

Shift-Invariant Gibbs Free Denoising Algorithm based on Wavelet Transform Footprints

Pier Luigi Dragotti¹ and Martin Vetterli^{1,2}

¹ Laboratoire de Communications Audiovisuelles (LCAV)
École Polytechnique Fédérale de Lausanne (EPFL)
CH - 1015 Lausanne, Switzerland.

email: dragotti,vetterli@lcavsun1.epfl.ch,
tel: +41 21 693 7663, fax: +41 21 693 4312.

²Electrical Engineering and Computer Science Department,
University of California at Berkeley, Berkeley CA 94720

ABSTRACT

In recent years wavelet have had an important impact on signal processing theory and practice. The effectiveness of wavelets is mainly due to their capability of representing piecewise smooth signals with few non-zero coefficients. Away from discontinuities, the inner product between a wavelet (with a number of zero moments) and a smooth function will be either zero or very small.⁸ At singular points, a finite number of wavelets concentrated around the discontinuity lead to non-zero inner products. This ability of wavelet transform to pack the main signal information in few large coefficients is behind the success of wavelet based denoising algorithms. Indeed, traditional approaches simply consist in thresholding the noisy wavelet coefficients, so the few large coefficients carrying the essential information are usually kept while small coefficients mainly containing noise are cancelled. However, wavelet denoising suffers of two main drawbacks: it is not shift-invariant and it exhibits pseudo Gibbs phenomenon around discontinuities.

In this work, we present a new denoising algorithm which does not present the pseudo Gibbs phenomenon and which is almost shift-invariant even if we do not use a frame expansion. In our analysis we focus on piecewise polynomial functions. For this class of signals we know that, if wavelets have enough vanishing moments, away from discontinuities the wavelet coefficients are exactly zero. Moreover the wavelet coefficients generated by a discontinuity are highly dependent across scales. Therefore, a good denoising algorithm should take advantage of this dependency. We thus introduce the notion of *footprints*, which are the traces left by time domain singularities in the wavelet domain. So a footprint is a vector containing all the significant wavelet coefficients generated by a singularity.

In our denoising algorithm instead of thresholding the noisy wavelet coefficients independently as in traditional approaches (*scalar thresholding*), we gather all the noisy coefficients around a discontinuity in a vector and compute the inner product with the closest footprint that represents that discontinuity. If this product is large enough we keep it, otherwise we cancel it (*vector thresholding*). With this approach we substantially improve the performance of traditional denoising algorithms in terms of SNR and we eliminate the pseudo Gibbs phenomenon since we force the denoised signal to belong to the class of piecewise polynomial functions.

The proposed algorithm can potentially be extended to more general signals like piecewise smooth signals and represents an effective solution to problems like signal denoising.

1. INTRODUCTION

Wavelets as time and frequency localized bases are particularly suited for the representation of piecewise smooth functions. That is, wavelets are able to represent such functions with few coefficients, thanks to the localization properties inherent to wavelets. This is in contrast with the Fourier bases, which are inadequate when discontinuities are present. This basic intuition is behind many of the successes of wavelets in applications, including compression or denoising.

In the present paper, we show that it is possible to improve wavelet methods even further, exactly in the case of piecewise polynomial, and approximately in the case of piecewise smooth functions. This is done by performing a

more precise analysis of the behaviour of wavelet coefficients around discontinuities and using a new data structure we call wavelet footprints.

This is a data structure that efficiently represents time domain discontinuities in the wavelet domain. Together with scaling functions which represent the polynomial parts, wavelet footprints can thus efficiently represent piecewise polynomial signals. More generally, piecewise smooth signals are well approximated by scaling functions and footprints.

The motivation for this work came from earlier work in wavelet denoising³ as well as in representation of piecewise polynomial signals.^{6,7} The large gap between the performance of oracle based or dynamic programming based methods in⁷ and the more usual wavelet methods, made us wonder if there was a way to improve the latter methods. This lead both to an understanding of the performance loss, and to the introduction of wavelet footprints, which close the gap and allow wavelet methods to perform close to optimality.

In case of denoising, the usual wavelet based algorithm consists in transforming the noisy signal into the wavelet domain, then in subjecting the wavelet coefficients to soft or hard thresholding, finally in inverse transforming.³ Despite its simplicity this approach has shown to be very efficient and has near optimal theoretical properties.³ However, this method suffers of a main drawback, it presents pseudo Gibbs effects around discontinuities. These artifacts are due to the shift-variance of the orthogonal wavelet transform and to the fact that wavelet coefficients are thresholded independently, so the dependency between wavelet coefficients is not exploited. An algorithm that partially reduces these effects is the *Translation Invariant Thresholding* proposed by Coiffman and Donoho.¹ This algorithm basically consists in applying a thresholding algorithm to the original signal and to all its translated versions and in averaging them; this is similar to using a frame expansion instead of an orthogonal wavelet transform.

In this paper we propose a new algorithm based on the notion of footprint and we show that it eliminates the pseudo Gibbs effects, even though we do not use a frame expansion. Moreover, this system performs better than traditional methods also in terms of mean square error and its performance is closer to the ideal bound.

In next Section, we introduce the notion of footprint, while in Section 3 we clarify the relationship between footprints and piecewise polynomial signals. Our denoising algorithm is presented in Section 4 and performance are assessed in Section 5. We conclude in Section 6.

2. FROM DISCONTINUITY CHARACTERIZATION TO FOOTPRINTS

We briefly review here the characterization of discontinuities in the wavelet domain. For a more detailed treatment, see.^{4,2}

Assume an orthonormal wavelet series with scale and shift parameters m and n , respectively. We use the convention that small scales correspond to large, negative m , that is:

$$\psi_{m,n}(t) = \frac{1}{2^{m/2}} \psi(2^{-m}t - n) \quad m \in \mathbb{Z}$$

In that case, discontinuities of order k lead to a decay of order $2^{-m(k+1/2)}$ of the wavelet coefficients in the vicinity of the discontinuity for small enough scales.

If the wavelet has compact support of length L , then an isolated discontinuity generates at most $\lfloor L \rfloor$ non-zero coefficients at each scale.

If the wavelet has N zero moments, and the signal is polynomial of degree $D < N$, then the inner product is zero. The polynomial is then exactly represented by the scaling coefficients and its shifts alone.

Finally, a compactly supported orthogonal wavelet with N zero moments is of length $L \geq 2N - 1$. Thus, we just saw the basic tradeoff in choosing a wavelet: many zero moments lead to long wavelets, which in turn lead to many non-zero coefficients around discontinuities. An illustration of the packing of wavelet coefficients and discontinuities is shown in Fig. 1.

The above discussion can be repeated for discrete-time filters and signals, and leads to similar conclusions. If the filter bank is orthogonal and the low-pass filter has N zeros at $\omega = \pi$ (and thus, the high-pass filter has N zeros at $\omega = 0$), then discrete-time polynomials up to degree $N - 1$ are reproduced by the low-pass or its iterates. Discontinuities lead to an order of $L - 1$ non-zero coefficients in the high-pass branch if the filters are of length L .

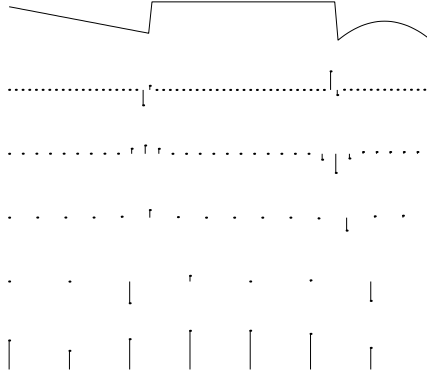


Figure 1. Wavelet decomposition of a piecewise polynomial function. The first 4 level are wavelet coefficients, while the last level are scaling coefficients. Filters: Daubechies four.

For our discussion, it is useful to introduce the following operators which act on $l_2(\mathcal{Z})$. The first is a J level orthonormal discrete time wavelet decomposition denoted W . Given a sequence $x[n]$, it generates $J + 1$ sequences $y_0[n], y_1[n], \dots, y_J[n]$ corresponding to highpass ($y_0[n]$), bandpass ($y_1[n], y_2[n], \dots, y_{J-1}[n]$) and lowpass $y_J[n]$ channels. This operator is linear periodically time-varying, with period 2^J . A time domain discontinuity generates a response of size J (the number scales) by $L - 1$ (the length of the filters after subsampling) which we call a footprint. This footprint depends on the location of the discontinuity, namely, a particular discontinuity has 2^J footprints, depending on its location. From now on $fp_k^{(d)}$ will denote a footprint; k is the discontinuity location in time domain and d the maximum degree of the two polynomials around the discontinuity. Note that given the type of discontinuity and its location, the footprint is a deterministic function.

The other discrete-time operator we need is the frame operator F corresponding to a non-subsampled version of W . In signal processing terms, it corresponds to the filter bank implementing W , but having all subsamplers shifted out (with corresponding equivalent filters). This is the non-subsampled equivalent filter bank.⁹

With J scale decomposition, an input $x[n]$ is transformed into $J + 1$ sequences $f_0[n], f_1[n], \dots, f_J[n]$ corresponding to highpass ($f_0[n]$), bandpass ($f_1[n], f_2[n], \dots, f_{J-1}[n]$) and lowpass $f_J[n]$ channels. Because W is related to F by sub-sampling, we have:

$$y_i[n] = f_i[n2^{i+1}] \quad (1)$$

The frame operator is linear and shift invariant. A discontinuity in the input generates a footprint which is shift-invariant and grows exponentially with J . The location dependent footprints and the operator W are simply the dyadic subsampled versions of the frame footprints, similarly to (1).

3. PIECEWISE POLYNOMIAL SIGNALS AND FOOTPRINTS

From now on we will mainly focus on the W operator and on the corresponding periodically shift-variant footprint. We aim at clarifying the relation between footprints and piecewise polynomial signals. We will first give a geometrical interpretation of the space spanned by piecewise polynomial signals in time domain, then we will move to the wavelet domain and the relation between footprints and piecewise polynomial signals will be explained. In each sub-section we start considering piecewise polynomial functions with only one discontinuity at a fixed position k , then we allow the position to be arbitrary, finally we consider the case of more than one discontinuity.

3.1. Piecewise Polynomial Signals in time domain

Consider piecewise linear signals with one discontinuity at a fixed position k . Clearly, this set of signals forms a linear space of dimension 4. A possible basis for that space is represented by the signals in Fig. 2 and the orthogonalized version is represented in Fig 3. This is not the only possible orthogonal space decomposition but it has the advantage of splitting the smooth and the transient behaviours of the signal in two orthogonal linear subspaces. From now on we will call these two sub-spaces: the *smooth sub-space* and the *transient sub-space*. We call $\{P^{(0)}, P^{(1)}\}$ the two signals that span the smooth sub-space and $\{T_k^{(0)}, T_k^{(1)}\}$ the two related to the transient one.

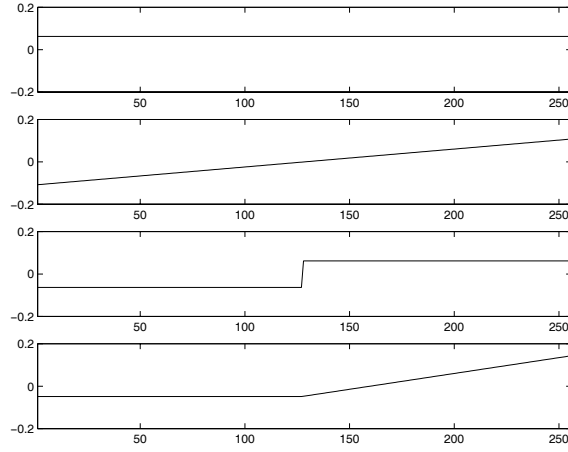


Figure 2. Basis for piecewise linear signal with one discontinuity at k and signal with $N = 256$ samples.

Analogously, the set of piecewise polynomial signals with maximum degree D and with only one discontinuity at a fixed position forms a $2 \cdot (D + 1)$ dimensional linear space. Following the same procedure as in the piecewise linear case, we can decompose this space in two orthogonal sub-spaces of dimension $D + 1$. One is related to the global behaviour of the signal and is spanned by a Legendre polynomials expansion, the other is related to the transient behaviour.

Finally, notice that if we allow the discontinuity to be in a different position then the space spanned by the new piecewise polynomial signal is different than in the other case, but the smooth part of the signal lies on the same sub-space as in the previous case, while the transient contribution forms a different transient sub-space.

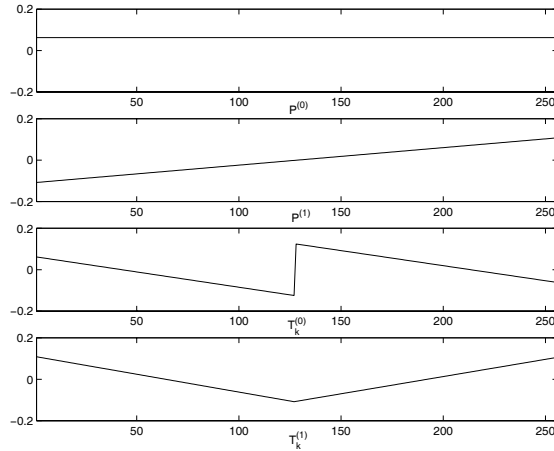


Figure 3. Orthonormal basis for piecewise linear signal with one discontinuity at k and signal with $N = 256$ samples.

In conclusion, we can say that any piecewise polynomial signal of maximum degree D and with several discontinuities can be seen as the combination of a smooth part lying on a fixed $D + 1$ dimensional space and transient parts each of them related to a different discontinuity and lying on different $D + 1$ dimensional sub-spaces.

3.2. Piecewise Polynomial Signals in wavelet domain and footprints

Consider again a piecewise linear signal X with one discontinuity at a fixed position k . Let us perform a J level wavelet decomposition of X ($Y = WX$). Let us perform the same wavelet transform on the four signals that represent a basis for X : $\{P^{(0)}, P^{(1)}, T_k^{(0)}, T_k^{(1)}\}$. Notice that only $T_k^{(0)}, T_k^{(1)}$ have non-zero wavelet coefficients, while

$P^{(0)}, P^{(1)}$ are completely characterized by their scaling coefficients. Let us call $fp_k^{(0)}$ and $fp_k^{(1)}$ the corresponding footprints generated by $T_k^{(0)}, T_k^{(1)}$ and $\hat{fp}_k^{(0)}, \hat{fp}_k^{(1)}$ the orthonormalized versions ($\|\hat{fp}_k^{(d)}\| = 1, d \in \{0, 1\}$ and $\langle \hat{fp}_k^{(0)}, \hat{fp}_k^{(1)} \rangle = 0$).

It clearly results with a little abuse of notation:

$$X = W^{-1} \left(Y_s + \sum_{d=0}^1 \langle Y, \hat{fp}_k^{(d)} \rangle \hat{fp}_k^{(d)} \right), \quad (2)$$

where Y_s are the scaling coefficients of Y and $\langle Y, \hat{fp}_k^{(d)} \rangle$ is the inner product between $\hat{fp}_k^{(d)}$ and the wavelet coefficients of Y located at the same spatial position as the coefficients of $\hat{fp}_k^{(d)}$. Equation (2) shows that the smooth sub-space is specified by the scaling coefficients, while any discontinuity at position k can be expressed in terms of a linear combination of the two elementary footprints $\hat{fp}_k^{(0)}, \hat{fp}_k^{(1)}$. That is $\hat{fp}_k^{(0)}$ and $\hat{fp}_k^{(1)}$ represent again an orthonormal basis for the transient sub-space related to the position k .

Now, consider the case where the discontinuity is located at position l different from k and $(l - k) < 2^J$. Since footprints are periodically shift-variant with period 2^J , a particular discontinuity generates 2^J different footprints depending on its location. So the pair of elementary footprints $\{\hat{fp}_k^{(0)}, \hat{fp}_k^{(1)}\}$ can represent only discontinuity at positions k and $k + m \cdot 2^J, m \in \mathbb{Z}$. To characterize the other discontinuities we need different pairs of elementary footprints. To characterize any discontinuity we need a dictionary of 2^J pairs of footprints: $\{\hat{fp}_l^{(0)}, \hat{fp}_l^{(1)}\} l = 1..2^J$. Let us call this dictionary \mathcal{D} . So if X has a discontinuity at position l then the pair of footprints that is adequate to represent this discontinuity is $\{\hat{fp}_k^{(0)}, \hat{fp}_k^{(1)}\}$ with $k = (l \bmod 2^J)$. From now on, for simplicity, we will keep denoting the elementary footprint related to a position l with $\hat{fp}_l^{(d)}$ even though the right subscript should be: $k = (l \bmod 2^J)$.

Consider, now, the case where the piecewise linear signal X has more than one discontinuity (i.e. two discontinuities at position k_1 and k_2). If these two discontinuities are far enough then the corresponding footprints generated by them do not overlap and we say that footprints are orthogonal and we can write:

$$X = W^{-1} \left(Y_s + \sum_{d=0}^1 \langle Y, \hat{fp}_{k_1}^{(d)} \rangle \hat{fp}_{k_1}^{(d)} + \sum_{d=0}^1 \langle Y, \hat{fp}_{k_2}^{(d)} \rangle \hat{fp}_{k_2}^{(d)} \right).$$

This clarifies the advantage of representing X in terms of footprints instead of a time domain basis. Since footprints give a sparse and local representation of the time domain basis, it is possible to represent a signal X with more discontinuities in an orthogonal basis (the footprints basis), this is not possible with the previously defined time domain basis.

However, if the two footprints overlap then they are not orthogonal and in order to find the right representation of X in terms of our footprints dictionary we use a greedy algorithm similar to matching pursuit.⁵

We compute the 4 inner products:

$$\langle Y, \hat{fp}_{k_1}^{(d)} \rangle \quad d \in \{0, 1\}$$

and:

$$\langle Y, \hat{fp}_{k_2}^{(d)} \rangle \quad d \in \{0, 1\}.$$

Then we choose the location $k \in \{k_1, k_2\}$ such that the quantity:

$$(|\langle Y, \hat{fp}_k^{(0)} \rangle|^2 + |\langle Y, \hat{fp}_k^{(1)} \rangle|^2)$$

is maximum. The two contributions $\langle Y, \hat{fp}_k^{(0)} \rangle, \langle Y, \hat{fp}_k^{(1)} \rangle$ are then subtracted from the signal

$$R^1 y = Y - \langle Y, \hat{fp}_k^{(0)} \rangle \hat{fp}_k^{(0)} - \langle Y, \hat{fp}_k^{(1)} \rangle \hat{fp}_k^{(1)}$$

and the process is iterated on the residue. Now, if we call $R^M y$ the residue computed after M iterations then it can be shown that $\|R^M y\|$ converges exponentially to 0 as $M \rightarrow \infty$.⁵ So after enough iterations, this algorithm gives an adequate representation of X . Usually a good approximation of the signal is obtained after few iterations.

Finally, notice that the results of this sub-section can be extended to the case of piecewise polynomial signals with maximum degree $D > 1$. In that case any discontinuity can be specified by a linear combination of $D + 1$ elementary footprints $\hat{f}p_k^{(d)}$ $d = 0..D$ and since footprints are shift-variant, we need a dictionary \mathcal{D} of $(D + 1) \cdot 2^J$ footprints to represent any discontinuity.

4. DENOISING OF PIECEWISE POLYNOMIAL SIGNALS BASED ON FOOTPRINTS

The term *denoising* is usually referred to the problem of removing noise from a corrupted signal without impairing the original signal too much.

The typical problem formulation is the following: one has to estimate an N -dimensional discrete signal X , whose samples $(x_i, i = 0..N - 1)$ have been corrupted by additive noise. One observes $f_i = x_i + e_i$ where e_i are independent and identically distributed (iid) zero mean Gaussian variables with variance σ^2 . The original signal is deterministic and it is independent of the noise.

Let us denote with F, X, E the vector representation of the signals under consideration. The goal of a denoising algorithm is to obtain an estimate \hat{X} of the original signal which minimizes a risk function usually the mean square error $E[\|X - \hat{X}\|^2]$.

Recently wavelet transforms have been successfully used for denoising. The most common method consists in thresholding the wavelet coefficients and it involves three main steps:

- Make a discrete orthogonal wavelet transform of the noisy signal $G = WF$.
- Set a threshold T and apply a soft or hard thresholding to the transformed coefficients.
- Transform back to the original domain the thresholded coefficients.

The *soft-threshold* function is defined as:

$$\eta_T(x) = \text{sgn}(x) \cdot \max(|x| - T, 0),$$

while the *hard-threshold* function is:

$$\phi_T(x) = x \cdot \mathbf{1}\{|x| > T\}.$$

This approach has near optimal theoretical properties. Donoho and Johnstone³ have shown that if we take a threshold $T = \sigma\sqrt{2 \ln N}$ then the function \hat{X} estimated with the proposed algorithm is the best estimator in the min-max sense as $N \rightarrow \infty$. Moreover for $N \geq 4$ the mean square error satisfies:

$$E[\|X - \hat{X}\|^2] \leq (2 \ln N + 1)(\sigma^2 + \sum_{m=0}^{N-1} \min(|y_m|, \sigma)),$$

where $y_m, m = 0..N - 1$ are the wavelet coefficients of X .

Wavelet thresholding is effective because the wavelet transform gives a sparse representation of the signal. The essential signal information is carried by few large coefficients which are usually preserved in the thresholding step, while small coefficients (carrying less information) are usually removed.

In particular if the signal is piecewise smooth (i.e. piecewise polynomial) then wavelet coefficients in the neighbourhood of sharp discontinuities tend to be large and are not affected by the thresholding step while coefficients around smooth regions are small or zero and then it is very likely that they are set to zero. For this reason wavelet thresholding is usually seen as an adaptive smoothing algorithm and it shows performance very close to a denoising algorithm that knows in advance the discontinuity positions.

For piecewise polynomial signals we know that non-zero wavelet coefficients appear only around discontinuities and that these coefficients are highly dependent. Moreover, if D is the maximum degree of any polynomial in the signal, then any discontinuity lies on a $D + 1$ dimensional sub-space.

In our denoising algorithm using the footprints dictionary \mathcal{D} , for each possible discontinuity position we project the corresponding set of noisy wavelet coefficients on the right sub-space, and if this projection is large enough we

keep it, otherwise we cancel it (thresholding step). This is equivalent to applying a vector threshold on the set of noisy coefficients rather than a scalar threshold as in the traditional approach.

So, consider a piecewise polynomial signal X of maximum degree D over the interval $[0, N - 1]$ and assume that F is the observed noisy version of X . Clearly there are $N - 1$ possible discontinuity positions ($k \in [1, N - 1]$). Let \mathcal{D} be the dictionary of $(D + 1) \cdot 2^J$ elementary footprints with $2^J < N$. Our denoising algorithm works this way:

1. Compute a J level wavelet transform of the noisy signal: $G = WF$.
2. Define a threshold $T = \sigma\sqrt{2 \ln N}$.
3. Given the dictionary of footprints \mathcal{D} , for each possible discontinuity position $k \in [1, N - 1]$ compute the $D + 1$ inner products $\langle G, \hat{f}p_k^{(d)} \rangle, \hat{f}p_k^{(d)} \in \mathcal{D}$.
4. Choose the location k_1 such that $\sum_{d=0}^D |\langle G, \hat{f}p_{k_1}^{(d)} \rangle|^2$ is maximum.
5. If

$$\exists d \in \{0..D\} : |\langle G, \hat{f}p_{k_1}^{(d)} \rangle| \geq T, \quad (3)$$

then compute the residue:

$$R_G^1 = G - \sum_{d=0}^D \langle G, \hat{f}p_{k_1}^{(d)} \rangle \cdot \mathbf{1}\{|\langle G, \hat{f}p_{k_1}^{(d)} \rangle| > T\} \hat{f}p_{k_1}^{(d)}.$$

6. Iterate step 3-4-5 on the residue until condition (3) is not verified anymore.

Finally the estimated signal \hat{X} is:

$$\hat{X} = W^{-1} \left[G_s + \sum_{m=0}^{M-1} \sum_{d=0}^D \langle R_G^m, \hat{f}p_{k_m}^{(d)} \rangle \hat{f}p_{k_m}^{(d)} \right], \quad (4)$$

where M is the total number of iterations, $R_G^0 = G$ and G_s are the scaling coefficients of G .

Notice that step 4 can be seen as a robust, multiscale singularity detection. The efficiency of this algorithm depends on the speed of convergence of the Matching Pursuit (MP). Clearly if the original signal has several discontinuities but the corresponding footprints do not overlap or overlap only with few coefficients, then MP has a fast convergence and our proposed algorithm performs well. However, performance is degraded when discontinuities are very close to each other and footprints overlap in many coefficients.

Finally, notice that the reconstruction formula (4) does not guarantee that \hat{X} is a piecewise polynomial signal. This is because the reconstructing is done in a noisy environment, so there is an inconsistency between the footprint estimation and the scaling coefficients G_s , this can lead to pseudo Gibbs phenomenon once the inverse wavelet transform is performed. To avoid this problem we reconstruct the signal directly in the time domain. Basically once the contribution $\langle R_G^m, \hat{f}p_{k_m}^{(d)} \rangle$ has been chosen, we directly consider the signal $\hat{T}_{k_m}^{(d)}$, that is the time domain version of $\hat{f}p_{k_m}^{(d)}$ and add all the contributions directly in the time domain. The reconstruction formula becomes:

$$\hat{X} = \sum_{m=0}^{M-1} \sum_{d=0}^D \langle R_G^m, \hat{f}p_{k_m}^{(d)} \rangle \hat{T}_{k_m}^{(d)} + \sum_{d=0}^D \langle F, P^{(d)} \rangle P^{(d)}.$$

5. SIMULATION RESULTS

In this section we are interested in assessing the performance of our denoising algorithms. We will only consider piecewise step or piecewise linear signals.

As a first experiment, we want to analyze the behaviour of our denoising algorithm with the size N of the signal. Recall that the signal to noise ratio $\text{SNR} = \left(\frac{\|X\|^2}{\|X - \hat{X}\|^2} \right)$ increases with the size of the signal. In case of an oracle method, where an oracle tells in advance the discontinuity locations, the SNR increases like N with the size of the signal.⁴ In,³ it was shown that for large N a hard-thresholding method behaves as $\frac{N}{\ln^2 N}$. In Figure 4 we compare our system against the theoretical behaviour of a hard-thresholding method and the ideal bound represented by the oracle method. In this first experiment we are considering piecewise step signals with no more than two discontinuities and we are making the hypothesis that footprints never overlap. It is interesting to notice that footprints based algorithm has a better behaviour than a hard-thresholding based method and it shows a behaviour quite close to the ideal one. This is a strong result, because it clarifies that the vector thresholding approach leads to a substantial improvement in performance. In Table 1, we compare the performance of our algorithm against a

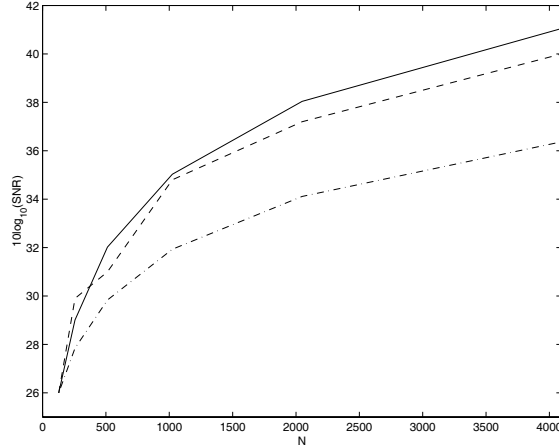


Figure 4. SNR of denoising algorithms in function of the size of the signal. Line: oracle method, dashed: footprint denoising, dashed-dotted: hard-thresholding theoretical behaviour.

real hard-thresholding system and in Figure 5 and in Figure 6 we make a visual comparison between our algorithm and hard-thresholding algorithm. The method based on footprints behaves well also in terms of visual quality, there are no artifacts around discontinuities. That is because footprints based algorithm knows in advance that the signal under consideration is piecewise polynomial and then forces the reconstructed one to belong to the same class of signals.

N	128	256	512	1024	2048	4096
Footprints	26dB	29.9dB	31dB	34.8dB	37.2dB	40dB
Hard-thresholding	25dB	27.6dB	29.5dB	32.2dB	34.4dB	36.1dB

Table 1. Footprints vs. Hard-thresholding. Piecewise step signals with no more than two discontinuities.

When footprints overlap the performance of our denoising algorithm is degraded. This is because Matching Pursuit does not represent a good way to reconstruct the signal. Assume that the signal has two overlapping footprints generated by two discontinuities at positions k_1 and k_2 . Recall that each discontinuity lies on a $D + 1$ dimensional sub-space and that these two sub-spaces are not orthogonal. Now, the discontinuity positions k_1 and k_2 can be estimated with Matching Pursuit, while the reconstruction can be performed using the right dual basis for that sub-spaces or performing a new wavelet transform with a number of decomposition level $L < J$ such that the corresponding footprints do not overlap anymore. The design of an adaptive denoising algorithm that efficiently deals with the problem of overlapping footprints is a topic under investigation.

6. CONCLUSIONS

In this work we have addressed the problem of efficiently denoising piecewise polynomial signals. We have analyzed the correlation existing between wavelet coefficients and have introduced the notion of footprints. Thanks to footprints we have designed a denoising algorithm which outperforms classical denoising algorithm based on scalar thresholding and which does not lead to the pseudo Gibbs effect.

We believe that the concept of footprint can be also extended to piecewise smooth signals and can represent an effective tool for denoising of *real-world* signals.

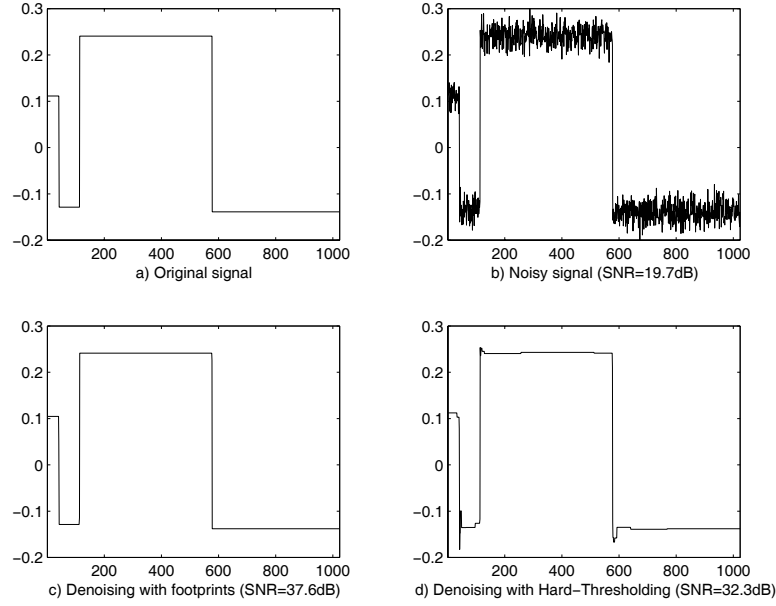


Figure 5. a) Original Signal, b) Noisy Signal (SNR=19.7), c) Denoised Signal using footprints (SNR=37.6dB), d) Denoised Signal with Hard-Thresholding (SNR=32.3dB)

REFERENCES

1. R.R.Coifmann and D.L. Donoho "Translation Invariant de-noising", Technical Report 475, Dept of statistics, Stanford University. May 1995.
2. I. Daubechies *Ten Lecture on Wavelets* SIAM, Philadelphia,PA,1992
3. D.L.Donoho and I.M. Johnstone "Ideal spatial adaptation via wavelet shrinkage," Biometrika, 81:425-455, 1994.
4. S. Mallat *A wavelet tour of signal processing* Academic Press, 1998.
5. S. Mallat and Z. Zhang "Matching pursuits with time-frequency dictionaries," IEEE Transactions on Signal Processing, 41 (12):3397-3415, December 1993.
6. P. Prandoni and M. Vetterli "Approximation and Compression of Piecewise Smooth Functions," Phil. Trans. R.Soc.Lond. 1999.
7. P.Prandoni "Optimal segmentation techniques for piecewise stationary signals" PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland. 1999.
8. G. Strang and T. Nguyen *Wavelets and Filter Banks* Wellesley-Cambridge Press.
9. M.Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, 1995.

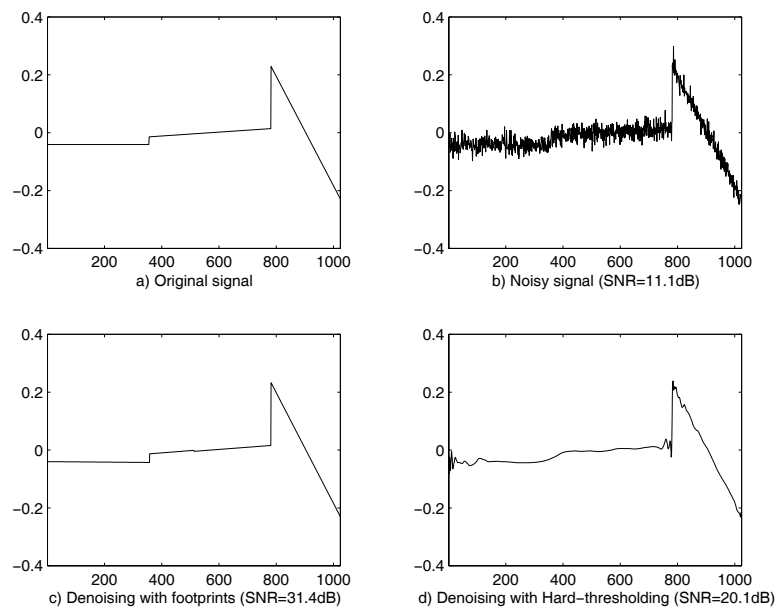


Figure 6. a) Original Signal, b) Noisy Signal (SNR=11.1), c) Denoised Signal using footprints (SNR=31.4dB), d) Denoised Signal with Hard-Thresholding (SNR=20.1dB)