ELSEVIER

Chemometrics and Intelligent Laboratory Systems 34 (1996) 187–202

Chemometrics and intelligent laboratory systems

Wavelet denoising of Gaussian peaks: a comparative study

C.R. Mittermayr *, S.G. Nikolov, H. Hutter, M. Grasserbauer

Institute of Analytical Chemistry, University of Technology, Getreidemarkt 9 / 151, Vienna 1060, Austria

Received 12 June 1995; revised 10 March 1996; accepted 8 May 1996

Abstract

In this paper we apply some recent results on non-linear wavelet analysis to simulated noisy signals of chemical interest. In particular, we compare the wavelet soft universal thresholding algorithm described by Donoho, to Fourier filters and to polynomial smoothers such as the Savitzky–Golay filters (SG). All reconstruction filters were evaluated on the basis of three different criteria: the mean squared error (MSE) both for the whole signal and for an interval centred around the peak, the signal-to-noise ratio (SNR) improvement and the change in the peak area. The simulated data consists of narrow Gaussian peaks with white noise. Signals with low SNR were investigated, since this is a challenging problem for each reconstruction filter. Four common wavelets (Haar, Daubechies, Symmlets and Coiflets) were selected for the wavelet denoising. Our results show that under the chosen conditions wavelet denoising (WD) gives in most cases superior performance over classical filter techniques.

Keywords: Signal processing; Filter; Chromatography; Fourier transform; Wavelet transform

1. Introduction

As known the main goal of the analytical chemist is the extraction of information from measured data [1]. The achievement of this goal is complicated by the presence of noise. Nowadays nearly every instrument is driven by a computer, it has become common practice to tackle the noise reduction problem by digital processing of signals. In general, digital signal processing methods are called filters, and the methods for the removal of noise are often called smoothers [2].

Since processing digital signals is not only a problem of analytical chemists but is of interest in all fields of science and technology, a large number of filters has been developed in the past decades. In spite of the existence of diverse filters only a few (e.g., Savitzky–Golay, Fourier and Kalman filters) are extensively used among chemists.

One of the most important problems that has to be solved with the application of digital filters is the correct choice of the filter type and the filter parameters. The most difficult choice is that of the cut-off frequency of the filter which has to be specified either explicitly or implicitly. It is often selected arbitrarily or by adopting a certain theoretical model [2–4]. Thus far, only a few comparative studies have been carried out in order to motivate the choice of the filter parameters that are most appropriate for a specific problem [2,4–9]. Two approaches based on criteria from information theory (maximum Shannon

* Corresponding author. Fax: +43-1-5867813; e-mail: cmitter@fbch.tuwien.ac.at

PII S0169-7439(96)00026-3

information entropy, Akaike's information theoretic criterion) [2,5] for the optimal selection of filters have been recently published.

During the last decade a new and very versatile technique, the wavelet transform (WT), has been developed as a unifying framework of a number of independently developed methods [10–12]. The WT has found a lot of applications in signal and image processing [13–15] and has already become a standard method in data compression (e.g., FBI finger print standard [16]). With a few exceptions [17–21] this development has practically escaped the attention of the chemists.

The WT decomposes a signal into a set of basis functions, called a wavelet basis, which are obtained from a single prototype wavelet by dilation and translation. By analogy with the Fourier transform (FT), WT maps the signal into another domain, in this case the time-scale (frequency) domain. However, the FT spreads the information of small or suddenly changing features over a wide frequency range, while the WT is localized in time, which makes it more appropriate for applications to transient, non-periodic signals [13]. This localization in time also allows the removal of noise-dependent high frequencies, while conserving the signal bearing high frequency terms of the signal. Failing to sort out the frequencies results in signal distortion and broadening of the signal after filtering, which can easily be explained in the terms of the FT [5]. Donoho [22,23] has developed reliable statistical criteria for wavelet filtering. By universal thresholding (an algorithm they have invented) nearly minimax mean square error (MSE) is achieved.

This study is intended to explore the wavelet universal thresholding algorithm for denoising data and to compare its performance with that of other commonly used smoothing filters in chemistry. First we shortly describe the WT and the other filter techniques (the Fourier filter and SG filters). By Fourier filtering (FF), boxcar filtering in the frequency domain is meant in this paper. Then, the parameters of the simulated data are given. The Gaussian function has been chosen as the ideal model of measured peaks because it is commonly used to represent peaks and bands in chromatography [24] and spectroscopy. As noise model white noise is used, since it is fundamental due to the nature of light and matter and it is present in every signal [25,26]. Several quantities, i.e., SNR, MSE and peak area preservation, are used to measure the quality of reconstruction by the different filters. Finally, the results are discussed with respect to these quality criteria.

2. Theory

2.1. Filters

If the measured data are both slowly varying and also corrupted by random noise, it is useful to smooth them by replacing each data point with some kind of local average of the surrounding data points. Since nearby points have almost the same values as the point of interest, averaging reduces the noise level without biasing the value obtained.

Let us have a series of equally spaced data points $y_i = f(t_i)$, where $t_i = t_0 + i\Delta$ for some constant sampling Δ and $i = \ldots -2, -2, 0, +1, +2, \ldots$ A smoothing filter replaces each data point y_i by a linear combination $\hat{y}_i$ of itself and some neighbouring points

$$\hat{y}_i = \sum_{n=-n_L}^{n_R} c_n y_{i+n} \tag{1}$$

where n_L is the number of points used to the left of the data point i while n_R is the number of points to the right [27].

2.1.1. Savitzky–Golay filters

The SG filters [7,28] approximate the underlying function by a polynomial of order M.

For each point y_i a least-squares fit is made for all the points in the moving window and then $\hat{y}_i$ is set to the value of the polynomial at position i. It may be easily shown that the filter coefficients c_n for which Eq. (2) accomplishes polynomial least-squares fitting inside a moving window may be precalculated and stored, since the polynomially fitted coefficients are themselves linear in the values of the data. The general formula for calculating the SG filter coefficients is [29]:

$$c_n = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{e}_n) \tag{2}$$

where $A_{ij} = i^j$, $i = -n_L, \ldots, n_R$, $j = 0, \ldots, M$, is the design matrix and M is the order of the polynomial.

Pairs of even and odd order filters such as 0 and 1, 2 and 3, etc. have the same coefficients. Some typical SG filter coefficients may be found in [28,30].

The simplest SG filter, for $M = 1$ is the moving window average (MWA) filter where all c_n are constant:

$$c_n = \frac{1}{(n_L + n_R + 1)} \quad (3)$$

If the underlying function is a constant or if it changes linearly with time, then no bias is introduced in the MWA result.

2.1.2. Fourier filter

The Fourier transform [31] is a technique for determining the frequency spectrum of a signal (this can also be a spatial frequency when spectra are considered). The signal of interest can be distinguished from the noise in the frequency domain due to the fact that the signal varies slowly in comparison to the noise. Therefore, the information on the signal resides in the low frequency coefficients. By setting all coefficients above a certain frequency f_c (the cutoff frequency) to zero one tries to remove only noise, while preserving the information on the signal. Gaussian white noise by definition has a constant contribution to all frequencies and thus one has to compromise when setting the cutoff frequency [5]. Some factors which influence the decision are the signal shape, the desired amount of filtering and the tolerable distortion. On the other hand, often it is difficult to guess whether a change in the filtered signal is due to the removal of noise or already due to removal of information from the signal.

2.1.3. Wavelet denoising via thresholding

2.1.3.1. Wavelets. A multiscale representation [12] that has recently attracted great interest in the signal processing community is the wavelet transform, which dates back to Strömberg [32] and Meyer [10]. The two-parameter family of translated and dilated functions

$$\psi_{a,b}(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right) \quad a, b \in R,\ a \neq 0 \quad (4)$$

defined from a single mother function ψ: $R \rightarrow R$ is called a wavelet basis (see Fig. 1). The function ψ is

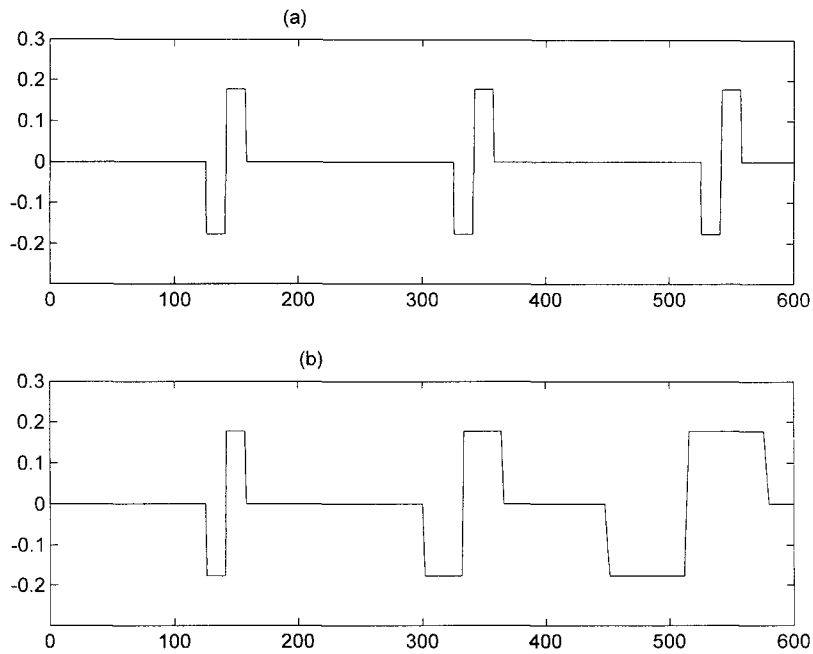

Fig. 1. Haar wavelet. (a) Translations (different b values, refer Formula 4), (b) dilations (different a values, refer Formula 4).

called a wavelet (a short wave) which comes from the requirements that it should integrate to zero, 'waving' above and below the x-axis. Throughout this paper several mother wavelets including the Haar wavelet, Daubechies wavelets, Coiflets and Symmlets have been used for signal denoising (see Fig. 2). The basic idea of the wavelet transform is to represent any arbitrary function f as a superposition of wavelets. The continuous wavelet transform (CWT) of f is given by

$$(Wf)(a,b) = \langle f, \psi_{a,b} \rangle = |a|^{-1/2} \int_{x \in R} f(x) \psi\left(\frac{x-b}{a}\right) \mathrm{d}x \quad (5)$$

A particularly well studied problem is the construction of orthogonal wavelets for discrete signals, which permits to obtain a compact non-redundant multiscale representation of the signal data. A large number of signals are better approximated by substantially fewer wavelet basis functions than by sine and cosine functions for instance. Fast implementation of the discrete wavelet transform (DWT) developed by Mallat [12] has made the wavelet methods a tool that is a particularly valuable for processing of analytical data.

2.1.3.2. Denoising via thresholding. In our present study we have used the signal denoising algorithm proposed by Donoho [22,23]. For a signal $y_i = f(t_i) + \sigma z_i$, $i = 1, \cdots, n$, where z_i is white Gaussian noise and $n = 2^{J+1}$, we can use a 1D pyramidal filtering by the following three steps: Perform the forward pyramid wavelet transform of Cohen et al. [33] to the normalized data $y_k/\sqrt{n}$, yielding noisy wavelet coefficients $w_{j,k}$, $j = j_0, \cdots, J$, $k = 0, \ldots, 2^j - 1$; apply universal soft thresholding to the noisy wavelet coefficients, yielding estimates $\hat{w}_{j,k}$; make the inverse wavelet transform, that produces the reconstruction $\hat{y}_i$.

Different thresholding methods like

$$\hat{w}_{jk}^{\text{hard}} = \begin{cases} 0, & w_{jk} < t \\ w_{jk}, & w_{jk} \geq t \end{cases} \quad \text{hard thresholding} \quad (6)$$

$$\hat{w}_{jk}^{\text{soft}} = \text{sign}(w_{jk})\left(|w_{jk}| - t\right)_+ \quad \text{soft thresholding} \quad (7)$$

and universal thresholding (hard or soft thresholding with $t = \sigma\sqrt{2\log(n)}/\sqrt{n}$, where σ is the variance of the Gaussian noise and n is the size of the data set) and others have been used to solve problems ranging

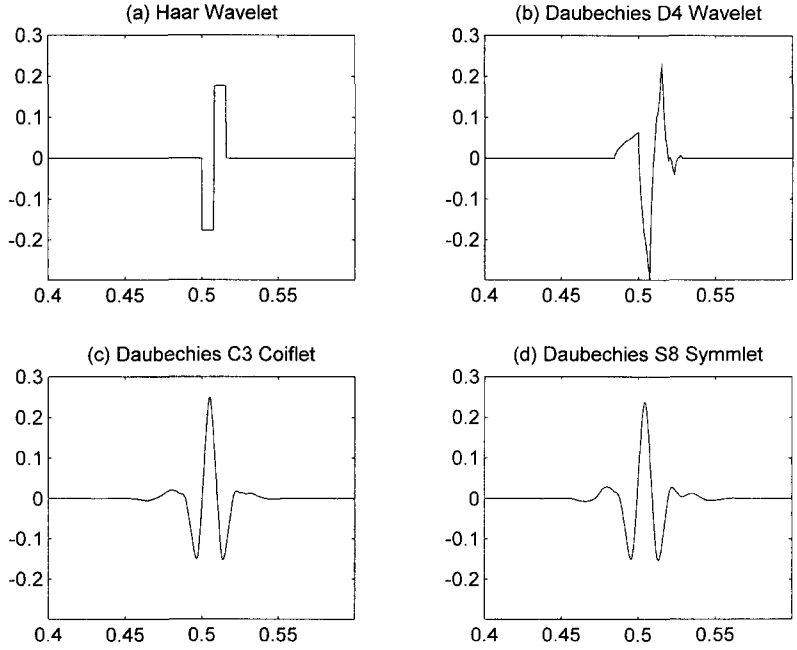

Fig. 2. Various wavelets. (a) Haar wavelet (HA), b) Daubechies wavelet (D8), (c) Coiflet (C3), (d) Symmlet (S8).

from Gaussian noise removal to density estimation and inverse problems [23].

The above-described wavelet domain filter shrinks the wavelet coefficients to zero. Thus, because the few large wavelet coefficients preserve almost the whole energy of the signal, the shrinkage removes the noise without distorting the signal peaks. By analogy with the FT, the narrower a peak, the higher the frequencies, that are required to describe it. Frequency information is contained in the different scales (though at an exponential scale) of the multiscale wavelet transform representation of the signal, where high frequency information resides in the fine levels and low frequencies in the coarse levels [12,21]. Therefore, optimal denoising depends on the amount of noise (usually contained in the finer levels) and the size of the features of interest.

2.2. Measures

The results which were obtained by different denoising filters were evaluated using several mathematical measures: (i) mean square error (MSE), (ii) signal-to-noise ratio (SNR) improvement, (iii) peak area change.

The MSE is an estimator showing how close the reconstruction is to the original signal

$$\mathrm{MSE} = \frac{1}{n}\sum (y - \hat{y})^2 \tag{8}$$

The SNR improvement is the ratio between the SNR of the reconstructed signal and the SNR of the original signal. The SNR is usually calculated by dividing the maximum of the signal by the standard deviation of the noise. The SNR improvement is of special interest, since it lowers the detection limit (which is defined as n times the standard deviation, where n is usually chosen to be equal to 3) and aids the interpretation of the signal.

Quantisation in many cases relies on integrating the area under the peak in a signal $A = \int_a^b y \, \mathrm{d}y$, therefore filtering should not change the area under the peak. Proper detection of the integration boundaries a and b, however, is often very difficult, especially in the presence of noise [34].

3. Experimental

3.1. Simulation

3.1.1. Peak model

Simulated data sets were used to compare the performance of the various methods. They should resemble the real data as close as possible. In chromatography for instance eluting peaks are modeled by a Gaussian function [34]:

$$f(t) = A\exp\left[\frac{1}{2}\left(\frac{t - t_0}{\sigma}\right)^2\right] \tag{9}$$

All the simulated signals have only one peak and a size of 1024 data points. For the time axis integer increments were used. All the peaks were centred at data point 512. Six Gaussian shaped signals with different standard deviations (1, 2, 4, 8, 16 and 24) were generated. The corresponding full width at half maximum (fwhm) can be calculated from the relation:

$$\mathrm{fwhm} = 2\sigma\sqrt{2\ln(2)} = 2.3548\sigma \tag{10}$$

3.1.2. Noise model

Ten sequences of Gaussian noise with zero mean and unit standard deviation were created using a routine provided by MATLAB [1]. Noisy signals were obtained by adding the noisy sequences to the signal generated as described above. A predefined SNR was achieved by scaling the noise record before adding it to the signal. The SNR of the six data sets, defined as the peak height divided by the standard deviation for the noise, was 2, 3, 4, 5 and 10, respectively. We decided to investigate the results of various filters at relatively low SNR because such signals are often recorded and they are very difficult to interpret.

3.1.3. Mean squared error

The MSE was calculated both for the whole signal ($\mathrm{MSE_s}$) and for the peak area ($\mathrm{MSE_p}$). By peak area in this case we mean the interval of $\pm 4\sigma$ centred around the peak maximum. For a Gaussian peak this interval encompasses approximately 99.99% of the peak area. $\mathrm{MSE_p}$ shows how close the recon-

[1] MATLAB, The Mathworks, Natick, MA.

structed peak is to the original peak. If the whole signal is taken when calculating the MSE, the result is dominated by the background noise.

In order to asses the overall quality of a filter for signals of different peak shapes we used the mean value of the MSE_p for all different peak widths (MSE_{6p}). This can be considered as the MSE of a signal consisting of 6 different Gaussian peaks, where only the $\pm 4\sigma$ regions around the peak centres were taken in account. Since the MSE_p of the narrowest peak is about 10 times larger than that of the broadest, there is a strong bias to the narrower peaks. That is why we also introduced the weighted mean of the MSE_p (MSE_{6wp}). Each weight is equal to the number of data points for calculation of the MSE_p.

3.1.4. Signal to noise ratio improvement

The standard deviation of the noise was obtained from the first 400 data points of the signal, which contain only noise. To evaluate the overall performance of a filter for signals of different peak shapes we used the mean of the SNR improvements for all different peak widths (SNR_6). This is possible since the SNR before filtering is constant and the standard deviation of the baseline noise of the filtered signals is independent of the peak widths (it is calculated from an area where no peak is present). Therefore, SNR_6 can be considered as a normalized sum of the peak maxima after filtering.

3.1.5. Area change

The reference area was calculated from the data points of the original signal which correspond to $h = 10\%$ of the peak height. Since the moving window average and the Savitzky–Golay filter shift in the peak location and cause peak broadening, different integration boundaries have to be used depending on the signal shape and the filter length in order to preserve the peak area. This problem is illustrated in Fig. 3, where (a, b) are the reference peak area boundaries and (a', b'), the integration boundaries determined after filtering the original signal, so that the difference

$$\left| \int_a^b y \, dy - \int_{a'}^{b'} y' \, dy \right| < \epsilon_d \tag{11}$$

will be minimal. Two functions a' and b'

$$a' = \theta_a(l,\sigma,h), \quad b' = \theta_b(l,\sigma,h) \tag{12}$$

where l is the filter width and σ is the peak parameter, were numerically determined, which give the relation between the reference boundaries and the new boundaries (a', b'). The same boundaries (a', b') were used to calculate the peak area after filtering the noisy signal. Therefore the peak area change is defined as

$$\Delta A = A - \hat{A} = \int_a^b y \, dy - \int_{a'}^{b'} \hat{y} \, dy \tag{13}$$

In the case of wavelet filtering the new integration boundaries are equal to the reference boundaries, because there is no shift and almost no broadening of the peak (ϵ_d is less than 0.5% of A in (Eq. (11)), when $a' = a$ and $b' = b$).

3.1.6. Filter parameters

The polynomial filters (PF) have been implemented by the FILTER function of MATLAB. The filter coefficients have either been taken from literature [28,30] or calculated as specified by Eq. (2). The

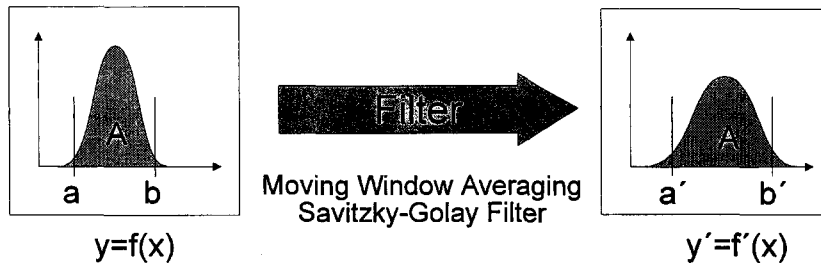

Fig. 3. Peak area change. A, a, b are the original peak area and boundaries, respectively; A', a', b' are the filtered peak area and boundaries, respectively.

filter widths used are: fw = 5 + 2n ($n = 1, \cdots, 30$) and fw = 69 + 8m ($m = 1, \cdots, 10$). The filter orders are 1, 3 and 5, respectively.

The cutoff frequencies f_c varied from $i = 2$ to 510 with an increment of 4 (i is the index of the Fourier coefficients). All frequency coefficients f_i, for which $|i| \geq c$, are set to zero.

To reconstruct the simulated data the following wavelets were chosen: Haar, Daubechies 4 and 6, Symmlet 6 and 8, and Coiflet 2 and 3 (Fig. 2). All these wavelets are part of the TeachWave Toolbox [2]. The level down to which universal thresholding was applied, was coded as the last symbol in the wavelet code. Since the signal length is $n = 1024 = 2^{10}$, the finest level of decomposition is $J = \log_2(2^{10-1}) = 9$. The coarsest level we have used is $j = 3$.

The following notation has been used when coding the filter parameters: the order of the polynomial filters is coded as the first digit after 'p', followed by an underscore and the width of the filter. The Fourier filter is represented by an 'f' followed by the cutoff frequency. The various wavelets are coded as follows: Haar (ha), Daubechies 4 and 6 (d4, d6), Symmlets 6 and 8 (s6, s8) and Coiflets 2 and 3 (c2, c3) and the lowest decomposition level is given after the underscore.

The results of this study are available on the WWW server: http://www.iac.tuwien.ac.at/~cmitter/wave_denoise/gauss_peak.html

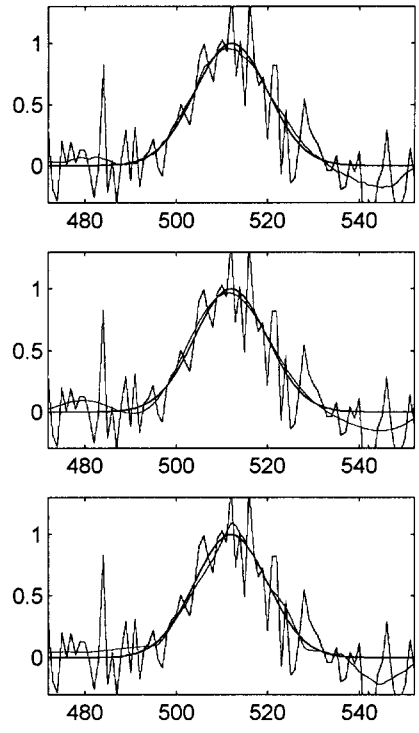

Fig. 4. Example reconstructions, SNR = 4. Top: Savitzky–Golay reconstruction (filter width = 45, $M = 5$); middle: Fourier reconstruction (cutoff frequency = 42); bottom: wavelet reconstruction (Daubechies 6, level 6). In each plot the original signal is drawn with a solid line (————), the noisy signal with a dashed line (- - -) and the reconstruction with a dotted line (· · ·).

3.2. Computation

All computations were performed on a 90 MHz Pentium DECpc computer with 32 MB RAM. The authors have used TeachWave, a MATLAB wavelet toolbox developed by Donoho and his team from Stanford University, to process all data sets. In addition, a library of scripts for simulation of artificial data, noise removal and calculation of various statistical measures was created. An example of three reconstructions is given in Fig. 4.

The computational complexity of the DWT is $O(n)$ while that of the FFT is $O(n \log_2(n))$. This makes wavelet denoising a reconstruction algorithm that is accessible on a wide range of computer platforms. The average denoising time for a discrete signal of $2^{10} = 1024$ data points on the above mentioned computer is about 0.38 s (including forward DWT, thresholding and inverse DWT).

4. Results

4.1. MSE

The MSE was calculated both for MSE_s and for MSE_p. Since in some cases the peak represents only a small part of the whole signal, there is a significant difference in the results obtained.

The MSE_s shows a maximum depending on the signal shape (Table 1). This trend is also affected by the SNR, so that the maximum is shifted to narrow

[2] D. Donoho et al., TeachWave version 0.550, Copyright (1993), Stanford University.

Table 1
Optimal mean and its standard deviation (using 10 replicates) for MSE_s obtained by MWA filters (polynomial filter $M = 1$) for different σ and SNR values

SNR σ	2	3	4	5	10	2	3	4	5	10
1	0.3321 ± 0.0835	0.249 ± 0.0381	0.2198 ± 0.0221	0.2058 ± 0.0128	0.1805 ± 0.0059	p1_149	p1_149	p1_149	p1_125	p1_49
2	0.5066 ± 0.0882	0.4218 ± 0.0351	0.386 ± 0.0221	0.3664 ± 0.0197	0.1799 ± 0.0139	p1_149	p1_125	p1_93	p1_63	p1_9
4	0.8142 ± 0.0779	0.6596 ± 0.0792	0.457 ± 0.0511	0.3387 ± 0.0370	0.13 ± 0.0127	p1_109	p1_25	p1_19	p1_15	p1_11
8	0.871 ± 0.1599	0.5006 ± 0.0745	0.3309 ± 0.0455	0.2393 ± 0.0320	0.0896 ± 0.0118	p1_37	p1_29	p1_23	p1_21	p1_15
16	0.6077 ± 0.1396	0.3328 ± 0.0666	0.2173 ± 0.0386	0.1563 ± 0.0253	0.0572 ± 0.0077	p1_47	p1_39	p1_33	p1_31	p1_23
24	0.4706 ± 0.1069	0.2549 ± 0.0541	0.1647 ± 0.0321	0.1176 ± 0.0214	0.0426 ± 0.0062	p1_57	p1_47	p1_41	p1_37	p1_29

The right part of the table displays the filter width, which gives optimal performance.

peaks for higher SNR, independent of the filter chosen.

In the case of the narrowest peaks almost the highest available width of the polynomial filter produces the smallest MSE_s (Table 1), which is not due to a better peak approximation but rather of the baseline (which results in flattening of the peak).

The broader the peaks become, the more they contribute to the MSE_s. When the SNR becomes larger the contribution of noise to the MSE_s is reduced, so that even narrow peaks have a major role in the MSE_s.

The Fourier filter fits also mainly the baseline by using very low cutoff frequencies, where especially the narrowest peaks ($\sigma = 1$ and 2) are completely extinguished when using $f_c = 2$. The narrower the peak the wider it spreads in the frequency domain, therefore, its contribution can not be distinguished from noise anymore. The f_c resembles the inverse polynomial filter width.

The results for the wavelets show the same tendency, although in this case different wavelet functions are used. For narrow peaks Symmlets produce best results, at the MSE_s maxima Daubechies are optimal whereas regarding broader peaks with higher SNR Coiflets give lowest MSE_s.

In order to investigate the reconstruction of the peak only, we have calculated the MSE_p (Table 2). It is easily seen that the MSE_p decreases when peaks become broader and when the SNR increases (Fig. 5). Our results agree with some publications [7,27] where the optimal filter width (for SG filters) is found to be between one and two times the fwhm of the peak (Table 2, top). For high SNR the best filter width is about equal to the fwhm, while for decreasing SNR larger filter widths are necessary to remove the noise more efficiently. Another trend is that the higher the order of the filter the broader the optimal filter, which follows from the impuls response of the filter. Again for the Fourier filter the same trend as for the polynomial filters can be observed (Table 2, middle).

Our investigation shows that Symmlets 6 or Daubechies 6 wavelets are a good choice in the case of narrow peaks, while Coiflets or Symmlet 8 are more appropriate for broader peaks (Table 2, bottom). When the SNR is very low, Daubechies 6 gives the best approximation.

In the case of wavelet denoising, broader peaks are optimally smoothed down to coarser levels, since the broader the peak is, the more information is located in coarser levels of the multiscale representation and hence more finer levels can be thresholded, without distorting the signal too much. To get optimal results at higher SNR it is necessary to smooth only the finer levels, where the noise is situated.

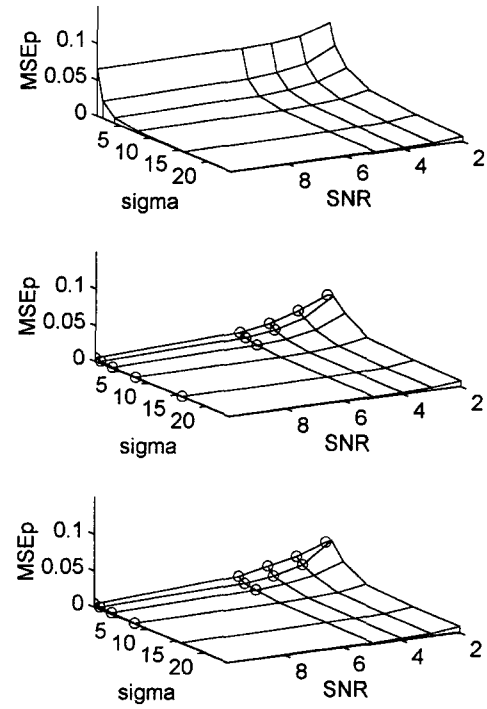

Fig. 5. This plot shows the MSE_p as a function of the peak width and the initial SNR for different filters. The points, where wavelet and Fourier MSE_p is significantly better than the SG MSE_p are marked on the plot. Top: MSE_p after filtering with a Savitzky–Golay filter; middle: MSE_p after filtering with a Fourier filter; bottom: MSE_p after wavelet filtering.

The filter widths of the polynomial filters can be compared to the f_c of the Fourier filter and the levels in the wavelet denoising. It follows from theory that high levels in the WT correspond to high frequencies in the FT [13]. When one looks at the impuls response of polynomial filter with different filter lengths [5], it is easily seen that a polynomial filter is a lowpass filter to which some kind of cutoff frequency can be attributed. This cutoff frequency is inversely proportional to the filter width.

Table 2
Optimal mean and its standard deviation (using 10 replicates) for MSE_p obtained

SNR σ	2	3	4	5	10	2	3	4	5	10
SG										
1	8.6722 ± 4.1361	7.3813 ± 2.8004	6.9178 ± 2.0642	6.7184 ± 1.6374	6.4967 ± 0.8087	p5_15	p3_9	p3_9	p3_9	p3_9
2	5.8207 ± 2.8343	3.9233 ± 1.9090	3.1754 ± 1.2175	2.8463 ± 0.9925	2.3958 ± 0.4680	p1_7	p3_11	p3_11	p5_15	p3_9
4	3.4475 ± 1.8219	2.0858 ± 1.1078	1.4879 ± 0.9006	1.1937 ± 0.7048	0.783 ± 0.3192	p1_11	p1_9	p3_17	p5_23	p3_13
8	1.7325 ± 0.9603	0.9789 ± 0.5955	0.6771 ± 0.4481	0.5183 ± 0.3473	0.2817 ± 0.1618	p3_35	p3_33	p5_45	p5_43	p5_39
16	1.0107 ± 0.5782	0.5285 ± 0.2666	0.3361 ± 0.1698	0.2401 ± 0.1198	0.0999 ± 0.0450	p5_101	p3_57	p5_85	p5_77	p5_69
24	0.695 ± 0.3456	0.3523 ± 0.2078	0.217 ± 0.1303	0.1518 ± 0.0879	0.0578 ± 0.0312	p3_93	p5_125	p5_117	p5_117	p5_101
Fourier										
1	4.961 ± 2.7409	3.3569 ± 2.4088	2.1635 ± 1.1249	1.4411 ± 0.7089	0.4868 ± 0.1494	f162	f166	f286	f286	f286
2	5.1695 ± 2.7097	2.7427 ± 1.7584	1.5817 ± 0.9735	1.0504 ± 0.6119	0.3269 ± 0.1824	f98	f146	f146	f146	f158
4	3.4884 ± 1.2930	1.7537 ± 0.7889	1.0114 ± 0.4672	0.6667 ± 0.3153	0.1874 ± 0.0951	f62	f86	f86	f86	f98
8	1.5301 ± 0.6210	0.7778 ± 0.3370	0.4638 ± 0.2526	0.3157 ± 0.1732	0.0842 ± 0.0450	f34	f38	f42	f42	f50
16	1.0051 ± 0.3999	0.4848 ± 0.1578	0.2886 ± 0.0868	0.1982 ± 0.0571	0.0541 ± 0.0218	f18	f22	f22	f22	f30
24	0.7582 ± 0.3925	0.3659 ± 0.1875	0.2076 ± 0.1055	0.1342 ± 0.0676	0.0363 ± 0.0173	f14	f18	f18	f18	f18
Wavelet										
1	4.7111 ± 1.1798	3.2625 ± 0.5351	2.5258 ± 1.6633	1.6455 ± 1.0599	0.4753 ± 0.2595	d6_8	d6_8	s6_9	s6_9	s6_9
2	5.2899 ± 2.7363	2.494 ± 1.5937	1.4644 ± 0.8985	0.9872 ± 0.5767	0.3491 ± 0.1477	d6_7	s6_8	s6_8	s6_8	s6_8
4	3.0664 ± 0.8560	1.6749 ± 0.8539	1.0248 ± 0.4859	0.7239 ± 0.3150	0.2261 ± 0.0976	d6_6	d6_7	d6_7	d6_7	s6_8
8	1.2775 ± 0.4890	0.6607 ± 0.2197	0.4458 ± 0.1253	0.3467 ± 0.0814	0.1081 ± 0.0360	d6_6	d6_6	d6_6	d6_6	s8_7
16	0.8644 ± 0.5462	0.4641 ± 0.2416	0.3225 ± 0.1367	0.257 ± 0.1471	0.069 ± 0.0368	c2_5	c2_5	c2_5	s8_6	s8_6
24	0.7896 ± 0.4334	0.4183 ± 0.2732	0.2561 ± 0.1541	0.1806 ± 0.0995	0.0656 ± 0.0355	d6_4	s8_5	s8_5	s8_5	s8_6

Top: polynomial filters ($M = 1,3,5$); middle: Fourier filter; bottom: best wavelet filter, for different σ and SNR.
The right part of the table displays the optimal filter parameters.

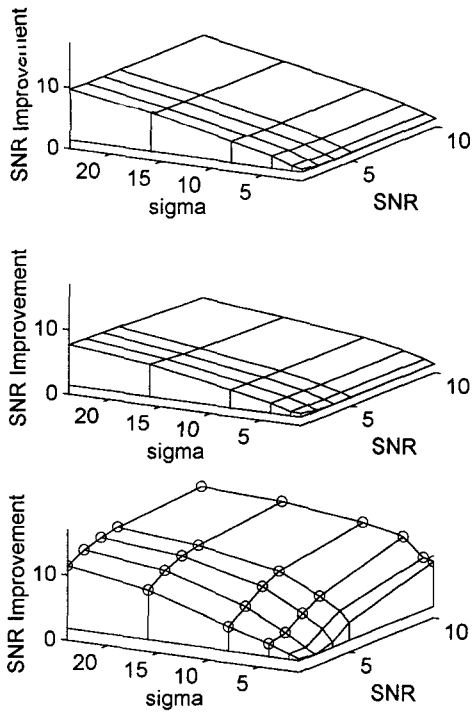

Fig. 6. This figure displays the SNR improvement as a function of the peak width and the initial SNR for various filters. The points, where wavelet SNR is significantly better than the other filters are marked on the plot. Top: SNR improvement after filtering with a Savitzky–Golay filter; middle: SNR improvement after filtering with a Fourier filter; bottom: SNR improvement after wavelet filtering.

4.2. SNR

Broader peaks yield a higher SNR improvement, independent of the filter used (Fig. 6 and Table 3). For the polynomial filters the SNR improvement is not correlated with the initial SNR (Table 3, top). The optimal filter widths are about 2 to 3 times the fwhm of the peak for the 3rd or 5th order filters, respectively, which agrees with Enke and Nieman [7]. There is nearly no dependence of the f_c of the Fourier filter on the initial SNR. The broader the peak the lower the f_c (Table 3, middle).

The best wavelet with respect to the SNR improvement is the Daubechies 4 (6 is best but one) at the coarsest possible level (we specified as 3), except of narrow peaks at low SNR, where smoothing only the finest levels gives the best SNR improvement (otherwise the signal would be distorted), and narrow peaks of high SNR where Symmlets are preferable (Table 3, bottom).

Wavelet denoising gives a 2–5 fold higher improvement of the SNR compared to all other filters in the case of narrow signals (1–4σ) and high SNR (3–10). This is achieved at the cost of a high distortion of the signal so the peak can be seen only qualitatively, which is fine for peak detection.

4.3. Area change

The error of the peak area determination can be expressed as

$$\Delta A = \epsilon_d + \epsilon_f + \epsilon_n + \epsilon_r \qquad (14)$$

where ϵ_d is the error due to the discrete approximation of A, ϵ_f is the error caused by the fitting of the boundaries with functions d' and b' (Eq. (12)), ϵ_n is the error due to the noise fluctuations and ϵ_r is the error introduced by the filter due to the imperfect reconstruction of the original signal. We have studied the first three types of error separately and our results show that their average magnitudes are 0.09%, 0.05% and 1% of A respectively for the MWA filter, and 0.37%, 0.22% and 1% of A, respectively, for the SG filter ($M = 3$). The results we have achieved are not robust, because the total error ΔA is dominated by the noise (Table 4). An increase of the initial SNR or/and peak widths significantly reduces the total peak area error, but the results are still not robust. All filters give similar performance, therefore no particular filter may be recommended as optimal in terms of the peak area preservation.

4.4. Comparison of the filters

To compare the results of the individual filter types a t-test has been performed between the optimal filters for each combination of σ and SNR. The critical t-value was computed for a significance level of 95% and 10 replicates.

Fourier filtering and wavelet denoising gave significantly lower MSEp than the best polynomial filter for the upper triangle of the data matrix (small σ and high SNR), while there was not significant difference between FF and WD (Table 3). In all other cases the optimized filters gave equally good performance in statistical terms.

Table 3
The optimal mean and its standard deviation (using 10 replicates) for the SNR improvement obtained

SNR σ	2	3	4	5	10	2	3	4	5	10
SG										
1	1.8587 ± 0.8450	1.4637 ± 0.1752	1.3988 ± 0.1520	1.3785 ± 0.1314	1.3438 ± 0.0959	p1_149	p3_5	p3_5	p3_5	p3_5
2	2.15 ± 0.2744	2.0066 ± 0.1956	1.9562 ± 0.1772	1.9281 ± 0.1674	1.8901 ± 0.1563	p3_11	p3_11	p3_11	p3_11	p3_11
4	3.1818 ± 0.3898	3.0162 ± 0.3105	2.9409 ± 0.2860	2.9023 ± 0.2840	2.8335 ± 0.3106	p5_45	p5_45	p5_45	p5_45	p5_45
8	4.5629 ± 0.9154	4.4756 ± 0.8683	4.4411 ± 0.7905	4.4214 ± 0.7866	4.3872 ± 0.7897	p3_59	p5_85	p5_77	p5_77	p5_77
16	7.3753 ± 2.1359	7.2375 ± 1.9999	7.1783 ± 1.9340	7.147 ± 1.9024	7.0885 ± 1.8397	p5_149	p5_149	p5_149	p5_149	p5_149
24	9.723 ± 3.3103	9.5665 ± 3.1244	9.5182 ± 3.1069	9.49 ± 3.0975	9.4361 ± 3.0819	p3_149	p3_149	p3_149	p3_149	p3_149
Fourier										
1	1.9305 ± 2.6481	1.492 ± 0.1455	1.4297 ± 0.1357	1.4062 ± 0.1217	1.3635 ± 0.0996	f2	f222	f222	f222	f222
2	2.0544 ± 0.2266	1.957 ± 0.2956	1.9289 ± 0.2356	1.9171 ± 0.2311	1.8954 ± 0.2213	f170	f102	f98	f98	f98
4	2.9534 ± 0.3399	2.8384 ± 0.2973	2.7952 ± 0.2862	2.77 ± 0.2915	2.7075 ± 0.2853	f50	f46	f46	f46	f70
8	4.1897 ± 0.6989	4.1131 ± 0.6795	4.0607 ± 0.6697	4.01 ± 0.6575	3.8861 ± 0.6495	f26	f26	f26	f26	f34
16	6.4496 ± 1.8478	6.1563 ± 1.5834	5.8442 ± 1.3632	5.6711 ± 1.0467	5.2095 ± 0.8676	f14	f14	f14	f18	f22
24	7.6778 ± 2.2575	7.5354 ± 2.0942	7.4055 ± 1.9712	7.2659 ± 1.8548	6.4113 ± 1.3149	f14	f14	f14	f14	f14
Wavelet										
1	2.1314 ± 0.4382	1.8094 ± 0.2120	2.5971 ± 4.2143	4.5677 ± 7.1483	8.6895 ± 7.5212	d4_8	d4_8	s6_3	s6_3	s6_3
2	2.6622 ± 0.3577	2.6027 ± 4.0808	4.4483 ± 6.4137	6.1522 ± 7.8723	9.0757 ± 2.6080	d4_7	s6_3	s6_3	s8_3	d4_3
4	3.7915 ± 0.5054	4.3835 ± 4.9530	6.3103 ± 2.3885	7.9781 ± 2.4886	11.844 ± 3.4909	d6_6	s8_3	d4_3	d4_3	d4_3
8	5.5507 ± 0.9542	7.599 ± 3.1648	9.6125 ± 3.5078	10.898 ± 3.7394	13.158 ± 4.0172	d6_5	d4_3	d4_3	d4_3	d4_3
16	9.365 ± 3.4632	11.411 ± 3.810	12.547 ± 4.0164	13.101 ± 4.0647	14.628 ± 4.2636	d4_3	d4_3	d4_3	d4_3	d6_3
24	11.429 ± 3.7850	12.811 ± 4.034	13.629 ± 4.0962	14.175 ± 4.1172	15.175 ± 4.1843	d4_3	d4_3	d4_3	d4_3	d4_3

Top: polynomial filters ($M = 1,3,5$); middle: Fourier filter; bottom: wavelet filter, for different σ and SNR.
The right part of the table displays the optimal filter parameters.

Table 4
The results obtained by polynomial filter ($M = 3$) for different σ and SNR values

SNR σ	2	3	4	5	10	2	3	4	5	10
1	3.3173 ± 7.5441	2.1857 ± 9.4119	0.7378 ± 7.0589	0.131 ± 5.6471	0.7167 ± 4.1139	p3_5	p3_9	p3_9	p3_9	p3_13
2	2.2342 ± 10.029	1.4493 ± 10.543	0.2304 ± 7.9077	0.2097 ± 24.092	0.7795 ± 3.0163	p3_9	p3_15	p3_15	p3_149	p3_13
4	0.1936 ± 27.826	0.2215 ± 9.7143	0.2141 ± 11.020	0.0633 ± 10.732	0.1391 ± 2.8354	p3_141	p3_15	p3_117	p3_133	p3_23
8	0.243 ± 7.6034	0.0516 ± 5.1094	0.0031 ± 3.8358	0.1657 ± 3.0686	0.5032 ± 1.5343	p3_21	p3_23	p3_25	p3_25	p3_25
16	0.1659 ± 7.9579	0.2004 ± 5.1033	0.0095 ± 3.8275	0.0133 ± 2.8395	0.2125 ± 1.3066	p3_125	p3_117	p3_117	p3_15	p3_29
24	0.0805 ± 7.1576	0.0632 ± 4.7718	0.0546 ± 3.5788	0.0494 ± 2.8631	0.039 ± 1.4315	p3_125	p3_125	p3_125	p3_125	p3_125

The optimal mean and its standard deviation (using 10 replicates) for the Area change (%) is given.
The right part of the table displays the filter width, which gives optimal performance.

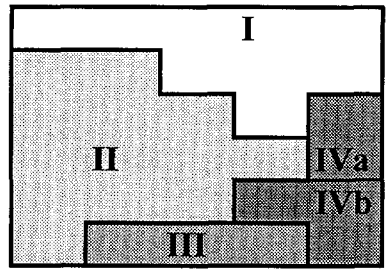

Fig. 7. SNR improvement for optimal MSEp. The following areas may be distinguished: I: where WD MSE is significantly smaller than that of the SG filters and at the same time WD SNR improvement is significantly higher than the SNR improvement of all other filters; II: where WD MSE is the same as all the other filters and at the same time WD SNR improvement is significantly higher than the SNR improvement of all other filters; III: where both MSE and SNR improvement is the same for all filters; IVa: where WD MSE is significantly larger than that of the SG filters and WD SNR improvement is smaller than that of the SG; IVb: where WD MSE is the same as all the other filters and at the same time WD SNR improvement is smaller than that of the other filters.

The highest SNR improvement was obtained by WD followed by PF and FF. WD gave the highest SNR improvement for $\sigma \geq 4$ and also when the SNR was 10. For narrow peaks ($\sigma = 1,2$) with one exception, when the SNR was 10, all filter performed equally well. With respect to broad peaks and high SNR ($\sigma = 16$ and 24, SNR = 4–10) PF gave significantly higher SNR improvements than FF. But it has to be noticed, as mentioned above, that maximizing the SNR improvement is not always desirable because of the distortion of the signal, therefore, we also compared the SNR improvement for filters minimizing the MSE_p (see Fig. 7).

For $\sigma = 1$ there is a clear hierarchy of filters producing significantly better SNR improvements: WD > FF > PF, except of SNR = 2 (Fig. 7, I). Apart from that, with the only two special cases:

1. SNR = 10 and $\sigma \geq 4$, where PD and FF are significantly better than WD (Fig. 7, IV), and
2. $\sigma = 24$ and SNR = 3–5, where all filters perform equally well (Fig. 7, III),

WD is significantly better than the other filters (Fig. 7, II).

Thus when the filter capabilities of minimizing MSE_p and maximizing SNR improvement are considered simultaneously, one sees that WD either gives superior performance with respect to both criteria or a better SNR while producing equally good MSE_p. There is no gain in very broad peaks and even a drawback for high SNR, where the SNR improvement is less than the one obtained by other filter types.

4.5. MSE_{6p}, MSE_{6wp} and SNR_6

Concerning the MSE_{6p} the filters can be ordered as follows WD < PF (Table 5, top). For the very low SNR (2 or 3) the differences are statistically not significant anymore. Among the optimal PF only filters of order $M = 1$ with very small filter widths perform best. There is a slight decrease of filter widths with increase of SNR. The already mentioned inverse relation for the optimal f_c of FF holds here too. This trends is reflected in an increase of the optimal level of the WD, where Daubechies 6 gives the best results.

Regarding the optimization of the MSE_{6wp} all filters perform nearly the same and there is a minor advantage of using WD over PF (Table 5, middle). The weighting of the MSE_p, which balances the influence of the different peak shapes (i.e., reduces the influence of the narrower peaks), results in larger filter widths (ca. 2 times) for PF, lower f_c (ca. $1/2$) for FF and a one level lower WD. All the changes in the filter parameters by a factor of two (remember 1 level is a dilation by 2) show the close relations between these parameters.

In optimizing the SNR_6 WD is clearly superior to the other filters (Table 5, bottom). The highest SNR_6 for PF is obtained by a 3rd order filter of maximal available length, which smooths out the narrow peaks. FF produces the highest SNR_6 using the same $f_c = 14$ for all SNR, while Daubechies 4 or 6 at level 3 or 4 give best wavelet denoising performance.

5. Conclusion

The selection of the proper measure and therefore of the optimal filter and filter parameters is determined by the specific task, whether we are interested in optimal reconstruction or the increase of the detection limit or the best peak area preservation, etc. of the experimental data. This works shows that among the various figures of merit used to quantitatively compare the performance of different filters only a set of criteria, namely MSE_p, MSE_{6p}, SNR and SNR_6 are relevant to analytical data. The peak area preser-

Table 5
The best results for MSE_{6p} (top), MSE_{6wp} (middle) and SNR_6 (bottom) obtained by different filters for different SNR

SNR	2	3	4	5	10	2	3	4	5	10
NSE_{6p}										
PF	0.0504 ± 0.0377	0.0339 ± 0.0292	0.0266 ± 0.0227	0.0229 ± 0.0229	0.0179 ± 0.0228	p1_9	p1_7	p1_5	p1_5	p3_9
FF	0.0648 ± 0.0255	0.0335 ± 0.0117	0.0245 ± 0.0085	0.0180 ± 0.0068	0.0054 ± 0.0019	f142	f154	f198	f218	f270
WD	0.0428 ± 0.0252	0.0276 ± 0.0236	0.0173 ± 0.0074	0.0125 ± 0.0069	0.0049 ± 0.0057	d6_7	d6_7	d6_8	d6_8	s6_8
MSE_{6wp}										
PF	1.8208 ± 0.9511	1.1724 ± 0.5293	0.8636 ± 0.4009	0.6937 ± 0.3231	0.3787 ± 0.1478	p1_17	p1_13	p1_9	p1_7	p1_5
FF	1.8505 ± 0.9076	1.1970 ± 0.7095	0.8219 ± 0.4634	0.6305 ± 0.3682	0.2521 ± 0.1695	f30	f54	f66	f86	f138
WD	1.5715 ± 1.2349	0.9687 ± 0.5135	0.7331 ± 0.4999	0.5329 ± 0.3098	0.2145 ± 0.1535	d6_6	d6_6	d6_7	d6_7	s6_8
SNR_6										
PF	4.5798 ± 3.6258	4.2764 ± 3.6037	4.1452 ± 3.6180	4.0767 ± 3.6304	3.9634 ± 3.6517	p3_149	p3_149	p3_149	p3_149	p3_149
FF	3.8760 ± 2.7958	3.6336 ± 2.7580	3.4710 ± 2.6784	3.3270 ± 2.5847	2.9728 ± 2.2112	f14	f14	f14	f14	f18
WD	5.3799 ± 4.4324	6.5040 ± 5.2311	7.6317 ± 5.5532	8.5808 ± 5.5769	11.752 ± 4.5991	d4_3	d4_3	d4_3	d4_3	d4_3

The optimal mean and its standard deviation (using 10 replicates) for the individual measures is given.
The right part of the table displays the optimal filter parameters.

vation was not very suggestive for parameter selection due to the sensitivity to noise. Since the peak represents only a small fraction of the whole signal, the MSE_s is mostly dominated by the baseline. Maximizing the SNR results in distortion of the signal and thus the results can be used only for qualitative purposes.

With respect to MSE_p wavelet denoising gives superior results to all other filters for narrow and high SNR. In the other cases the optimized filters give equally good performance, while WD shows a higher SNR improvement than the other filters. MSE_{6p} indicates, how to get a filter which is optimal to a signal that contains peaks of different widths. Daubechies 6 gives the best overall performance in this respect.

Maximizing the SNR improvement can be achieved by using Daubechies 4 wavelets at level 3, especially when the signal contains various peak widths, which was shown by the SNR_6. Only in the case of broad peaks and high SNR polynomial filters show better results than WD.

Other wavelet bases may turn out to be more suitable for specific peak widths and SNR. Although the wavelets known thus far are numerous, new wavelets can be constructed to suit specific purposes.

Since wavelets give promising results for the reconstruction of Gaussian peaks in the presence of white noise, we have started investigating more suitable peak models (like exponential modified Gaussian function and Lorentzian function) as the next step in the application of wavelet denoising to experimental data, such as chromatograms and Raman spectra.

Acknowledgements

This work was supported by the Austrian Scientific Research Council (projects S6205 and S5902). The authors would like to thank Prof. Donoho for providing the TeachWave toolbox to the scientific community and the reviewers for their useful comments.

References

[1] K. Cammann, Fresenius' Z. Anal. Chem. 343 (1992) 812.
[2] O. Lee, A.P. Wade and G.A. Dumont, Anal. Chem. 66 (1994) 4507.
[3] M.U.A. Bromba and H. Ziegler, Anal. Chem. 56 (1988) 2052.
[4] A. Felinger, T.L. Pap and J. Inczedy, Anal. Chim. Acta 248 (1991) 441.
[5] R.J. Larivee and S.D. Brown, Anal. Chem. 64 (1992) 2057.
[6] A.M. Rzhevskii and P.P. Mardilovich, Appl. Spectrosc. 48 (1994) 13.
[7] C.G. Enke and T.A. Nieman, Anal. Chem. 48 (1976) 705A.
[8] L.S. Greek, H.G. Schulze, M.W. Blades, B.B. Gorzala and R.F.B. Turner, Appl. Spectr. 49 (1995) 425.
[9] B. van den Bogaert et al., Anal. Chim. Acta 274 (1993) 71.
[10] Y. Meyer, Analysis at Urbana I, Wavelets and Operators, London Math. Soc. Lecture Notes Ser. (Cambridge University Press, Cambrdige, 1989).
[11] I. Daubechies, Ten Lectures on Wavelets (SIAM, Philadelphia, PA, 1992).
[12] S.G. Mallat, IEEE Trans. PAMI 11 (7) (1989) 674.
[13] O. Rioul and M. Vetterli, IEEE SP Mag. 8 (1991) 14.
[14] J.-L. Starck and A. Bijaoui, Signal Processing 35 (1994) 195.
[15] S.G. Mallat and W.L. Hwang, IEEE Trans. Info. Theory 38 (1992) 617.
[16] J.N. Bradley and C.M. Brislawn, SPIE Proc. 1961 (1993).
[17] M. Bos and E. Hoogendam, Anal. Chim. Acta 267 (1992) 73.
[18] P.B. Stark, M.M. Herron and A. Matteson, Appl. Spectrosc. 47 (1993) 1820.
[19] D.N.S. Perman and H.J. Teitelbaum, J. Phys. Chem. 97 (1993) 12670.
[20] M. Bos and J.A.M. Vrielink, CILS 23 (1994) 115.
[21] S.G. Nikolov, H. Hutter and M. Grasserbauer, Chemom. Intell. Lab. Syst. 34 (1996) 263
[22] D.L. Donoho, IEEE Trans. Info Theor. 41 (1995) 613.
[23] D.L. Donoho, Proc. of Symp. in Appl. Math. (1993) 47.
[24] J.P. Foley and J.G. Dorsey, J. Chromatogr. Sci. 22 (1984) 40-46.
[25] D.A. Skoog and J.J. Leary, Principals of Instrumental Analysis (Harcourt, Brace and Jovanovich, Philadelphia, 1989) Ch. 4.
[26] J.D. Ingle and S.R. Crouch, Spectrochemical Analysis (Prentice Hall, Englewoods Cliffs, NJ, 1988) Ch. 5.
[27] H. Press et al., Numerical Recipes in C, The Art of Scientific Computing, 2nd Ed. (Cambridge University Press, Cambridge, 1992).
[28] A. Savitzky and M.J.E. Golay, Anal. Chem. 36 (1964) 1627.
[29] S.E. Bialkowski, Anal. Chem. 61 (1989) 1308.
[30] J. Steinier, Y. Termonia and J. Deltour, Comments on smoothing and differentiation of data by simpliefied least square procedure, Anal. Chem. 44 (1972) 1906.
[31] A.V. Oppenheimer and R.W. Schäfer, Discrete-Time Signal Processing (Prentice Hall, Englewood Cliffs, NJ, 1989).
[32] J.O. Strömberg, Proc. Conf. in Harmonic Analysis in Honor of Antoni Zygmund, Vol. II (Wadworth Mathematical Series, 1983).
[33] A. Cohen, I. Daubechies, B. Jawerth and P. Vial, Comptes Rendus Acad. Sci. Paris (A) 316 (1992) 417.
[34] J.M. Laeven and H.C. Smit, J. Chromatogr. 176 (1985) 77.