Adaptive Wavelet Transforms with Application in Signal Denoising

M. Tomic

***Abstract*—This papers gives an overview of wavelet transforms application to signal denoising problem. Wavelets perform well for this application and have been used before with great success. To further improve the transform efficiency adaptivity is introduced to traditional wavelet transforms. The lifting framework as a method to construct wavelets facilitates the effort to develop efficient adaptation algorithms. The most important papers that cover the topic of adaptive wavelet transforms in signal denoising are presented in brief and also the successful non-wavelet domain method for signal denoising using the intersection of confidence intervals method is described.**

***Index Terms*—wavelets, adaptive lifting scheme, denoising, ICI**

I. Introduction

Wavelet transforms have gained wide acceptance as a valuable tool for common signal processing tasks. The most important difference between the wavelet transforms and transforms such as a Fourier transform is that the wavelets are localised in both the frequency and the time. This makes it possible to better localise properties of the analysed signal. The result is a well known ability of the wavelet transforms to pack the main signal information into a very small number of large wavelet coefficients. Another advantage of wavelet transforms is that there is an indefinite number of basis functions. An appropriate wavelet can then be chosen for a specific signal which makes the transform adjustable and adaptable. Because of these excellent properties, wavelet transforms have been used with great success in many different applications, such as a signal denoising and compression or a feature detection.

We will focus on the application in signal denoising, where wavelets are used extensively. Since there is an indefinite number of wavelet transform basis functions (wavelets) possible the efficiency of noise removal is greatly influenced by the choice of the wavelet. Shorter wavelets perform better for signals with many discontinuities or lots of high frequencies. On the other hand, longer and smoother wavelets perform better for smoother signals consisting of primarily lower frequency components. Figure 1(a) shows an example of a short wavelet, while in Fig. 1(b) an example of a longer and smoother wavelet can be seen. Ideally, shorter wavelets should be used for higher frequency parts of the signal, while longer wavelets should be used for lower frequency parts of the signal. To achieve this, different adaptation techniques are used widely with varying success.

This paper will provide a brief overview of current research achievements in signal denoising using adaptive wavelet transforms.

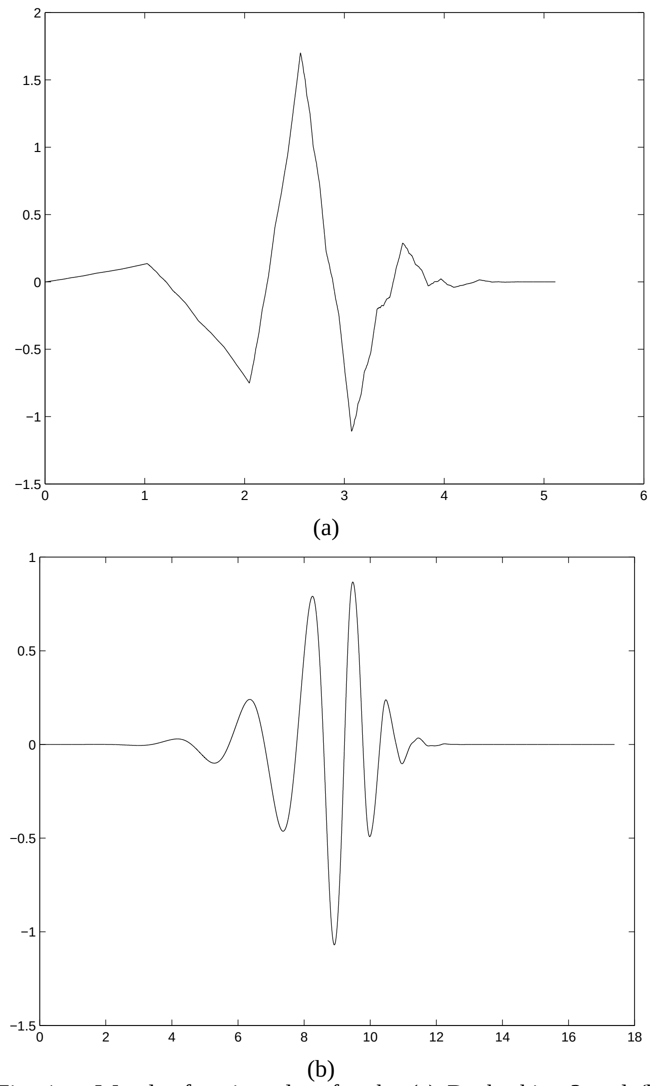

Fig. 1. Wavelet function plots for the (a) Daubechies 3 and (b) Daubechies 9

II. Wavelets and Wavelet Shrinkage

For efficient application of the wavelet theory in signal processing it was necessary to simplify the calculations and allow the adaptivity to be introduced into the transform. In [1]–[4] Mallat successfully connected the theory with the concept of the multiresolution analysis. He also showed that the wavelet analysis can be performed using iterated filter banks of certain properties. Calculating the discrete wavelet transform (DWT) became as easy as performing a simple signal filtering. In a typical scenario, the analysis filter bank consists of a low-pass and a high-pass channel. Output from the low-pass channel represent the coarse approximation of the signal, while the output from the high-pass channel contain the fine signal details. Both, the approximation and the details, are decimated by a factor of 2 and low-pass channel output is iteratively filtered with the same filter bank until the desired level of decomposition is reached. This is illustrated in Fig. 2. Details coefficients at each decomposition level are also called the wavelet coefficients. To reconstruct the original signal, synthesis filter bank is used, with its inputs upsampled by a factor of 2, as shown in Fig. 3.

The above procedure covers the basics of the wavelet analysis of the given signal but it does not provide a tool to remove a noise from the noisy signal. For denoising purposes, the wavelet shrinkage technique is used. It is based on the fact that the magnitude of the wavelet coefficients is important. Coefficients of higher magnitude are more likely to represent the main signal properties, while the coefficients of smaller magnitudes are more likely to be caused by the noise. To remove the noise, a thresholding is applied to wavelet coefficients before signal reconstruction. A work was done in this area by Donoho and Johnstone [5]–[8]. They successfully used the wavelet shrinkage technique for signal denoising and done significant research of the topic.

Although it has been showed that the wavelet transforms perform well in signal denoising they still exhibit unwanted visual artifacts, such as a Gibbs phenomena around signal discontinuities. Since the traditional wavelet transform is not translation invariant the size and the impact of the artifacts might heavily depend on the position of the discontinuity within the signal. To reduce the artifacts regardless of the discontinuity position, Coifman and Donoho in [9] proposed the translation-invariant denoising, by means of the undecimated wavelet transform. In this transform the decimation step is not used, which results in an overcomplete representation of the signal. This was proved to both, significantly lower the overall root mean square error (RMSE) of the denoised signal, and also improve it's visual appearance - subjective quality. The undecimated wavelet transform is now a common approach to a signal denoising problem.

An example of denoising using undecimated wavelet transform is shown in the Fig. 4. For denoising the noisy "Blocks" signal a Db3 wavelet is used and a 4-level decomposition tree constructed. Denoised signal is shown at the bottom plot and it can be seen that the Gibbs phenomena is still clearly visible and greatly influence the signal quality. To further reduce these artifacts, many adaptive techniques were developed.

III. Lifting Scheme

Traditional wavelet transforms implemented with filter banks were widely accepted and used in many applications, however, introducing adaptivity to such transforms or dealing with irregularly spaced data is not always an easy task. In [10], [11], Sweldens proposed a new framework for the wavelets construction - the lifting scheme. The lifting scheme is a computationally efficient way of implementing the wavelet transforms and overcomes the usual filter bank approach shortcomings. It makes it easy to introduce adaptivity to wavelet transforms or process irregularly spaced data. Several detailed case study examples were given in [12] by Sweldens and Shcroder. A simple two-band decimated filter bank based on the lifting scheme is shown in Fig. 5. Before filtering, the signal $X(z)$ is split into even-indexed and odd-indexed components. There are two different types of lifting steps – the dual lifting step and the primal lifting step, which are commonly referred to as the predict and the update step, respectively. The filter $P(z)$ is used in the predict step and it predicts the value of one of the odd-indexed components, based on the values of the even-indexed components. The prediction error of the prediction will represent the values in the high-pass channel of the filter bank, i.e. the wavelet coefficients. The filter $U(z)$ is used in the update step. It's purpose is to update the even-indexed components, based on the wavelet coefficients. The net effect will actually be the low-pass filtering of the even-indexed components. Output of the filter bank are coarse approximation, or average $A(z)$ and signal details $D(z)$. Daubechies and Sweldens in [13] showed that any discrete wavelet transform can be decomposed into a finite sequence of lifting steps. Also, in [14], Kovacevic and Sweldens presented generalisation of the lifting scheme approach to wavelet transforms to arbitrary dimensions.

If signal denoising is a target application of the filter bank then the decimators can be removed and filters

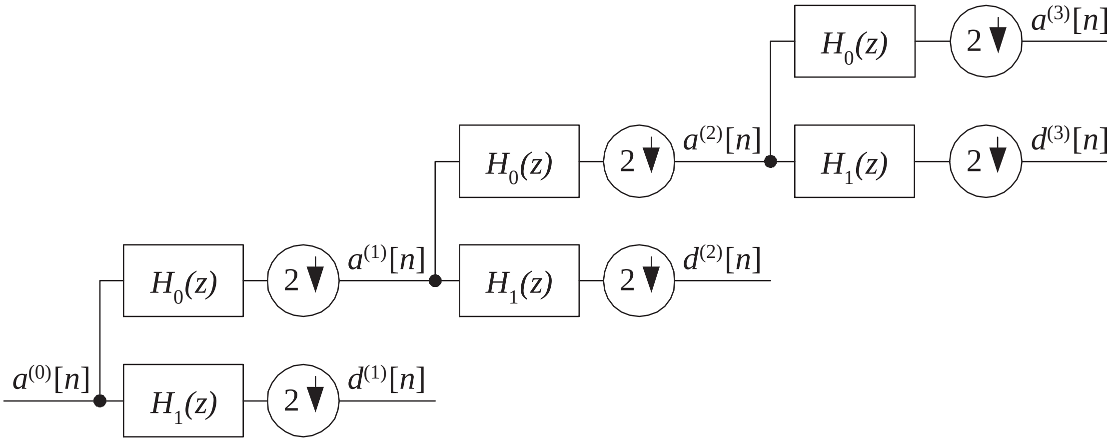

Fig. 2. Wavelet filter bank decomposition

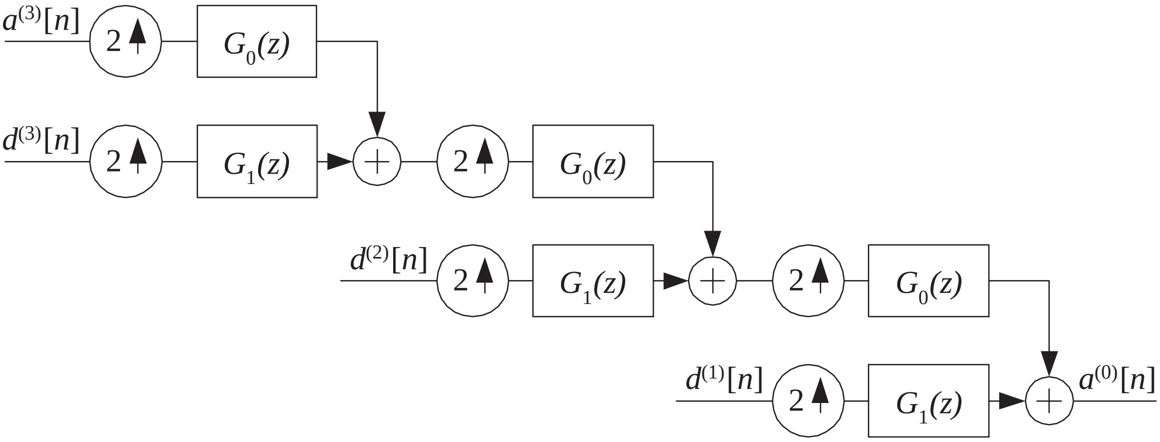

Fig. 3. Wavelet filter bank reconstruction

P(z) and *U(z)* upsampled, resulting in an undecimated transform.

IV. Adaptive Wavelet Transforms Using the Lifting Scheme

Since all calculations in the lifting framework are done in the spatial domain, it is easy to introduce adaptivity into transform. The idea of adaptation is commonly focused on modifying the transform properties based on a certain signal features or local properties. Typically, shorter wavelets should be used in the neighbourhood of discontinuities or generally higher frequencies in a signal, while longer wavelets should be used for the signal regions where lower frequencies prevail.

Claypoole et al. [15]–[17] proposed two adaptive transforms - the scale-adaptive transform (ScAT) and the space-adaptive transform (SpAT). In the ScAT, an adaptation is performed on a scale-by-scale basis. From a set of predictors they choose the one which better fits the prevailing signal properties. The objective measure of the quality of fit is a simple sum of squared prediction errors. The predictor which yields the smallest sum of squared errors is chosen as the predictor for the entire signal at a given scale. The approach in the SpAT transform is to perform adaptation on a point-by-point basis. The predictor is chosen for each data point as the one which gives the smallest prediction error. It is to be expected that such predictor will often fit well to the local signal properties and make it possible for the wavelets to contract in the neighbourhood of signal discontinuities. The shortcoming of SpAT is that it makes the predictor space-varying. Since the update step depends on the predictor it may complicate its design or introduce the necessity of bookkeeping of the predictor selection. To resolve this, Claypoole proposed the update first framework where the update step is performed before the predict step.

Unlike Claypoole et al. who adapt the predictor, Piella et al. in [18] proposed an adaptive update transform where only the update filter is adapted, while the predictor is fixed. In the general form presented, choice of the update filter depends on the decision maps which can be based on any meaningful criteria. In the original

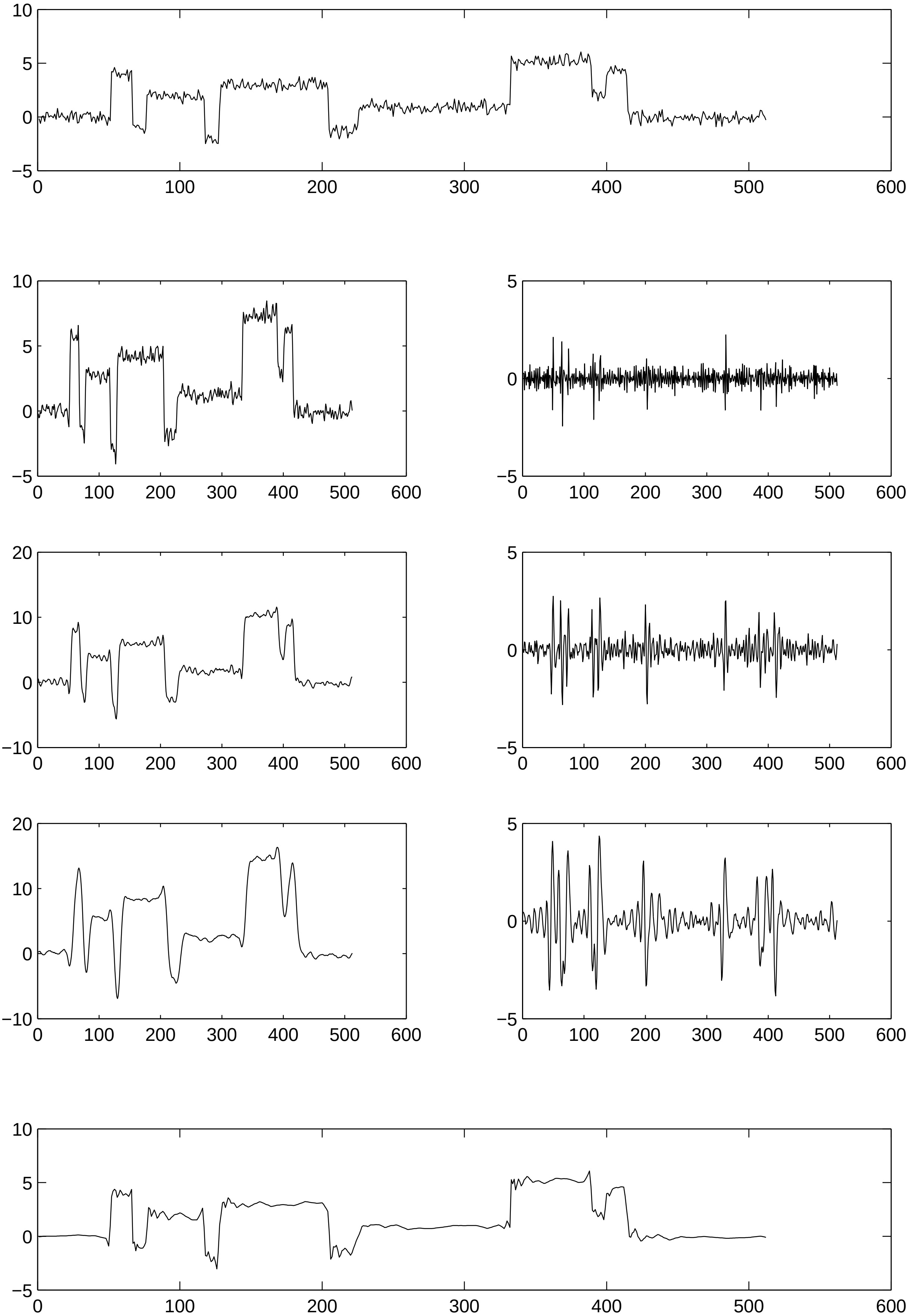

Fig. 4. Noisy Blocks signal denoised using Db3 undecimated wavelet transform

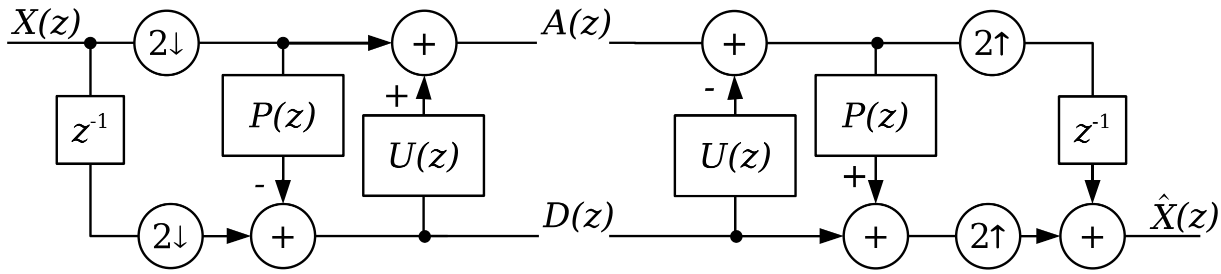

Fig. 5. Filter bank based on the lifting scheme

work, an example of a gradient based decision maps is suggested while in [19] derivative based decision maps are proposed. However, the method is only briefly investigated and no comparison to other methods are provided.

Wu et al. in [20] proposed the algorithm called the switch thresholding. It is a very simple adaptation where two conventional wavelets are used - Haar and CDF(2,2). The adaptation is performed on a point-by-point basis by choosing the appropriate wavelet for the current point. For the step edges, Haar transform is to be used, while for the other parts of the signal CDF(2,2) is chosen. Whether the current point is a step edge is determined by the difference to its neighbouring point. If the difference is larger than the threshold it is considered to be an step edge. Although some coarse limits for the threshold are suggested to follow from the statistics theory, the actual threshold value is determined empirically.

Instead of modifying the transform Chan and Zhou [21] took the opposite approach. They showed that the signal itself can be modified in the place of discontinuities. After the discontinuity is detected, signal values from one side of the discontinuity is used to extrapolate its values to the other side of the discontinuity. Traditional wavelets may now be applied to such a signal region because the discontinuity is eliminated and the whole region is smooth. If we record how the changes are made to the signal, it should be possible to recover the discontinuities at synthesis, using the inverse filters.

V. Intersection of Confidence Intervals

In a separate field, Katkovnik [22] used local polynomial approximation as a filter design tool for denoising. Katkovnik uses kernel estimators to estimate the true value of the noisy signal. Again, the idea is to use larger estimator support for the smooth signal regions and to contract the support for regions which contain high frequencies. To achieve adaptive support selection, he proposed the intersection of confidence intervals (ICI) method. The algorithm calculates confidence intervals (CI) for estimators of growing support. If the order of estimation is kept constant and estimator support grows, the CI gets smaller, as can be seen in Fig. 6. The ICI rule states that as long as the support spans data points with similar statistics the CIs should at least partially intersect. Ones there is no more intersection, it is likely that a change in a signal statistics occurred and estimator support should not include these points. For smaller supports estimation bias is low but the variance is high. As the support grows, estimation bias is getting larger and the variance is getting smaller (Fig. 7) Stankovic proved in [23] that the ICI rule gives an optimal bias-to-variance tradeoff.

VI. Conclusion

Real world signals are often corrupted by noise which may severely limit their usefulness. For this reason, signal denoising is a topic that continually draws great interest. Wavelets perform very well in the area but the denoising efficiency can be significantly improved by introducing adaptivity to traditional wavelet transforms. The lifting framework allows researcher to do it in a straightforward way and many adaptive algorithms were developed to explore this capability. Commonly, the adaptation algorithm tries to contract the wavelet support in the neighbourhood of discontinuities in the signal in order to avoid visual artifacts in the denoised signal, such as a Gibbs phenomena. For smooth signal regions a longer wavelets are used. There are many ways to further improve the transform efficiency and the focus will be put on connecting the ICI rule with the wavelet transforms.

References

[1] S.G. Mallat. Multiresolution approximation and wavelet orthogonal bases of l2(r). 315, 1989.

[2] S.G. Mallat. Multifrequency channel decompositions of images and wavelet models. 37:2091–2110, 1989.

[3] S.G. Mallat. Multiresolution representations and wavelets. 1988.

[4] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(7):674–693, 1989.

[5] David L. Donoho and Iain M. Johnstone. Minimax estimation via wavelet shrinkage. *Annals of Statistics*, 26(3):879–921, 1998.

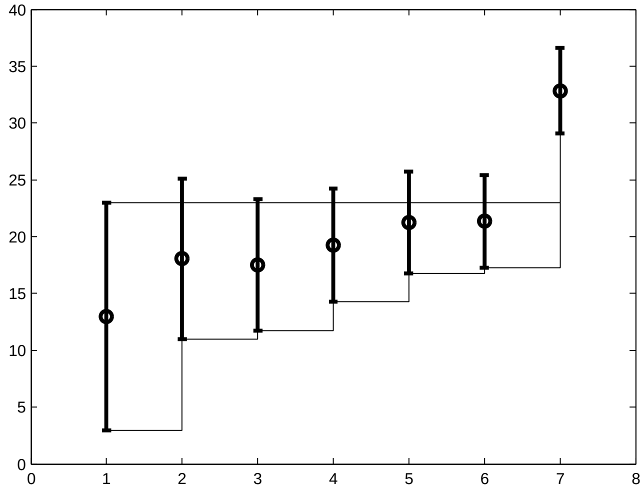

Fig. 6. Confidence intervals for estimators of growing support h

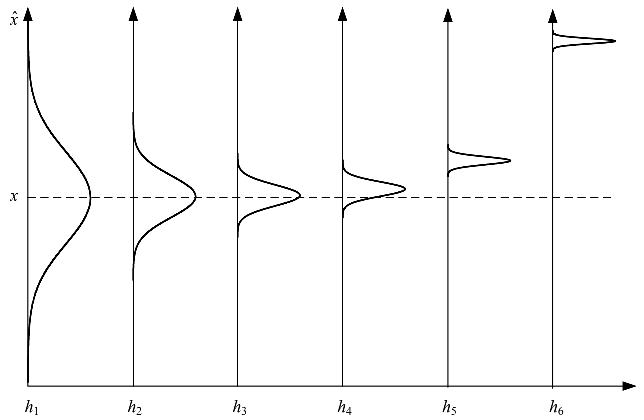

Fig. 7. Probability density functions of $\hat{x}$ for growing estimation support h

[6] David L. Donoho and Iain M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432):1200–1224, ???? 1995.

[7] David L. Donoho and Iain M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[8] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: Asymptopia? *J. R. Statist. Soc. B.*, 57(2):301–337, 1995.

[9] R. R. Coifman and D. L. Donoho. Translation-invariant denoising. 1996.

[10] Wim Sweldens and Peter Schroder. Building your own wavelets at home. 1997.

[11] W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.*, 3(2):186–200, 1996.

[12] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.

[13] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):245–267, 1998.

[14] J. Kovacevic and W. Sweldens. Wavelet families of increasing order in arbitrary dimensions. *IEEE Transactions on Image Processing*, 9(3):480–496, 2000.

[15] R.L. Claypoole, G.M. Davis, W. Sweldens, and R.G. Baraniuk. Nonlinear wavelet transforms for image coding via lifting. 12(12):1449–1459, 2003.

[16] R. Claypoole, G. Davis, W. Sweldens, and R. Baraniuk. Nonlinear wavelet transforms for image coding. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, volume 1, pages 662–667 vol.1, 1997.

[17] R.L. Claypoole, R.G. Baraniuk, and R.D. Nowak. Adaptive wavelet transforms via lifting. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages 1513–1516 vol.3, 1998.

[18] G. Piella and H. Heijmans. Adaptive lifting schemes with perfect reconstruction. 2002.

[19] G. Piella, B. Pesquet-Popescu, and H. Heijmans. Adaptive update lifting with a decision rule based on derivative filters. *IEEE Signal Processing Letters*, 9(10):329–332, 2002.

[20] Yonghong Wu, Quan Pan, Hongcai Zhang, and Shaowu Zhang. Adaptive denoising based on lifting scheme. In *Proc. 7th International Conference on Signal Processing ICSP '04*, volume 1, pages 352–355 vol.1, 2004.

[21] T. F. Chan and H. M. Zhou. Adaptive ENO-wavelet transforms for discontinuous functions. Technical Report 99-21, 1999.

[22] V. Katkovnik. A new method for varying adaptive bandwidth selection. *IEEE Transactions on Signal Processing*, 47(9):2567–2571, 1999.

[23] L. Stankovic. Performance analysis of the adaptive algorithm for bias-to-variance tradeoff. *IEEE Transactions on Signal Processing*, 52(5):1228–1234, 2004.