. 9, pp. 1–18, 2000.

[2] R.A. Devore and B. Lucier, "Fast wavelet techniques for near optimal image processing", *IEEE Military Communications Conference*, vol. 2 (12), 1992.

[3] D.L. Donoho, "De-noising by soft-thresholding", *IEEE Trans. Inf. Theory*, vol. 41 (3), pp. 613–627, 1994.

[4] S. Durand and J. Froment, "Reconstruction of wavelet coefficients using total variation minimization", *SIAM Journal of Scientific Computing*, vol. 24(5), pp. 1754–1767, 2003.

[5] A. Echelard, "Analyse 2-microlocale et application au débruitage", *Ph.D. Thesis, Universite de Nantes*. December 2007.

[6] Fraclab: a software toolbox for fractal processing of signals. http://www2.irccyn.ec-nantes.fr/FracLab/

[7] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C.K. Peng, H. E. Stanley, "Fractal dynamics in physiology: Alterations with disease and aging", *PNAS*, vol. 99, pp. 2466–2472, 2002.

[8] C. Ivanov, L.Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z. R. Struzik, H. E.Stanley, "Multifractality in human heartbeat dynamics", *Nature*, vol. 399, June 1999.

[9] S. Jaffard and Y. Meyer, *Wavelet methods for pointwise regularity and local oscillations of functions*, American Mathematical Society, 1996.

[10] J. Lévy Véhel and S. Seuret, "The 2-microlocal formalism", *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot, Proc. Sympos. Pure Math*, 2004.

[11] W. Härdle, G. Kerkyacharian, D. Picard and A. Tsybakov, *Wavelets, approximation, and statistical applications*, Springer, 1998.

[12] B. Vidakovic, *Statistical Modeling by Wavelets*, Wiley, 1999.

[13] Y. Yan, *Back Scattering from Fractal Surface of Sea*, Int. J. Infrared Millimeter Waves, vol. 21 (6), pp. 979–985, June 2000.