

Adaptive splitting for stabilizing 1-D wavelet decompositions on irregular grids

Ward Van Aerschot

Maarten Jansen

Adhemar Bultheel

Report TW 413, Dec 2004



Katholieke Universiteit Leuven

Department of Computer Science

Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Adaptive splitting for stabilizing 1-D wavelet decompositions on irregular grids

Ward Van Aerschot

Maarten Jansen

Adhemar Bultheel

Report TW 413, Dec 2004

Department of Computer Science, K.U.Leuven

Abstract

This paper proposes a scheme to improve the stability of wavelet decompositions on 1-D irregular grids. Wavelet transforms on irregular grids are constructed using the lifting scheme. The filters in this scheme take the structure of the grid into account. Nice as it is, however, we undoubtedly bump into numerical stability issues directly related to the irregularity of the grid. Existing stabilizing methods concentrate on the filters used in the lifting scheme itself. While this may be effective in reducing the instability, they are inadequate when a highly irregular grid is involved. The approach presented in this paper is different, as it concentrates on the subsampling or subdivision. Grid locations in the multiscale transform are inserted in such an order that the irregularity of the grid at coarse scales is kept under control. This way, the proposed algorithm prevents instability at coarse scales, rather than healing it. Simulations illustrate that the proposed multiscale decomposition scheme is much more stable than the currently available transforms, especially at coarse scales, where effects of instability have a wide range.

Keywords : wavelet, stable, irregular.

AMS(MOS) Classification : Primary : 42C40, Secondary : 65T99.

Adaptive splitting for stabilizing 1-D wavelet decompositions on irregular grids

Ward Van Aerschot, Maarten Jansen, Adhemar Bultheel

December 9, 2004

Abstract

This paper proposes a scheme to improve the stability of wavelet decompositions on 1-D irregular grids. Wavelet transforms on irregular grids are constructed using the lifting scheme. The filters in this scheme take the structure of the grid into account. Nice as it is, however, we undoubtedly bump into numerical stability issues directly related to the irregularity of the grid. Existing stabilizing methods concentrate on the filters used in the lifting scheme itself. While this may be effective in reducing the instability, they are inadequate when a highly irregular grid is involved. The approach presented in this paper is different, as it concentrates on the subsampling or subdivision. Grid locations in the multiscale transform are inserted in such an order that the irregularity of the grid at coarse scales is kept under control. This way, the proposed algorithm prevents instability at coarse scales, rather than healing it. Simulations illustrate that the proposed multiscale decomposition scheme is much more stable than the currently available transforms, especially at coarse scales, where effects of instability have a wide range.

Introduction

The lifting scheme [14, 13] has proven to be an excellent tool for extending the scope of applications of wavelet or general multiscale decompositions. The filters used in this scheme can be made grid-adaptive (i.e., non-stationary), nonlinear [1], or even data-adaptive [2, 8, 9]. Grid-adaptive, non-stationary filters [4, 10, 16] are interesting in the extension of wavelet transforms for data on non-equidistant grids.

The lifting filters can be designed according to several criteria. These criteria include existence and smoothness of the associated wavelet basis functions. It is also important that those basis functions can be found numerically from a convergent multiscale subdivision scheme. Another criterion is sparsity. A wavelet representation is sparse if a large proportion of its coefficients are not significant. Coefficients that are not significant, can be omitted in a compression scheme. In classical wavelet decompositions, compression is achieved by thresholding the smallest coefficients. Compression rates are determined by both the proportion of small coefficients and by closeness to zero of the individual coefficients. An important parameter in this context is the (dual) number of vanishing moments, defined in Section 1. A fourth criterion is numerical stability. As explained in Section 2, instability means that a transform is ‘far from orthogonal’. As a consequence, it is difficult to control the effect of operations on wavelet coefficients — such as thresholding — after taking the inverse wavelet transform.

Although stabilizing lifting filters have been designed in previous contributions [10, 16, 17], the stability of the resulting transform still depends on the irregularity of the underlying grid. This is formalized in a homogeneity constant, defined in Section 2.1. The goal of this paper is to develop a grid-adaptive subsampling scheme such that subgrids on intermediate levels are as regular as possible. If the grid of the observations is highly irregular, this irregularity inevitably shows up at the finest resolution level. At the finest level

local instabilities, however, have a limited range. In the classical subsampling scheme not only some of the irregularity persists at coarser scales, but also new irregular structures may show up. Although a locally adaptive subsampling procedure [16] may help to reduce the effect of instabilities, this scheme still prefers the classical, regular subsampling idea, with local exception handling. The scheme presented in this paper is globally grid-adaptive in the sense that it does not a priori prefer the classical regular subsampling. Thereby, it can easily be combined with any sort of lifting filters.

This paper is organized as follows: in Section 1 the key concepts about lifting are summarized. Section 2 introduces the idea of grid-adaptive subsampling, called splitting in a lifting context. Section 3 details the proposed method. The results are given in Section 4.

1 Second generation wavelets

The lifting scheme as proposed in [12] makes it possible to carry out discrete wavelet transforms without leaving the spatial domain. This scheme can easily be extended for wavelet transforms on irregular grids leading to what are called *second generation wavelets* [13]. The lifting scheme consists of three stages that we briefly discuss.

In the first stage the input data s_j are subdivided into two disjunct sets: the *even* and the *odd* set. In the remainder of this paper we keep referring to the two sets in terms of *even* and *odds* – as originated from the classical setting – although later on the words *even* and *odd* will lose their meaning¹.

In the second stage the value of each odd gridpoint is predicted in terms of values of even gridpoints. This predicted value is subtracted from the real value and represents the detail or wavelet coefficient $d_{j-1,k}$:

$$d_{j,k} = s_{j+1,2k+1} - \nu_{j+1,2k+1} * s_{j+1}, \quad (1)$$

with $\nu_{j+1,2k+1}$ the impulse response of the prediction filter, depending on the prediction location $x_{j+1,2k+1}$. Note that the filter coefficients are non-stationary unlike in the regular setting. At this step the *coarser* scaling functions connected to the even grid locations are built out of the *finer* scaling functions involved:

$$\varphi_{j,k}(x) = \varphi_{j+1,2k}(x) + \sum_m \nu_{j+1,2(k-m)+1}^{[m]} \varphi_{j+1,2(k-m)+1}(x). \quad (2)$$

With $\nu_{j+1,2(k-m)+1}^{[m]}$ the m th coefficient of $\nu_{j+1,2(k-m)+1}$. This equation is better known as the *dilatation equation*. Properties as compression characteristics of a wavelet analysis depend on the order of the predictor.

Definition 1 *The dual order \tilde{N} of a multi-resolution analysis is given by:*

$$\tilde{N} = \max \{N | \exists j_0 \text{ so that } \forall j \geq j_0 : \Pi_N \subset V_j\},$$

with Π_N the space spanned by polynomials of degree less or equal to $N - 1$ and V_j the space spanned by the scaling functions φ_j .

This means that all polynomials up to degree $\tilde{N} - 1$ lie in the spaces V_j ($j \geq j_0$), and can be expressed in terms of scaling functions φ_j .

The third stage is an update step. This step assures that several properties such as the mean value and higher order moments of the incoming high resolution $f_j = P_j f$ representation are being preserved in the output $f_{j-1} = P_{j-1} f$, with P_j the projection operator on space V_j .

¹The odd gridpoints will be assigned a detail coefficient while the evens pass the lifting scheme carrying a scale coefficient.

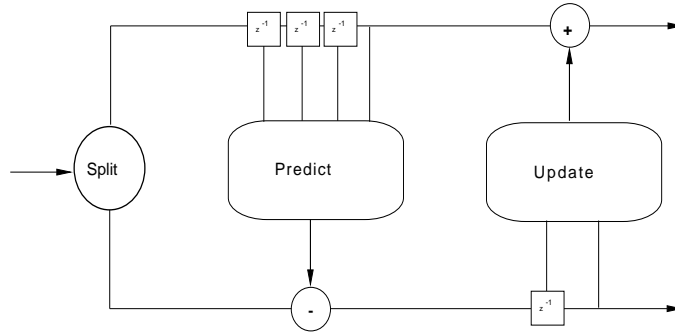


Figure 1: The lifting scheme. Here a 4-taps prediction filter and 2-taps update filter is presented. The prediction can be seen as an high-pass filter when it's output has been subtracted from the odd samples. The update filter can be understood as an anti-aliasing filter by adding its output to the decimated signal. This shows why the lifting scheme implements a filter bank in a very efficient manner.

Definition 2 We define a p -th order vanishing moment of $\psi_{j,k}$ as:

$$\int_{-\infty}^{\infty} x^p \psi_{j,k}(x) dx = 0.$$

Definition 3 The primal order N of a multi-resolution analyses is defined as:

$$N = \max \{n | \exists j_0 \text{ so that } \forall j \geq j_0 : \Pi_N \perp W_j\}.$$

Considering definition 2, we learn that wavelet functions must have vanishing moments of order $p = 0, \dots, N - 1$ to have a multi-resolution analyses of order N . In order to have a numerical stable transform the primal order must be at least one [6]. The number of vanishing moments given to the wavelet functions determine the approximation speed of the inverse wavelet transformation to the original signal. Analogous to the prediction phase we have:

$$\psi_{j,k} = \varphi_{j+1,2k+1} - \sum_m \mu_{j+1,k-m}^{[m]} \varphi_{j,k-m,m} \quad (3)$$

The scaling coefficients are updated accordingly:

$$s_{j,k} = s_{j+1,2k} + \mu_{j+1,k} * d_j \quad (4)$$

By cascading different lifting schemes, letting the scaling coefficients of the previous scheme be the input of the next one, we can perform a full wavelet decomposition over multiple scales. Note that in the irregular setting the filter coefficients are adapted to the underlying grid locations while the stencil moves along. Because of this non-stationary behaviour, base functions are no longer translates of each other. Dilation, however, still applies locally (see eq. (2)).

2 Grid adaptive splitting

When dealing with irregular grids the classical even/odd splitting scheme can not prevent scale mixing. Mixing of scales undoes the benefits and even the concept of a multi scale analysis. Multi-resolution analysis wants to represent a function f by $f_j = P_j f$ with the same degree of resolution along the interval. In the passed, there have been proposed some ad hoc methods [15] that worked under specific settings, but failed in general.

For applications as noise reduction where some wavelet coefficients are being subjected to thresholding, keeping the transformation stable is of utmost importance. Otherwise a small perturbation on a coefficient could have a huge impact on the rectified signal. The cause of this behavior lies in the fact the base functions of the different wavelet spaces are too strangled up with each other, such that slightly perturbing a coefficient on a high resolution level could have its effect on coarser levels. Avoiding scale mixing is a necessary step to enforce stability. The instability caused by scale mixing can not be undone by stabilizing methods trying to orthogonalise the subsequent wavelet spaces afterwards (see [10]). Therefore we propose a splitting scheme that tries to handle both scale-mixing and instability problems. In this manner we become a fully adaptive second generation wavelet decomposition tool. Not only the prediction and the update, but also the splitting is made adaptive with respect to the underlying grid. Furthermore, we want to nicely fit the algorithm into the lifting scheme by keeping it local. A consequence is that we can only approximate the most stable splitted grid. On the other hand the complexity of the algorithm does not destroy the fast time complexity property of the lifting scheme.

2.1 Irregularity and Stability

There exists a strong connection between the irregularity of the grid and the stability of the wavelet transform. Some notions of stability are briefly introduced. We first define uniformly stable bases:

Definition 4 *The sequence of bases Φ_j is called uniformly stable if there exists constants c and C , not depending on j , such that*

$$c \sum_{k=1}^{2^j} |s_{j,k}|^2 \leq \|f\|^2 \leq C \sum_{k=1}^{2^j} |s_{j,k}|^2,$$

for any function $f_j = \sum_{k=1}^{2^j} s_{j,k} \varphi_{j,k}$.

Let T_j be the multi-scale transform defined as a chain of inverse two-scale transforms :

$$T_j := A_{j-1} \begin{bmatrix} T_{j-1} & 0 \\ 0 & I \end{bmatrix}, \quad T_1 := A_0.$$

If the scale bases $\Phi_j = \{\varphi_j\}$ are uniformly stable and the multiresolution analysis is dense in L_2 , it can be proven that [3] on the scale bases the corresponding wavelet bases are stable bases (Riesz bases) if and only if there exists an upper bound for the sequence of condition numbers $\kappa(T_j)$.

The irregularity of a grid can be expressed in what we call a homogeneity constant. Let the gridpoints form a strictly increasing sequence: $x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N$.

Definition 5 *The homogeneity constant is defined as*

$$\gamma_j^* = \sup_{j,k} \frac{\max(x_{j,k+2} - x_{j,k+1}, x_{j,k} - x_{j,k-1})}{x_{j,k+1} - x_{j,k}}. \quad (5)$$

As long as $\gamma < \infty$ the multilevel grid is called homogeneous. Stability of second generation wavelets can be expressed in terms of this homogeneity constant γ . Earlier work provides an upperbound of the condition number in terms of the homogeneity constant. Figure 2 illustrates what happens if we still apply the classical splitting scheme. It shows the reconstruction of a denoised signal after a simple treshold procedure. The result is far

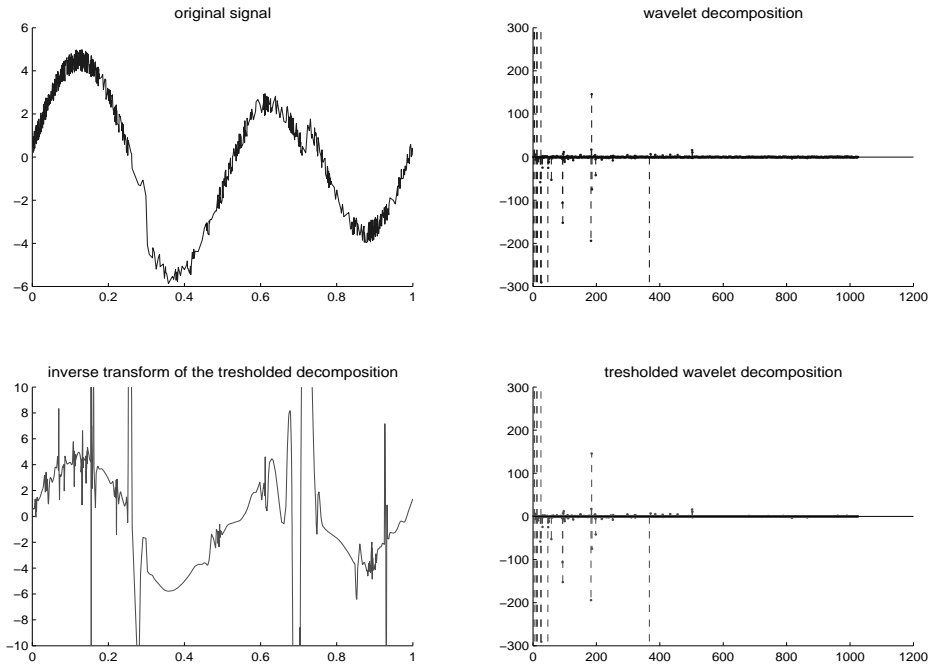


Figure 2: (Top left) The original noisy signal. (Top right) The wavelet decomposition exists for a certain amount of tiny wavelet coefficients. Thresholding those coefficients one can assume to reduce noise without causing a huge bias in the reconstructed signal. (Bottom right) Those wavelet coefficients which have a value less than some threshold and lying on the three highest resolution levels are being thresholded. The rest of the wavelet and scaling coefficients are left untouched. (Bottom left) The figure shows the result of the inverse transformation of the thresholded decomposition. Small changes in the decomposition have been blown up after inverse transformation and gave rise to enormous peaks in the reconstructed signal.

worse than the noisy signal from which we started. Let δc be the difference between the exact decomposition and the thresholded one, perturbation theory² states that

$$\delta f \approx \kappa(T_j) \delta c.$$

The large condition number of T_j explains why a small threshold can cause such a bias. Constructing more stable bases leads to a better conditioning of the multi-scale transform T_j , allowing thresholding to do its job right.

3 The proposed method

Recall from equation (2) that scale functions φ_j spanning V_j are formed out of scale functions φ_{j+1} spanning V_{j+1} . But also V_j will be split up resulting into subspaces W_{j-1} and V_{j-1} . If the scale functions spanning V_{j-1} already form an unstable base, — the scale functions are far from orthogonal — this instability will only further accumulate during the following decompositions and also affect the stability of the wavelet bases. By carefully adapting the construction of the scaling functions one can improve the conditioning of the corresponding MRA. As a consequence the proposed algorithm cooperates with the predictor.

²for further reading concerning condition numbers and perturbation theory we refer to literature like [7], [5]

Suppose the predictor has at least one vanishing moment, then we know that the prediction coefficients sum up to one. If, in addition, all coefficients are positive, we have:

$$\| \nu_{j,k} \|_1 = \sum_m |\nu_{j,k,m}| = 1. \quad (6)$$

In this case, the predicted value is a linear convex combination of scale coefficients $s_{j,k}$, which can be proven to be stable. Incoming errors ϵ on the values $s_{j,k}$ causes a prediction error ν as:

$$|\nu_k| \leq \sum_m |\nu_{j,k,m}| |\epsilon_m| \leq \max_m |\epsilon_m| \sum_m |\nu_{j,k,m}| = \| \epsilon \|_\infty.$$

However, in case of an interpolating prediction scheme only a linear predictor has positive coefficients. Higher order polynomials give rise to prediction coefficients lying outside the interval $[0, 1]$. The coefficients $\nu_{j,k,m}$ of an interpolating filter are simply the corresponding Lagrange weights.

Each time a scaling coefficient $s_{j+1,2k}$ has contributed to the prediction of an odd coefficient $s_{j+1,2k+1}$, its corresponding intermediate scale function $\hat{\varphi}_{j+1,2k}$ is added to the scaling function $\varphi_{j+1,2k}$ connected to the odd coefficient. After its participation in all the predictions involved its scale function is that as stated in eq. (2). These contributions diminish the angle³ between subspaces V_j and W_j . Large prediction coefficients causes heavy side lobes into the scaling functions, causing a huge overlap between the scaling functions and the later formed wavelet functions. That is what has to be prevented from happening. Carefully constructing subgrids selecting a gridpoint $x_{j+1,s}$ between each two coarser gridpoints $x_{j,k}$ and $x_{j,k+1}$ such that:

$$x_{j+1,s} = \inf_i \left\{ \max_m |\nu_{j,k,m}(x_{j+1,i})| : x_{j+1,i} \in [x_{j,k}, x_{j,k+1}] \right\} \quad (7)$$

These selected gridpoints form the set that we would call the odds, O_j , that are predicted from the evens, E_j . Merging those sets gives us the even set at a higher resolution level, E_{j+1} . We keep performing this algorithm until we eventually have our original set of grid points at the highest resolution level. Note how the upper bound of the prediction error is being controlled:

$$|\nu_k| \leq \sum_m |\nu_{j,k,m}| |\epsilon_m| \leq 4 |\nu_{j,k,m}(x_{j+1,s})| \| \epsilon \|_\infty.$$

In contrast to what happens at the classical splitting, we work our way bottom up starting from the lowest resolution grid gradually construction finer and finer grids ending up with the original grid. So, the initial grid consists of both end points of the original irregular grid. The bottom-up creation of subgrid has the advantage of keeping the transform as stable as possible at low resolution levels, so the instabilities do not even occur when a high resolution reconstruction is not required.

Note the similarity of the dependency of stability in terms of the homogeneity constant γ of our splitting scheme where expression (7) can be seen as an extension of the homogeneity constant which we try to minimize.

³for readers not familiar with the notion angle between spaces please consult [11]

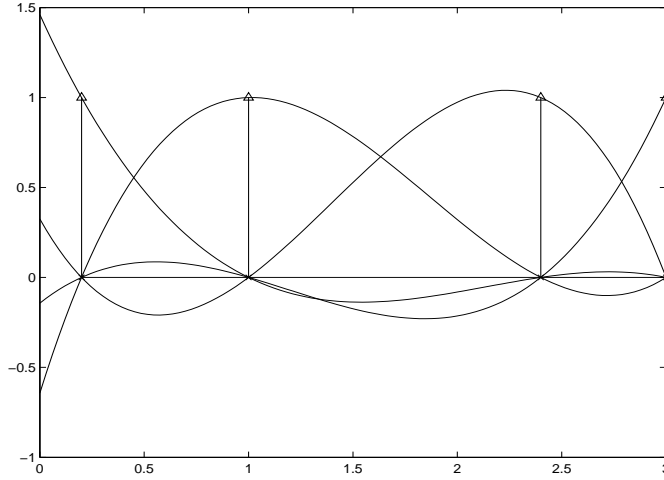


Figure 3: The figure presents the contribution of each even grid point to the prediction of an in-between grid point when a cubic polynomial prediction is used. The new grid location will be chosen between the even grid points where $\max_m |\nu_{j,k,m}(x_{j+1,i})|$ is minimal.

4 Results

On figure 4 the resulting subgrids are plotted for both the even/odd splitting and those of the adaptive splitting. The figure shows that more regular subgrids are *created* with the adaptive method compared to the subgrids *resulting from* even/odd splitting. Eventually both end up with the highest level grid. As already been stated, this bottom-up creation of subgrid keeps the transform as stable as possible at low resolution levels, so the instabilities do not even occur when a high resolution reconstruction is not an issue. This can be seen from the picture in figure 6. We have tested our method on numerous irregular grids of which the homogeneity constants can be read from figure 6 (end points), and have a range between 2^4 and 2^{10} . As we can see in the same figure the adaptive splitting method outperforms the classical even-odd splitting. Although not proven, when using grid adaptive splitting the homogeneity constants versus resolution level seem non-decreasing. The classical method even/odd splitting, however, could create more irregular subgrids than those lying on a higher (or even highest) resolution level. On the highest grid, of course, the γ_J will be the same for all methods.

Figure 7 presents the results on the scaling functions and wavelet functions resulting from the proposed splitting scheme: the scale functions behave more like in the regular setting and also scale mixing is greatly diminished. The presented result lets us already sense the better stability conditions of the constructed bases with the proposed grid adaptive splitting method. Figure 8 shows a boxplot that gives information about the condition number of T_j produced by different lifting schemes. We see that our proposed algorithm performs much better than an unstabilized and a stabilized method without adaptive splitting. Finally Figure 9 shows the output of our algorithm applied to the same data as used in Figure (2) subjected to a simple threshold. Mark how the decomposition behaves more well-behaved than that of figure 2. The range of the wavelet coefficients keeps the same order of the range as that of the input signal. A simple hard thresholding of a lot of wavelet coefficients does not result in enormous peaks in the rectified signal, while most of the noise is removed.

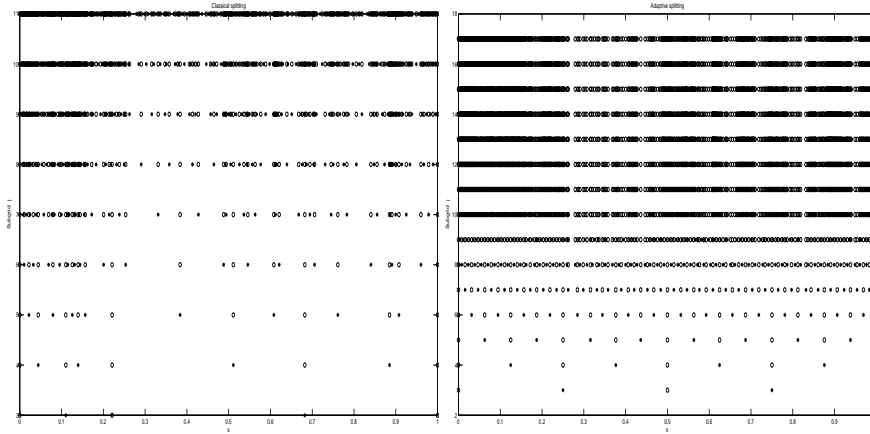


Figure 4: (Left) Top-down creation of subgrids by the classical even-odd splitting. (Right) Bottom-up creation of subgrid by our grid adaptive splitting method.

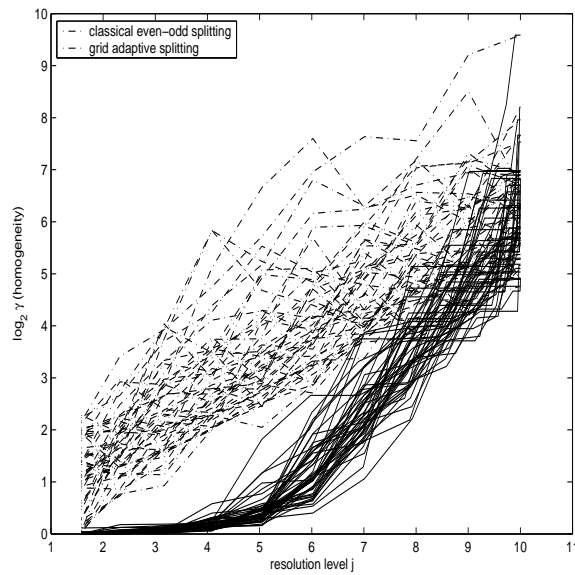


Figure 5: The homogeneity constants of fifty irregular grids are plotted against their corresponding resolution level. Notice that those plots are non-decreasing by application of the adapted method. By the classical method, however, more irregular subgrids are created than those lying on a higher resolution level, so better splitting would be recommended.

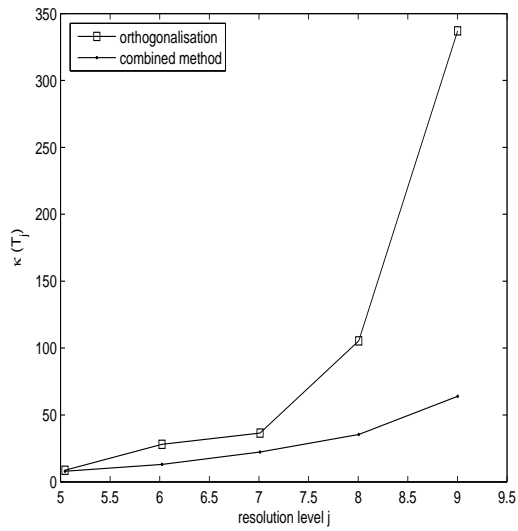


Figure 6: The condition numbers of the inverse transformation matrix T_j of an orthogonalisation method and a combined method including adaptive splitting.

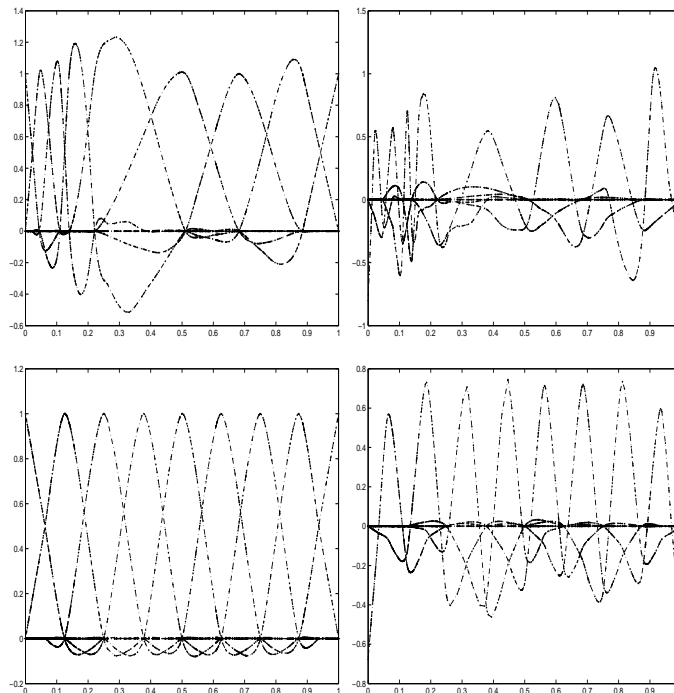


Figure 7: (Top) Scaling and wavelet functions generated by a semi-orthogonalisation method with classical splitting scheme. (Bottom) Scaling and wavelet functions generated by the same stabilisation method, this time with adaptive splitting.

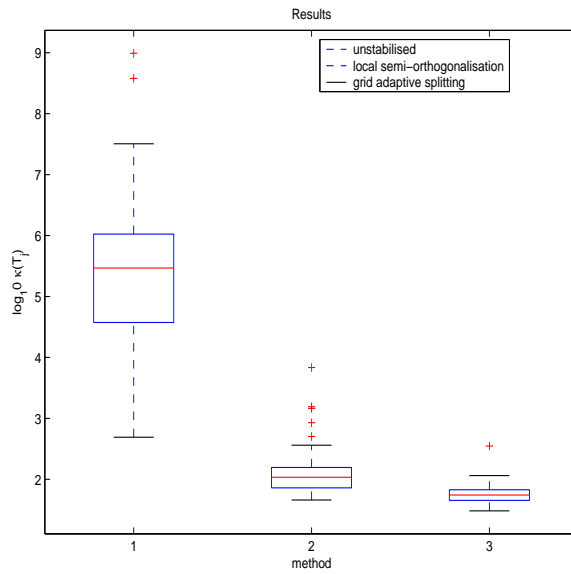


Figure 8: A comparison between the unstabilised transform, a local semi-orthogonalisation method and a combined semi-orthogonalisation method with the grid adaptive splitting phase.

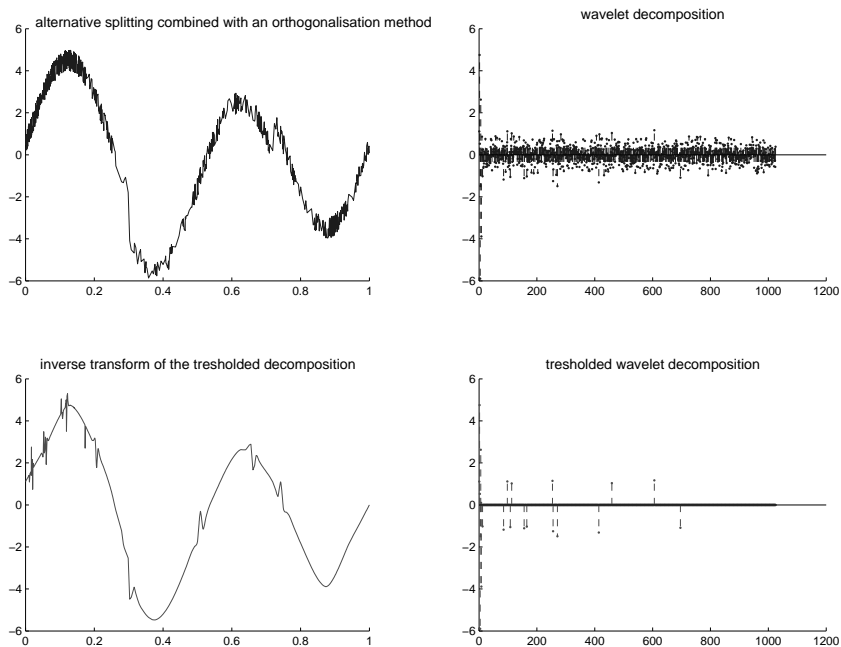


Figure 9: Note how the decomposition behaves less versatile than that of figure 2. The range of the wavelet coefficients is of the order of the range of the input signal range. A simple hard thresholding of a lot of wavelet coefficients does not result in enormous peaks in the rectified signal, while most of the noise is removed.

References

- [1] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo. Wavelet transforms that map integers to integers. *Appl. Comput. Harmon. Anal.*, 5(3):332–369, 1998.
- [2] R. Claypoole, G. M. Davis, W. Sweldens, and R. Baraniuk. Nonlinear wavelet transforms for image coding via lifting. *IEEE Transactions on Image Processing*, 12(12):1449–1459, 2003.
- [3] W. Dahmen. Some remarks on multiscale transformations, stability and biorthogonality. *P.J. Laurent, A. Le Mehaut, and L.L. Schumaker*, 1994.
- [4] V. Delouille, J. Simoens, and R. von Sachs. Smooth design-adapted wavelets for nonparametric stochastic regression. *J. Amer. Statist. Assoc.*, 2004.
- [5] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, PA, 1996.
- [6] M. Jansen and P. Oonincx. *Second generation wavelets and applications*. Springer, 2005.
- [7] Tosio Kato. *Perturbation theory for linear operators*. Die Grundlehren der mathematischen Wissenschaften. Springer Berlin, Berlin, 1984.
- [8] P. J. Oonincx and P. M. de Zeeuw. Adaptive lifting for shape-based image retrieval. *Pattern Recognition*, 36(11):2663–2672, November 2003.
- [9] G. Piella and H. J. A. M. Heijmans. Adaptive lifting schemes with perfect reconstruction. *IEEE Transactions on Signal Processing*, 50(7):1620–1630, July 2002.
- [10] J. Simoens and S. Vandewalle. A stabilized lifting construction of wavelets on irregular meshes on the interval. *SIAM J. Sci. Comput.*, 24(4):1356–1378, 2003.
- [11] Carbone A. Watson B. Singh, S.P. *Approximation theory, wavelets and applications*. NATO ASI series, Series C: Mathematical and physical sciences. Kluwer Dordrecht, 1995.
- [12] W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.*, 3(2):186–200, 1996.
- [13] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.
- [14] W. Sweldens and P. Schröder. Building your own wavelets at home. In *Wavelets in Computer Graphics*, ACM SIGGRAPH Course Notes, pages 15–87. ACM, 1996.
- [15] E. Vanraes, M. Jansen, and A. Bultheel. Stabilizing wavelet transforms for non-equispaced data smoothing. TW Report 319, Department of Computer Science, Katholieke Universiteit Leuven, Belgium, Feb 2001.
- [16] E. Vanraes, M. Jansen, and A. Bultheel. Stabilizing wavelet transforms for non-equispaced data smoothing. *Signal Processing*, 82(12):1979–1990, December 2002.
- [17] E. Vanraes, J. Maes, and A. Bultheel. Powell-sabin spline wavelets. *International Journal of Wavelets, Multiresolution and Information Processing.*, To Appear, 2004.