

Adaptive noise cancelling of multichannel magnetic resonance sounding signals

E. Dalgaard,¹ E. Auken¹ and J. J. Larsen²

¹HydroGeophysics Group, Department of Geoscience, Aarhus University, Denmark. E-mail: esben.dalgaard@geo.au.dk

²Aarhus School of Engineering, Aarhus University, Denmark

Accepted 2012 July 19. Received 2012 June 12; in original form 2012 March 13

SUMMARY

Adaptive noise cancelling of multichannel magnetic resonance sounding (MRS) signals is investigated. An analysis of the noise sources affecting MRS signals show that the applicability of adaptive noise cancelling is primarily limited to cancel powerline harmonics. The problems of handling spikes in MRS signals are discussed and an efficient algorithm for spike detection is presented. The optimum parameters for multichannel adaptive noise cancelling are identified through simulations with synthetic signals added to noise-only recordings from an MRS instrument. We discuss the design and the efficiency of different stacking methods. The results from multichannel adaptive noise cancelling are compared to time-domain multichannel Wiener filtering. Our results show that within the experimental uncertainty the two methods give identical results.

Key words: Time-series analysis; Fourier analysis; Hydrogeophysics Magnetic field; Instrumental noise.

INTRODUCTION

Since the late 1980s magnetic resonance sounding (MRS) has steadily evolved towards being a competitive technique for non-invasive investigations of groundwater resources. In particular, the method allows for a direct determination of the water content of the subsurface (Legchenko & Valla 2002). The main obstacle currently limiting a more widespread use of the method is the very low signal-to-noise ratio of MRS signals recorded in proximity to anthropogenic installations.

The first generations of MRS instruments were single channel instruments. With single channel instruments, both magnetic resonance excitation and signal recording is done with a single loop and various forms of filtering can be used to suppress noise, in particular powerline harmonics (Legchenko & Valla 2003).

Recently, a new generation of multichannel MRS instruments with multiple loops have been constructed (Walsh 2008; Dlugosch *et al.* 2011). The primary loop is still used for magnetic resonance excitation and signal recording. In addition, a number of reference loops, physically displaced from the primary loop, measure only noise. Parts of the noise recorded by the reference loops will be correlated with the noise in the primary loop. With proper signal processing, the noise in the reference coils can be filtered into a replica of the noise in the primary coil and when the replica is subtracted from the primary loop record, the desired MRS signal remains.

One method of performing the signal processing is adaptive noise cancelling (Farhang-Boroujeny 1998). The basic set-up for a two-channel adaptive noise cancelling system is shown in Fig. 1. The primary sensor measure the signal of interest, $s(k)$, corrupted by

additive and uncorrelated noise, $n_1(k)$. A reference sensor, physically displaced from the first sensor, measures only noise, $n_2(k)$. If the two noise signals are correlated, an adaptive algorithm can adjust the filter to make the output $y(k)$ as similar to $n_1(k)$ as possible by minimizing the error signal, $e(k)$. Under optimum circumstances the error signal is then equal to the original signal, $s(k)$.

Adaptive noise cancelling can be applied using different algorithms and set-ups but it is only one particular method of performing signal processing of multichannel MRS signals. Another method is multichannel Wiener filtering but it not yet clear which technique performs best under which circumstances. The first purpose of this paper is to discuss the optimization of adaptive signal processing algorithms and compare the results with multichannel Wiener filtering. MRS signals are often corrupted by very intense noise spikes from electrical discharges. The second purpose of this paper is to give an analysis of this noise source.

All the signals in this paper have been recorded with NUMIS Poly (IRIS Instruments, Orleans Cedex 2, France). This is a multichannel MRS instrument with one primary and three reference data channels. The data are sampled with a frequency of 19200 Hz. All channels are equipped with tunable analogue bandpass filters with a bandwidth of ~ 150 Hz and sampled by 16 bit analogue to digital converters.

ANALYSIS OF NOISE SOURCES

MRS measurements are disturbed by a number of different noise sources. The two most important are powerline harmonics and discharges from both natural and anthropogenic sources, that is, thunderstorms and electrical installations such as motors and fences.

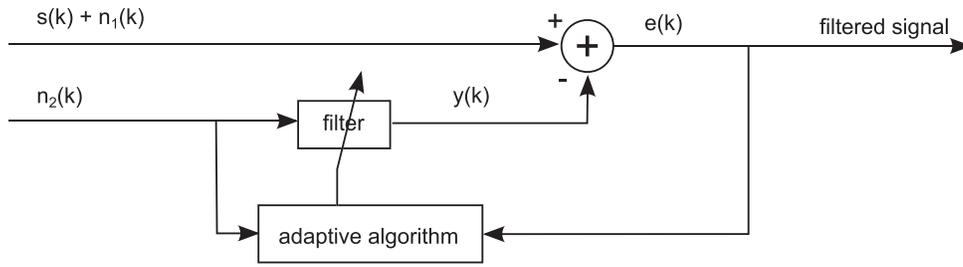


Figure 1. Schematic overview of adaptive noise cancelling with a two-channel set-up.

Other noise sources include low-level atmospheric noise and noise in the MRS receiver electronics. For multichannel MRS measurements the noise sources must be classified according to whether the noise in the primary loop and the reference loops is correlated or not, as this determines the primary way of attenuating the noise.

Short electrical discharges gives rise to an impulsive excitation of the bandpass filters with a subsequent near-exponential decay known as spikes. An example of spikes is shown in Fig. 2. It could be anticipated that spikes are linearly correlated between the loops, however, our experiments have shown that the situation is more complex than that. Fig. 2 shows an example of a noise-only, that is without excitation, simultaneous record of the primary loop and the three reference loops. The spike at 115 ms (labelled ‘B’) is found in all four channels whereas the spike at 20 ms (labelled ‘A’) is present in the primary loop and reference channels 2 and 3 but completely absent in reference channel 1.

One possible explanation for this phenomenon is that two different discharge mechanisms are responsible for the two spikes. The ‘B’ spike most likely originate from a discharge in the far-field as it has been recorded by all channels, whereas the ‘A’ spike most likely originate from a discharge in the near-field as it is only recorded by a subset of the possible channels. From a signal processing viewpoint this means that the transfer functions relating the observed signals in the individual channels depend on the particular noise event. The assumption that the transfer functions are identical for all noise sources is therefore not true. When an adaptive filter is tracking the transfer functions, large filtering errors will occur when different types of noise events occur. Thus transfer function based noise cancelling can only be performed with stationary or near-stationary noise conditions.

The observation in Fig. 2 implies that spikes cannot always be adaptively cancelled. For instance, if a spike is only found in a reference channel, the adaptive algorithm will remove a non-existent spike and thereby actually induce a spike in the primary loop signal. Similarly the large filtering error will disturb the filter coefficients and, depending on the convergence speed of the adaptive algorithm, the disturbance and hence the error cancelling will be diminished even after the spike has decayed. This is because of the non-stationary behaviour of spike events. It should also be noted that large spikes can even saturate the analogue-to-digital converters making adaptive noise cancellation useless. The conclusion of these observations of spike behaviour is that whenever a spike is found in any of the channels, adaptive noise cancelling is potentially unreliable.

The different noise sources can also be interpreted in the frequency domain. Figs 3(a) and (c) show an example of the time-series and spectrum of a noise-only record, with a large spike present in the time domain. In this particular case, the spike contains approximately 85 per cent of the energy of the signal. Because of the large impulsive excitation of the bandpass filter, the spectrum is mainly

dominated by the frequency response of the filter. In contrast, Figs 3(b) and (d) give an example of a time-series and spectrum of a spike-free noise-only section. In this case the spectrum is dominated by the powerline harmonics at multiples of 50 Hz. The inset shows the same spectrum on a logarithmic scale. Here, the bandpass filtering of wideband noise is clearly visible as the broad feature centred at 2150 Hz.

For adaptive noise cancelling to work efficiently, the primary signal and the reference signal must be highly linearly correlated. The linear correlation between two signals x and y is measured as a function of frequency by the magnitude squared coherence function defined as (Kuo & Morgan 1996):

$$|\gamma_{xy}(f)|^2 = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)}. \quad (1)$$

In this equation, $P_{xy}(f)$ is the complex cross-power spectral density and $P_{xx}(f)$ and $P_{yy}(f)$ are the power spectral densities of the individual signals. The magnitude squared coherence function returns a value of 1 for perfectly linearly correlated signals and a value of 0 for completely uncorrelated signals. A value between 0 and 1 indicates a partly linear correlation.

It can be shown that the maximum possible attenuation of an adaptive noise canceller as a function of frequency is given by (Kuo & Morgan 1996):

$$\text{attenuation}(f) = -10 \log_{10}(1 - |\gamma_{xy}(f)|^2) [\text{dB}]. \quad (2)$$

For example, an attenuation of 20 dB can only be obtained if $|\gamma_{xy}(f)|^2 = 0.99$ or larger.

A plot of the coherence between the primary channel and a reference channel for a noise-only signal without spikes is shown in the upper part of Fig. 4. The lower part of the figure shows the power spectral densities of the signal in the two channels. The figure shows that the 50 Hz harmonics in the two channels are nearly perfectly linearly correlated as the coherence is near unity. The resolution of our coherence measurements is limited by a number of factors; the maximum signal length which is recorded by the NUMIS Poly instrument, frequency drifts and jitter in the instrument and short time variations in the fundamental 50 Hz frequency of the powerline harmonics. Away from the harmonics, the coherence basically vanishes. This implies that the harmonic noise can be efficiently cancelled with adaptive noise cancelling, whereas the broad-band noise is incoherent, that is random and must be diminished by standard averaging of multiple experiments.

An example of the coherence of a cut out section of a noise-only signal record containing a single spike is shown in Fig. 5. In this case the power spectra are dominated by the broad feature of the bandpass filter. The value of the coherence function in the interesting region around 2150 Hz is well below 1. This implies that there is not a simple linear correlation between the primary channel

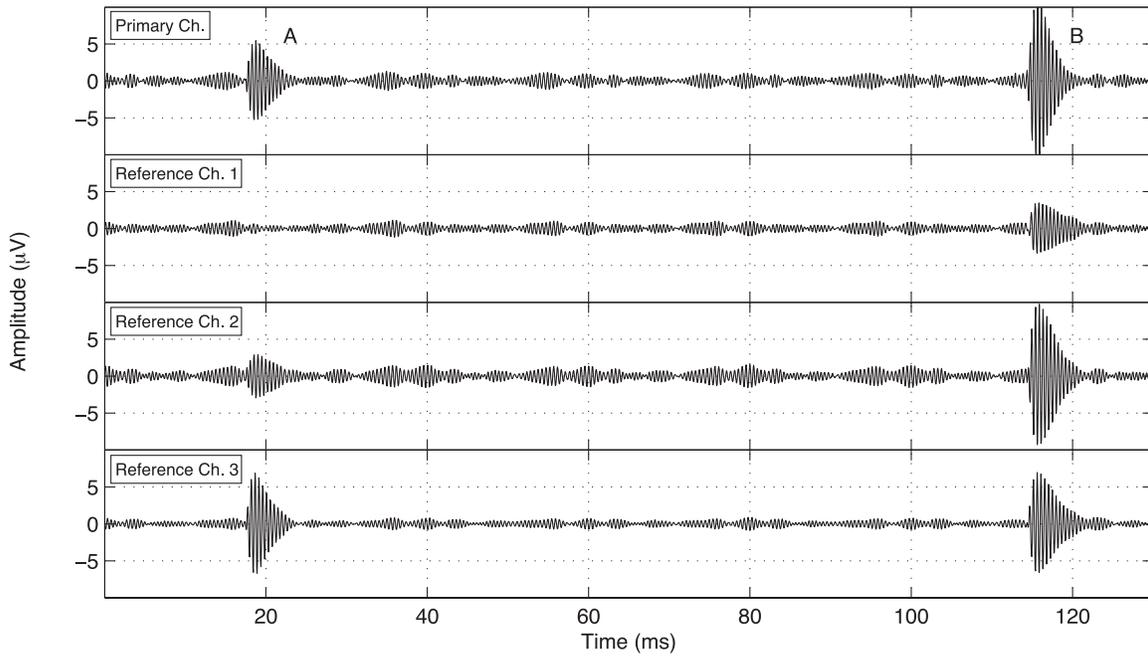


Figure 2. The figure shows the simultaneously measured noise-only signal in the primary channel (top graph) and the three reference channels. All reference channels were 10×10 mm with 10 turns. The spike, B, at 115 ms is present in all four channels whereas the spike, A, at 20 ms is completely absent in one of the reference channels.

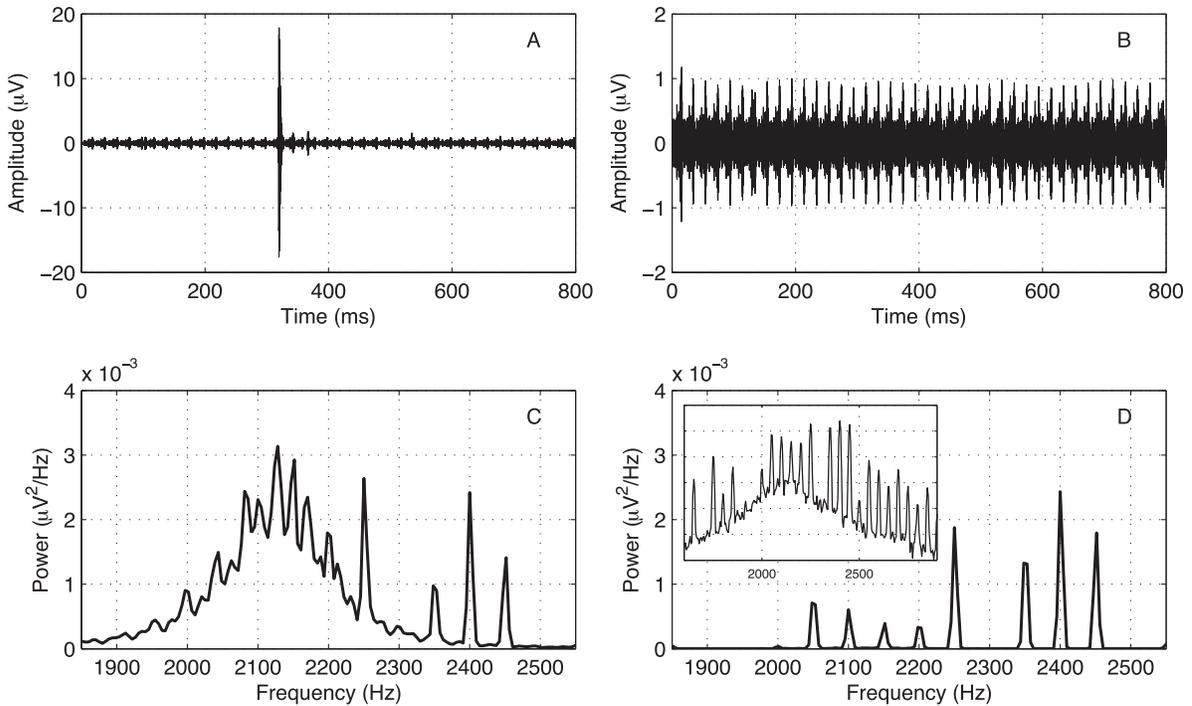


Figure 3. Examples of noise-only recordings and their corresponding spectra. Panel (a) time-series with a single spike. Panel (b) time-series without spike, note the different amplitude scales of panels (a) and (b). Panels (c) and (d) power spectra of panels (a) and (b). The inset in panel (d) shows the same spectrum on a logarithmic scale.

and the reference channel throughout the signal record and it is not feasible to use adaptive noise cancelling to remove the spike. If the spike is removed from the time-series, the coherence plot of the remaining data completely resembles Fig. 4. Similarly, if the coherence of a short section containing only the spike is measured, a high coherence above 0.9 is found in a broad region of several 100 Hz across the centre of the bandpass filter. However, the transfer func-

tion relating the spike in the two channels is different from the transfer function relating the powerline harmonics and this causes the coherence of the entire time-series to fall well below 1 in the region around 2150 Hz.

Based on the above observations we can conclude that efficient removal of powerline harmonics is possible through adaptive noise cancelling. However, the spikes in the MRS recordings cannot be

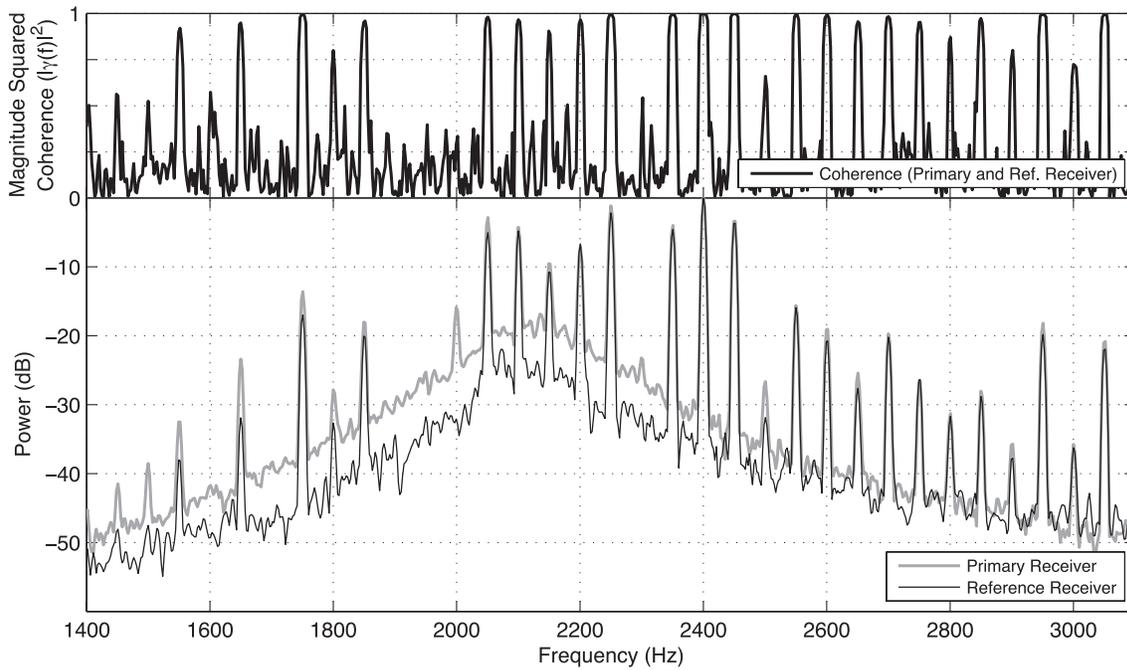


Figure 4. Top graph: Magnitude square coherence function between the primary channel and a reference channel for a noise-only, spike free signal. Lower graph: the corresponding power spectra normalized to the same peak value.

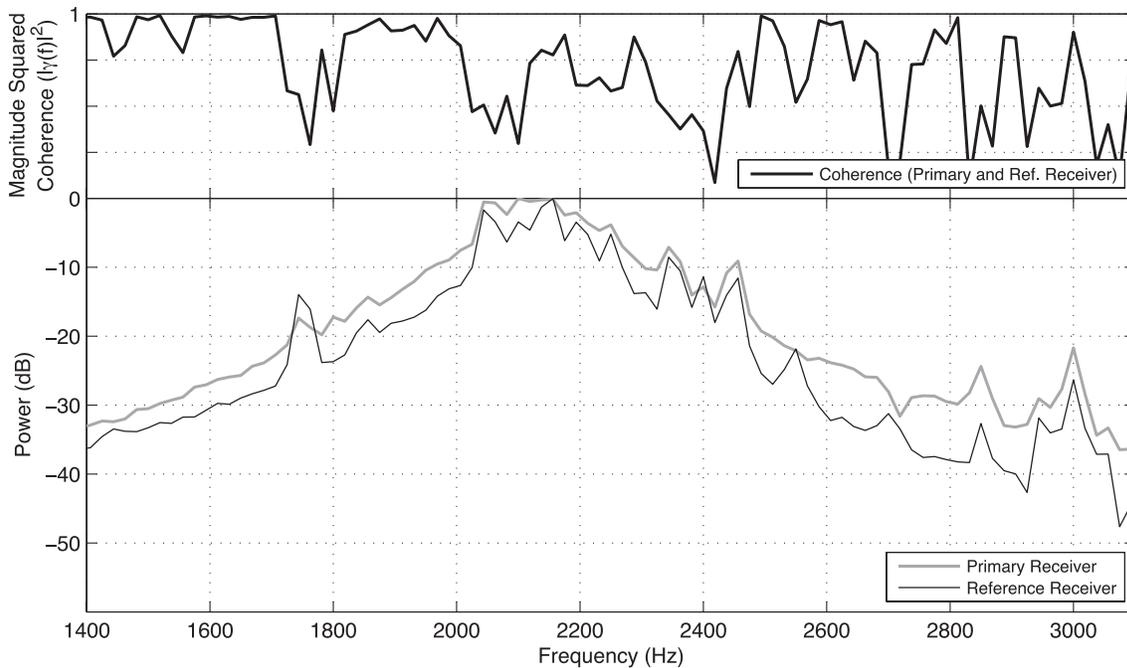


Figure 5. Top graph: Magnitude square coherence function between the primary channel and a reference channel for a noise-only signal containing a single spike. Lower graph: the corresponding power spectra normalized to the same peak value.

expected to be cancelled with the same transfer functions but must be treated differently. Normally, MRS recordings contain only a limited number of spikes therefore a simple and efficient solution of the problem is to discard the sections of the data containing spikes.

On top of the harmonic noise and the spikes we also observe a smaller random noise component, originating in wideband noise and receiver electronics noise. This random component can be suppressed by standard signal averaging.

In the above sections the emphasis has been on the correlation between the primary channel and one of the reference channels. We have also analysed the correlation in between the reference channels and the results are identical: (1) the powerline harmonics are highly correlated between the reference channels. (2) Spikes are also correlated but through different transfer function. (3) The random noise is not highly correlated between the reference channels.

The figures and conclusions are taken from a random site. They are, however, representative for all the different sites where

measurements have been performed. The noise characteristics, that is the ratio between power in the different harmonics or the power in the random noise components obviously differs from site to site but the conclusions remains valid for all sites.

The analysis shows that it is normally necessary to remove spikes before using adaptive noise cancelling. After despiking and noise cancelling the MRS signal is still corrupted by random noise. The random noise can be suppressed by standard averaging of multiple recordings. Finally, the parameters of the MRS signal can be extracted by envelope detection. The entire signal processing workflow for a MRS recording before inversion is then:

- (1) Despiking
- (2) Noise cancelling
- (3) Stacking
- (4) Envelope detection

SEMI-AUTOMATIC DETECTION OF SPIKES

The presence of spikes of different origins in MRS signal records will distort noise cancelling of the powerline harmonics as concluded in the previous paragraph. The spikes in the signals from both the primary and reference receivers must therefore be removed before further noise cancelling. The spikes will contaminate only a part of the signal section, leaving a major part of the section uncontaminated, thus deleting entire sections is not necessary. Instead a local detection and removal of spikes is more appropriate. Because of the huge data amount, a spike detection performed by hand is very time consuming and automatic spike detection would be preferable.

A number of different spike detection algorithms have been proposed by different authors, often lending inspiration from the closely related problem of finding peaks in a signal. The basic principle of

spike detection algorithms is a comparison of some feature of a signal with a predefined or statistically defined threshold. For instance Jiang *et al.* (2011) suggest the Romanovsky criteria for detection of spikes containing only a few samples of high amplitude. Their algorithm is applied in the stacking process after noise cancelling. This method is therefore not applicable in our case where we aim to perform the noise cancelling on despiked data. Strehl (2006) suggested the use of wavelets for identifying spikes, but further work is needed in this direction before a robust method can be devised. We have developed an efficient and simple procedure for automatic spike detection.

Inspired by the work of Mukhopadhyay & Ray (1998) we have implemented a spike detection algorithm based on the non-linear energy operator (NEO). For a discrete time signal, $x(k)$, the NEO is defined as:

$$\varphi[x(k)] = x^2(k) - x(k-1)x(k+1). \quad (3)$$

The purpose of the NEO is to pre-emphasize spikes before a threshold criterion is applied. Two examples of this are shown in Fig. 6. The left part of the figure shows an example from Risby, a noisy site more or less surrounded by electrical fences and in close proximity to powerline installations. The top graph shows a section of the signal containing one large spike and two smaller spikes, marked by the spike detection algorithm. On the lower graph the corresponding NEO signal is shown. The most interesting feature of the graph is the smallest spike at ~ 100 ms. The ratio between the amplitude of the spike and the amplitude of the base signal is increased from ~ 4 to ~ 8 by the NEO making the spike much more visible in the NEO plot.

The right part of Fig. 6 shows results recorded in Skive, a site almost completely devoid of powerline interference and spikes. The MRS signal is clearly visible in both the time-series and the NEO signal. Again, the application of the NEO results in a significant emphasize of the spikes.

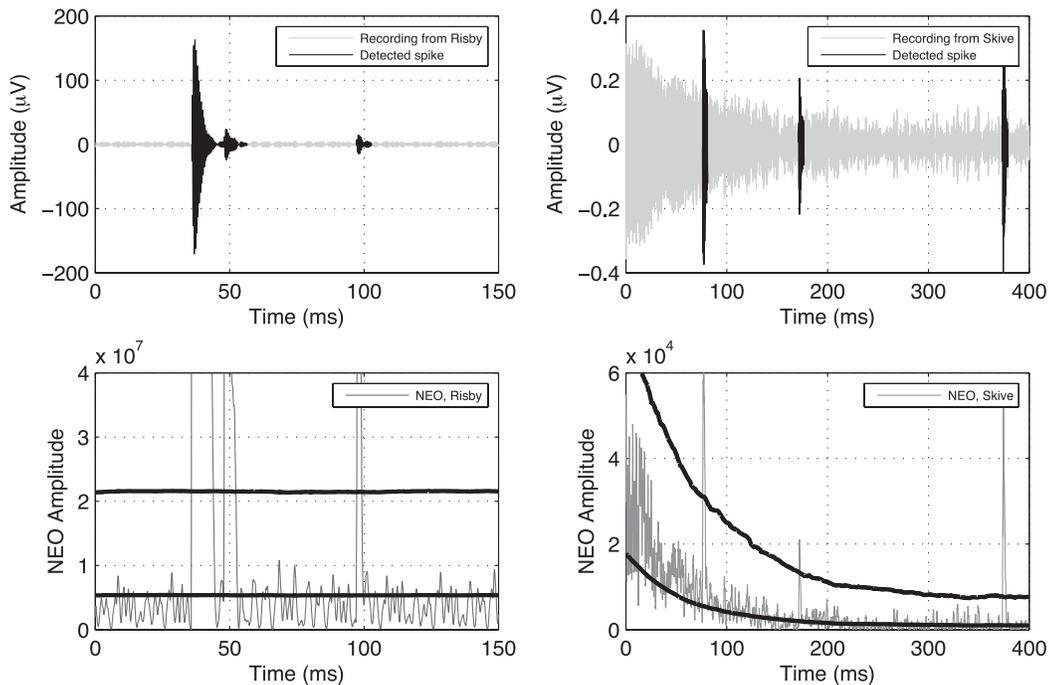


Figure 6. Examples of semi-automatic spike detection. Top graphs: MRS recordings from Risby and Skive. Detected spikes are marked. Note the very different amplitude scales. Lower graphs: corresponding plots of the NEO. The upper threshold for Risby is defined as the median +10 MAD and as the median +20 MAD in the Skive case.

Once the spikes are emphasized by the NEO, an appropriate threshold criteria must be defined. From the right part of Fig. 6 it is clear that a simple threshold criteria is insufficient as it will either mark the signal in the beginning as spike or fail to recognize the small spikes. A criterion based on the variance of the time-series or NEO signal is also sensitive to the presence of MRS signal. Moreover, the variance depends on the number and amplitude of spikes and hence detection of small spikes will be problematic when large spikes are present. To overcome these problems an ensemble-based threshold is used with the threshold determined by the median absolute deviation (MAD). For a set of measurements x_1, \dots, x_n the MAD is defined by (Hoaglin *et al.* 2000):

$$\text{MAD} = \text{median}_i \{ |x_i - \text{median}_j \{x_j\}| \}. \quad (4)$$

j indicating the ensemble of data at a given point, and the median related to i is taken from the just modified ensemble. The main benefit of the MAD is that it is much less sensitive to outliers than the variance and hence provides a very robust measure of the variability of the signal without the influence of spikes.

The workflow of the proposed spike detection algorithm contains three steps and is as follows: (1) For a given moment of the MRS recording, the NEO of all stacks and coils are calculated and appropriately low-pass filtered. For each sample in each coil the MAD is calculated across the ensemble of all stacks. (2) Based on the MAD an upper threshold and a lower threshold is defined for each sample. The thresholds are shown as two black curves on the lower graphs in Fig. 6. (3) Subsequently, the individual time-series from each coil are analysed. When the MAD crosses the upper threshold from below it defines the beginning of a spike. The end of the spike is defined as the instance the MAD crosses the lower threshold from above. The optimum choice of the thresholds depends on the noise characteristics of the specific site and is adjusted accordingly. As an example the upper threshold is defined as the median+10 MAD in the upper part of Fig. 6 and as the median+20 MAD in the lower part of Fig. 6. The bottom right graph shows how the MAD defined threshold automatically adjusts to the presence of the MRS signal. When the samples containing spikes are identified, these are ignored during the estimation of the transfer function and they are not applied in the stacking procedure.

The proposed algorithm has been found to be stable and efficient for all sites measured when used with appropriately user-selected thresholds. The thresholds are adjusted by inspection of the results from a few stacks. The algorithm has been tested on a number of synthetic signals containing either no spikes, a very large number of spikes or saturated spikes and was found to work efficiently in all tested scenarios.

ADAPTIVE NOISE CANCELLING

For optimum performance of an adaptive noise cancelling system it is important to choose a well-suited algorithm and operate it efficiently. The processing of MRS signals is normally carried out offline, therefore, the constraints imposed on a real-time adaptive noise canceller in terms of memory and power consumption, causality and required speed of processing does not apply in this case. The algorithm can be chosen solely based on noise cancelling performance.

A schematic overview of the adaptive MRS multichannel noise cancelling system is shown in Fig. 7. The signals from the three reference coils are sent through three adaptive filters and their combined output is used for noise cancelling.

Choice of adaptive algorithm and convergence properties

Since the invention of adaptive signal processing a large number of different algorithms have been proposed, evaluated and optimized for a number of scenarios. In this work we concentrate on two different type of adaptive algorithms; the normalized least mean square (NLMS) and the recursive least-square (RLS) algorithm, see for example Farhang-Boroujeny (1998) and appendix A. The major benefit of the least mean square type algorithms are their low computational cost and ease of implementation. However, if the input signal to be filtered, that is the MRS reference channel signal, has a very uneven distribution of power across its frequency spectrum the convergence properties of these algorithms can be quite poor. The RLS algorithm is computationally more expensive, but converge much faster independent of the frequency distribution of power. The distribution of power in the spectrum is quantified by the eigenvalue spread, that is, the ratio of the largest and smallest eigenvalue, of the autocorrelation matrix of the signal to be filtered. For MRS signals the typical input spectrum looks like the examples in Figs 3 and 4. The high power of the narrow harmonic signals as compared to the broad feature results in a large eigenvalue spread that often exceeds 10^4 .

To identify the best algorithm and its limitations a number of simulations have been carried out using both real MRS noise recordings and model noise. Three important results have emerged from these simulations. First of all, the adaptation of the filters should preferably not be done during sections including MRS signal. When the adaptive filter seeks to minimize the error signal, $e(k)$ in Fig. 1, the non-linear action of the adaptive filter can create a sharp peak in the filter output exactly at the frequency of the MRS signal. A simulation of this phenomenon is shown in Fig. 8 where the parameters of the simulation were adjusted to enhance the visibility

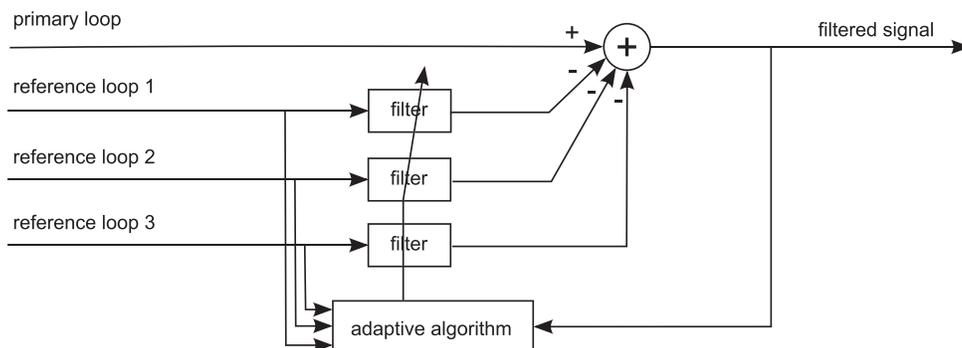


Figure 7. Schematic overview of multichannel adaptive noise cancelling.

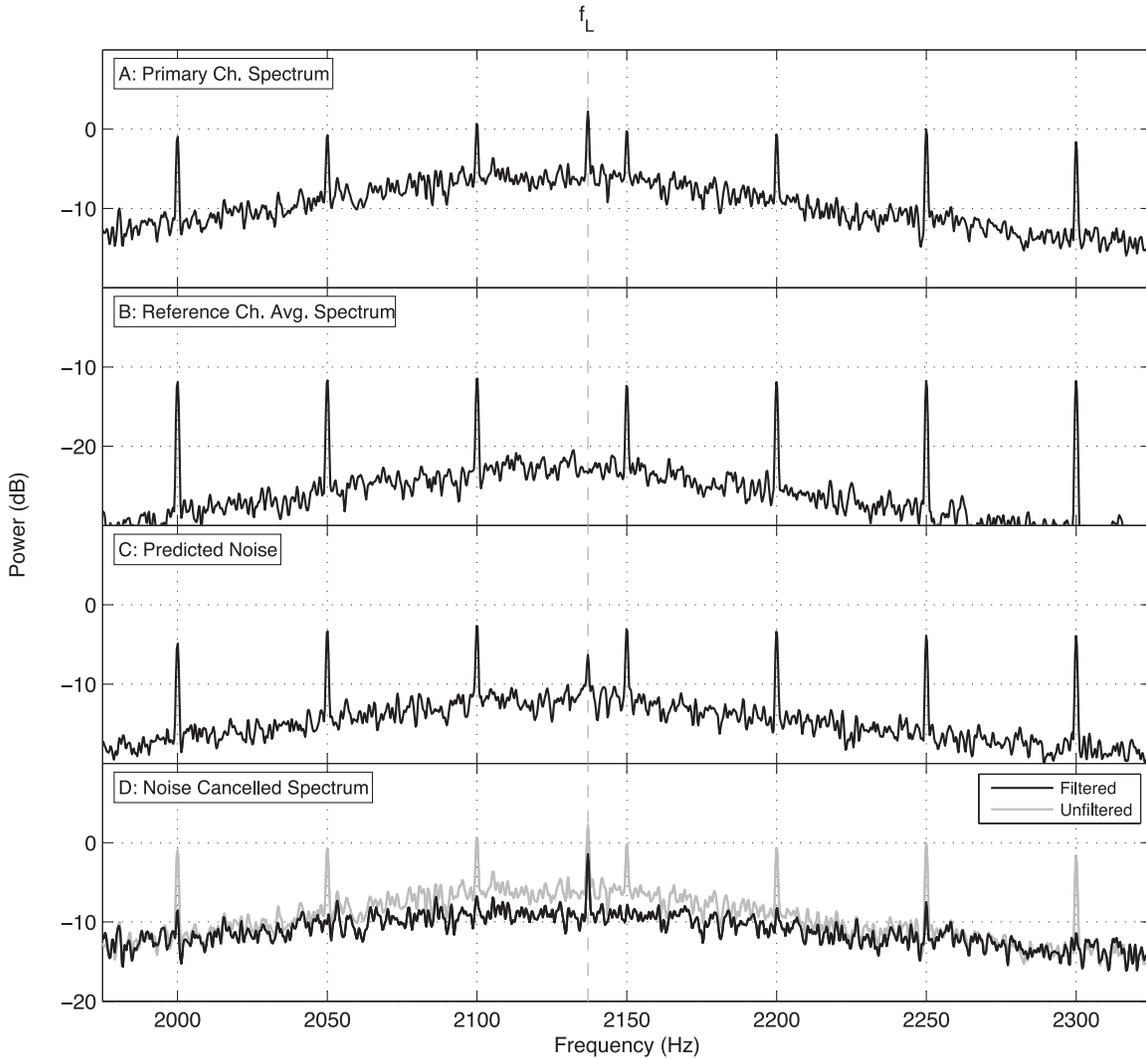


Figure 8. Simulation of noise cancelling with adaptation of filters during the presence of MRS signal. The Larmor frequency, f_L , is marked with a grey line. (a) Spectrum of the signal in the primary channel. (b) Averaged spectrum of the three reference channels. (c) Spectrum of the predicted noise at the primary loop. Note the new artificial spectral feature in the filtered signal at f_L . (d) Spectrum of noise cancelled MRS signal, black and the unfiltered signal, grey.

of the effect. The interpretation of the new peak is that the adaptive filter seeks to cancel both noise and MRS signal. This leads to a reduction of the MRS signal amplitude and a distortion in MRS signal shape. Simulations have shown that the effect can be minimized by adapting the filter very slowly. The only means of completely avoiding this problem is to adapt the filters on noise-only sections of the data and lock the filter coefficients during MRS signal recording.

The second important result is that the convergence of the adaptive filter must be performed very slowly. If the filter is allowed to converge quickly, it can cancel noise very efficiently through fast and continuous adjustments of the filter coefficients. However, if the filter coefficient values are suddenly fixed, as necessary during recordings with MRS signal, the efficiency of the noise cancellation is greatly diminished, as the coefficients are not adjusted for optimum average noise cancelling.

Fig. 9 shows simulations based on noise-only recordings from two different sites. A synthetic MRS signal was added to every other noise-only stack. The adaptive filters were trained on the stacks without synthetic signal and locked during the stacks with synthetic signal. The signal-to-noise ratio after noise cancellation was calcu-

lated as

$$\frac{S}{N} = \frac{\sum \text{syn}(t)^2}{\sum (y(t) - \text{syn}(t))^2}, \quad (5)$$

where $\text{syn}(t)$ is the synthetic signal and $y(t)$ is the filtered signal. The simulation was repeated for different step sizes of the NLMS algorithm. For each site the signal-to-noise measurements have been normalized to the maximum value. The figure shows that for both sites the signal-to-noise ratio peaks for a step-size parameter, β , of approximately 5×10^{-4} to 10^{-3} . If larger values of β is used, the signal-to-noise ratio quickly diminishes, whereas only a slight decrease is found for smaller values of β . The conclusion is that the step-size parameter should be in the range from 5×10^{-4} to 10^{-3} and importantly, the optimum value is only slightly dependent on the specific site.

Similar experiments with the RLS algorithm shows that the exponential forgetting factor λ must be very close to 1 for good signal-to-noise performance. The implication of the large λ value is that the potentially fast convergence of the RLS algorithm is detrimental to the average noise cancelling performance. Our result shows that we get similar noise cancelling performance with the two algorithms.

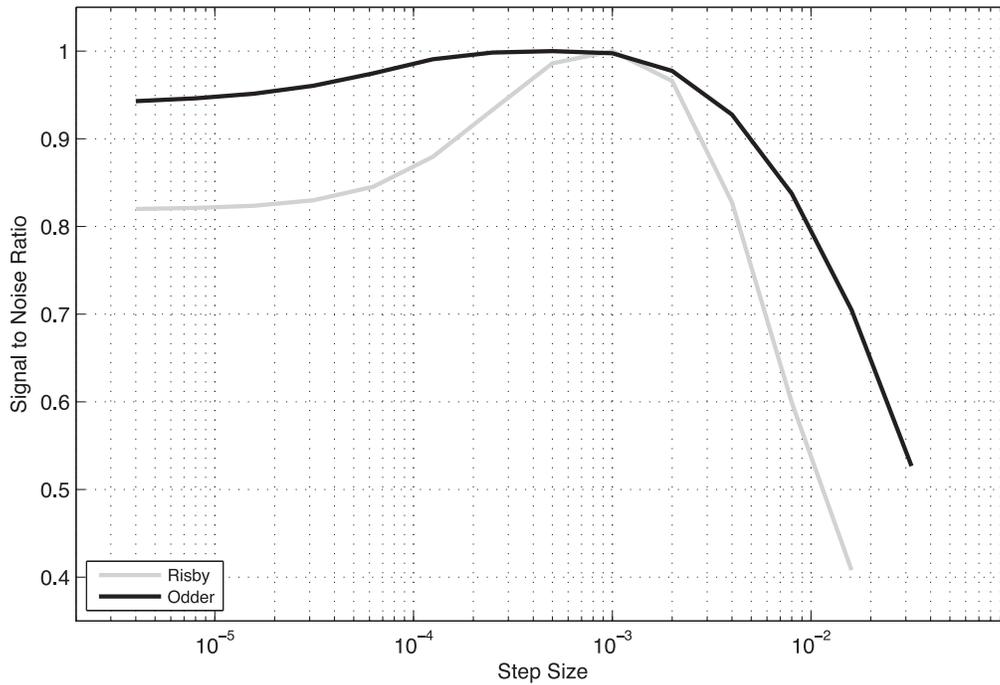


Figure 9. Determination of optimum step size with synthetic signals added to noise-only recordings from Risby, (grey) and Odder (black). The graphs show the normalized signal-to-noise ratio as a function of the step-size parameter for the NLMS algorithm.

The computational overhead of the RLS algorithm is therefore not justified and the NLMS is a well-suited algorithm for noise cancelling.

The final result is that the adaptation of the filters must be turned off during spikes. This problem was stressed above where it was shown that different types of spikes could occur. For different sources of powerline harmonics and spikes, the transfer functions between the reference channels and the primary channel are different. If the filter is allowed to adapt on both types of noise the result will be a distortion of the optimum harmonic noise cancelling whenever a spike occurs.

Choice of filter length and non-causal filtering

The coherence analysis of the different noise sources showed that the adaptive filter will primarily cancel the 50 Hz harmonics in the MRS signal and the adaptive filter must therefore be designed with this knowledge in mind. In a worst-case scenario, the relative amplitudes and phase of each harmonic in the primary and reference channels are completely arbitrary. The transfer function of the adaptive filter must therefore have a frequency resolution, F_{res} , of 50 Hz or better so that the amplitude and phase of each harmonic can be individually controlled. The minimum filter length, number of taps, n_{taps} , in a finite impulse response filter for this resolution can be calculated as

$$n_{\text{taps}} = \frac{F_s}{F_{\text{res}}}, \quad (6)$$

where F_s denotes the sampling frequency. For the NUMIS Poly system with a sampling frequency of 19 200 Hz, at least 384 taps are needed to achieve the desired 50 Hz frequency resolution. However, the worst-case scenario of an arbitrary relationship between amplitude and phase of harmonics in the primary and reference channels is not observed in practice.

Measurements have shown that between the n 'th and $n + 1$ 'th harmonic, the phase of the transfer function normally changes less

than ± 0.2 rad. Likewise, the magnitude of the transfer function measured at the n 'th and $n + 1$ 'th harmonic seldom varies by more than a factor of 2. These observations relax the constraints on the adaptive filter and it is therefore possible to significantly reduce the necessary number of taps in the adaptive filters. However, this conclusion might be different in other sites, for example in a number of Western Europe countries the powerline frequency is 50 Hz but the power for electrical trains has a fundamental frequency of $16^{2/3}$ Hz.

From a theoretical viewpoint it is also desired to minimize the length of the filters as the excess mean squared error for an adaptive filter is proportional to the length of the filter (Kuo & Morgan 1996). The optimum length of the adaptive filters can be found through simulations with synthetic signals added to noise-only recordings. Fig. 10 shows results of such simulations from two different sites, Risby and Odder. The simulations are carried out analogous to the simulations with variable step size except that the step size is fixed at $\beta = 10^{-3}$ and the filter length is adjusted.

For both sites a peak in the signal-to-noise ratio is found. The exact position and width of the peaks are different for the sites. For Risby, the optimum filter length is close to 20 taps whereas for Odder filter lengths in the 10–20 tap range gives almost identical results. For filter lengths longer than the optimum the signal-to-noise ratio decreased as expected. The oscillations found in the graphs are attributed to the high correlation of the input signals because of their large harmonic content.

For other sites investigated but not shown in the figure, similar results are achieved. The optimum filter length changes slightly between sites but remains in the 10–20 tap range. In conclusion, the filters should be long enough to give the necessary frequency resolution but for filter lengths increased beyond the optimum the noise cancelling properties decreases slowly. For a specific site the recovery of a synthetic signal added to noise-only recordings can be used to adjust both step size and filter length for optimum noise cancelling performance.

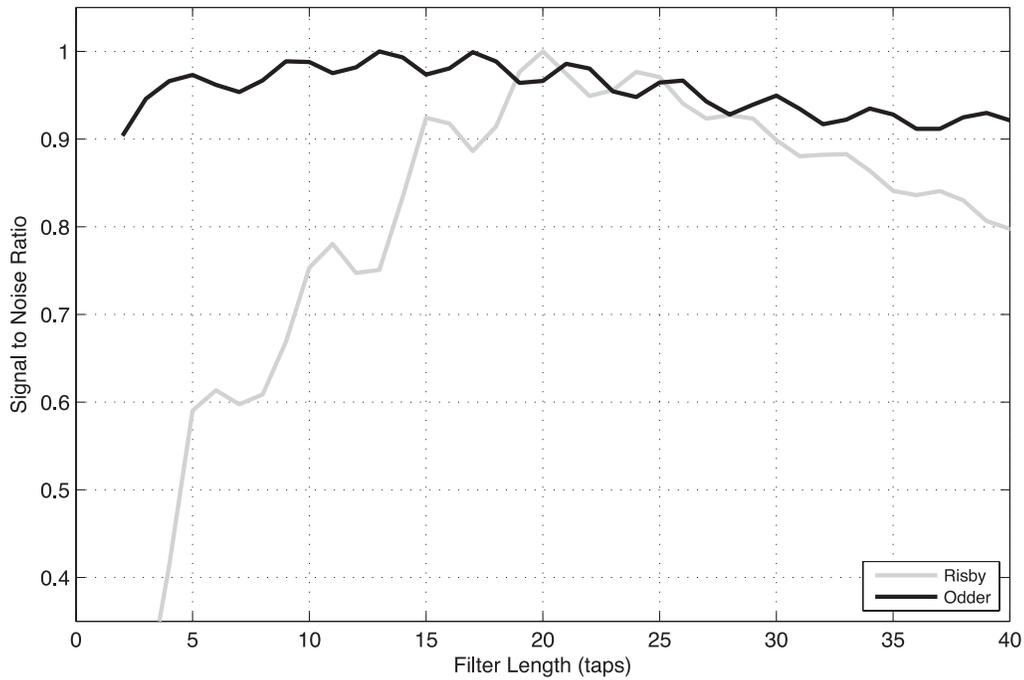


Figure 10. Determination of optimum filter length with synthetic signals added to noise-only recordings from Risby (grey) and Odder (black). The graph shows the normalized signal-to-noise ratio as a function of the filter length for the NLMS algorithm.

The adaptive noise cancelling of MRS signals is done offline. This implies that it is possible to use future values of the reference channel signals for noise cancelling. The analysis of the coherence properties showed that the only noise source that can be efficiently removed with adaptive noise cancelling is the powerline harmonics. Inherently, a harmonic signal is predictable and therefore non-causal filtering cannot improve the noise cancelling. Simulations on noise-only recordings have confirmed this result.

STACKING

After removal of spikes and coherent noise from the MRS recordings a random noise component still remains. The random noise can be suppressed by appropriate stacking of multiple recordings. Stacking can be done by a number of methods, the most simple being averaging to find the mean value. The stacking result can be greatly affected by outliers in the data and it can often be improved by using a method that is insensitive to outliers, for example by using the median as the average, using weighted median methods or using a Romanovsky criterion for discarding outliers (Jiang *et al.* 2011). The quality of the stacked result is quantified by the variance, which should be as small as possible.

A handle on the appropriate stacking method can be found by investigating the probability density function of the MRS recordings after despiking and filtering. Measurements on noise-only data show that the probability density function is nearly Gaussian. Yin *et al.* (1996) showed that for Gaussian distributed white noise the variance of the estimate of the mean value is a factor of $\pi/2$ better with averaging than using the median value.

In Table 1 a comparison of the variance of five different stacking methods employed on noise-only data from two different sites are shown. The five methods are mean value, median value, the mean of the middle 80 and 90 per cent of the data and the mean of the sample with outliers removed by the Romanovsky method. Stationary noise-only recordings were employed so that the ensemble averag-

Table 1. Comparison of the variance of different stacking methods used on data from two different sites, Risby and Odder. The variance is normalized to the mean.

	Risby	Odder
Mean	1	1
Median	1.15	1.15
Median: 80 per cent	0.99	0.99
Median: 90 per cent	0.99	0.98
Romanovsky	0.99	0.99

ing could be extended to neighbouring samples. This avoids the conclusions being limited by the typically small number (30–70) of stacks available for averaging. The data is normalized to stacking based on the mean value. The table shows that stacking based on the median value have the poorest performance as expected. The difference between averaging and median value is not as large as the theoretical prediction because of the only nearly Gaussian distribution of the data. The three stacking methods that are insensitive to outliers are slightly better than simple averaging. They have almost identical performance with small variations between different sites. It can thus be concluded that an outlier insensitive method provides the best stacking result but the result is not sensitive to the exact implementation.

COMPARISON OF ADAPTIVE AND WIENER FILTERING METHODS

In the above paragraphs the optimum parameters for an adaptive noise cancelling algorithm was identified. In this paragraph, this method will be compared with multichannel time-domain Wiener filtering as used by for example Neyer (2010). A full account of the method is given by Treitel (1970) and described in appendix B.

In time-domain Wiener filtering, see Fig. 11 for a schematic overview of the method, the noise-only sections of the data are used to estimate the optimum transfer functions between the reference

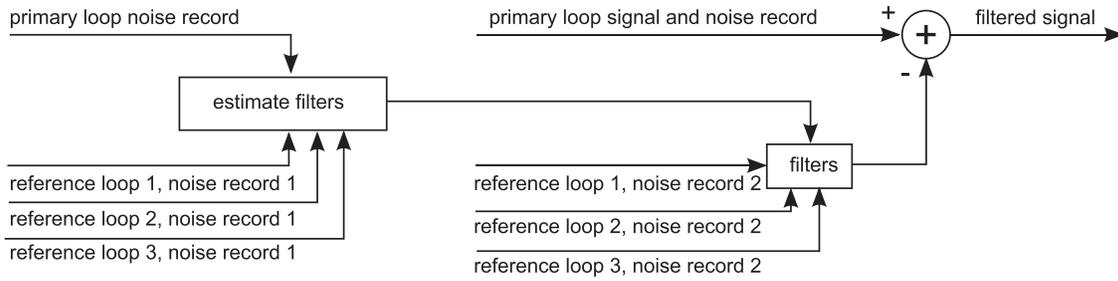


Figure 11. Schematic overview of multichannel Wiener filtering.

channels and the primary channel. The transfer functions are subsequently used to filter the noise recorded in the reference channels during MRS signal recording and subtracted from the primary channel signal. The estimate of the transfer functions is influenced by noise in the measurements and is thus only optimum in the average sense. One of the benefits of the adaptive method is the continuous update of the transfer functions hereby effectively averaging over extended periods of time thus minimizing the influence of noise in the estimate of the transfer functions. It should be noted that under stationary conditions, the adaptive filter will converge to the Wiener filter. The main difference between the two methods in this context is therefore their ability to handle the noise in an environment that is only approximately stationary.

To compare the methods, simulations have been performed where a single exponential MRS signal $s(t) = S_0 \cos(2\pi f_0 t + \varphi) e^{-t/T_2^*}$ is added to noise-only recordings. The noise records are only approximately stationary, for example the power of a particular harmonic can vary by more than a factor of two within minutes. The synthetic MRS signal is then recovered by first despiking the records. Subsequently, harmonic noise is removed with either Wiener or adaptive filtering. Finally, the data are stacked and the MRS signal are retrieved by envelope detection, described in appendix C and standard peak finding and fitting methods using the procedure of Neyer (2010), also described by Mueller-Petke *et al.* (2011). The ability of the different methods to retrieve the MRS signal parameters can then be compared. Representative results from two different sites, Risby and Odder, are shown in Table 2. For comparison the MRS signal has also been retrieved in the case where no filtering has been done. For both filtering methods filter lengths of 20 taps was used. A step size $\beta = 10^{-3}$ was used in the NLMS algorithm.

As expected the results show that the MRS signal is generally better retrieved by using either Wiener filtering or adaptive filtering than no filtering. Within the experimental uncertainty we find that Wiener and adaptive filtering gives equal overall performance with

small and random variations depending on the specific site and noise record used.

In Fig. 12 the spectrum of a single, unprocessed noise-only measurement with synthetic signal added and the corresponding spectrum of the de-spiked, filtered and stacked signal are shown. The figure shows that a noise reduction of approximately 20 dB are obtained across the interesting part of the spectrum and that up to 35 dB of the most intense powerline harmonics are removed by the noise cancelling algorithm. The peak at the Larmor frequency is left undisturbed. A realization of Fig. 12 in the time domain is shown in Fig. 13, in this figure it is possible to observe the characteristics of a MRS signal.

The ability of both Wiener filtering and adaptive filtering is currently limited by the instrument employed: The four analogue-to-digital converters used in the primary loop and the three reference loops are controlled by four independently generated clock signals. This implies that the four channels are not sampled at the exact same frequency and that the timing of the signal acquisition is affected by jitter. The jitter is tantamount to a random delay in the timing between channels and corresponds to a random phase shift of the transfer function. Neither Wiener filtering nor adaptive filtering can react to the random phase shift or the difference in sampling frequencies and the effect is a reduction in the noise cancelling efficiency of both methods. An example of this is seen in Fig. 12 where a number of the powerline harmonics are not as efficiently cancelled as potentially possible.

CONCLUSION

In this paper, we have investigated the applicability of adaptive noise cancelling of multichannel MRS signals. One of the most important results is that an analysis of the noise recorded with the multichannel MRS instrument showed that the relationship between noise, that is spikes, recorded in the primary channel and in the reference

Table 2. Comparison of retrieval of synthetic signal parameters using different methods. Data from two sites, Risby (top) and Odder (bottom) are used.

	Model	Despiking stacking	Stacking	Despiking Wiener stacking	Despiking adaptive stacking
S_0 (nV)	200	192.7 ± 2.2	190.3 ± 2.3	202.4 ± 1.0	203.1 ± 0.8
f_0 (Hz)	2137	2136.91 ± 0.08	2137.07 ± 0.08	2136.99 ± 0.08	2136.91 ± 0.08
φ (rad)	2	2.05	1.96	2.03	1.92
T_2^* (ms)	200	189.9 ± 4.3	201.2 ± 4.7	187.9 ± 1.8	186.3 ± 1.5
S_0 (nV)	200	197.3 ± 1.6	237.0 ± 4.5	198.9 ± 0.5	199.0 ± 0.6
f_0 (Hz)	2137	2137.02 ± 0.08	2137.26 ± 0.08	2137.04 ± 0.08	2137.04 ± 0.08
φ (rad)	2	2.00	1.92	1.99	1.99
T_2^* (ms)	200	205.8 ± 2.3	196.1 ± 5.2	200.9 ± 0.8	200.3 ± 0.8

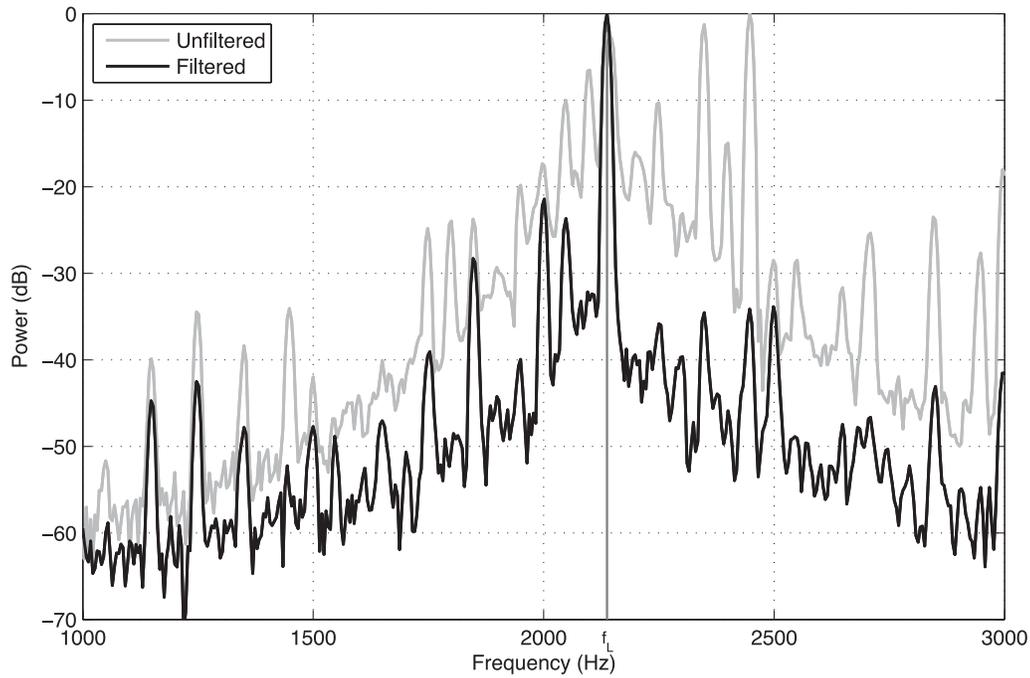


Figure 12. Spectrum of a single unprocessed measurement with synthetic signal added (grey) and the spectrum of the processed signal (black). The Larmor frequency, f_L , is marked with a grey line.

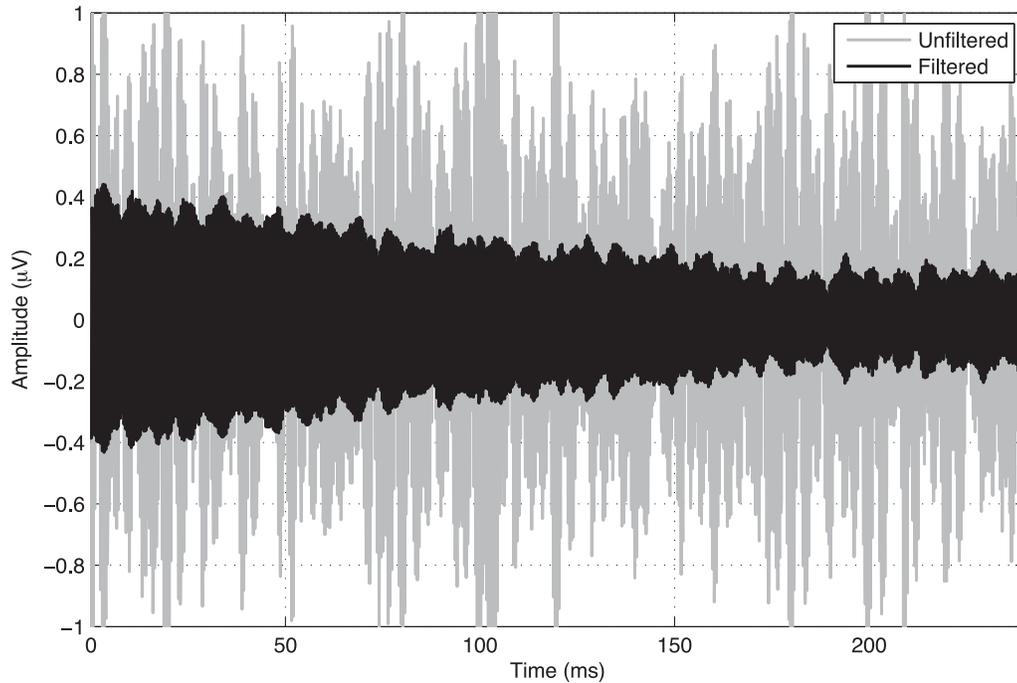


Figure 13. Time-series plot of Fig. 12. The single unprocessed measurement with synthetic signal added is grey and the processed signal is black.

channels typically depends on the particular noise event. This implies that different filters must be used for efficient noise cancellation of different noise events or sections of data containing spikes should be discarded. An efficient algorithm for semi-automatic spike detection was presented.

We have investigated the optimum choices of adaptive filtering algorithm, step size and filter length and devised rules of thumb for practical implementation. We found that adaptation of the filter coefficients should be avoided on time-series including

signal, when this can end up in cancellation of the signal. Adaptation of the filter coefficients should only be done on noise-only sections.

The paper contained a comparison of the noise cancelling properties of Wiener filtering and adaptive noise cancellation. It was found that the two methods provided identical noise cancelling performance within the experimental uncertainty. Our results are currently limited by differences in the sampling frequency of the four channels and jitter in the MRS instrument. These problems affect

both Wiener filtering and adaptive noise cancelling. We anticipate that better noise cancelling results than the ones presented here can be obtained with future generations of the instrument and that a more thorough comparison of the different methods can be made at that stage.

It should be stressed that more elaborate methods are needed to appropriately deal with noise originating from different sources. In particular, it is undesirable to discard noisy sections of the data. This prevents the application of MRS in very noisy areas where entire series of measurements would be discarded.

ACKNOWLEDGMENTS

The authors would like to thank Nicklas S. Nyboe, Ahmad Behroozmand and Lars G. Johansen for fruitful discussions on magnetic resonance sounding and signal processing. We would also like to thank Oliver Ritter (editor), Dr. Mike Mueller-Petke (reviewer) and the anonymous reviewer for their positive and helpful comments.

This work was supported by the Danish Council for Strategic Research funded project NiCA—a project searching to identify robust agricultural areas for which only a limited part of the nitrate leached from the root zone reaches streams. The data collected at Risby was supported by the project Riskpoint, Danish Council for Strategic Research—a project investigating point source contaminations.

We are grateful to the people who helped us collecting the data presented in this paper, August 2010 in Risby, June 2010 near Skive and June 2011 near Odder, which all are locations in Denmark.

REFERENCES

- Dlugosch, R., Mueller-Petke, M., Günther, T., Costabel, S. & Yaramanci, U., 2011. Assessment of the potential of a new generation of surface nuclear magnetic resonance instruments, *Near Surf. Geophys.*, **9**, 89–102.
- Farhang-Boroujeny, B., 1998. *Adaptive Filters: Theory and Applications*, John Wiley & Sons, Inc, New York, NY.
- Hoaglin, D.C., Mosteller, F. & Tukey, J.W., 2000. *Understanding Robust and Exploratory Data Analysis*, Wiley, Hoboken, NJ, 447pp.
- Jiang, C., Lin, J., Duan, Q., Sun, S. & Tian, B., 2011. Statistical stacking and adaptive notch filter to remove high-level electromagnetic noise from MRS measurements, *Near Surf. Geophys.*, **9**, 459–468.
- Kuo, S.M. & Morgan, D.R., 1996. *Active Noise Control Systems/Algorithms and DSP Implementations*, Wiley Interscience, New York, NY.
- Legchenko, A. & Valla, P., 2002. A review of the basic principles for proton magnetic resonance sounding measurements, *J. Appl. Geophys.*, **50**, 3–19.
- Legchenko, A. & Valla, P., 2003. Removal of power-line harmonics from proton magnetic resonance measurements, *J. Appl. Geophys.*, **53**, 103–120.
- Mueller-Petke, M., Dlugosch, R. & Yaramanci, U., 2011. Evaluation of surface nuclear magnetic resonance-estimated subsurface water content, *New J. Phys.*, **13**, 165–185.
- Mukhopadhyay, S. & Ray, G.C., 1998. A new interpretation of nonlinear energy operator and its efficacy in spike detection, *IEEE Trans. Biomed. Eng.*, **45**, 180–187.
- Neyer, F.M., 2010. Processing of full time series, Multichannel surface NMR signals, *MSc thesis*, ETH Zurich.
- Strehl, S., 2006. Development of strategies for improved filtering and fitting of SNMR-signals, *MSc thesis*, Technical University of Berlin.
- Treitel, S., 1970. Principles of digital multichannel filtering, *Geophysics*, **35**, 785–811.
- Walsh, D.O., 2008. Multi-channel surface NMR instrumentation and software for 1D/2D groundwater investigations, *J. Appl. Geophys.*, **66**, 140–150.
- Yin, L., Yang, R., Gabbouj, M. & Neuvo, Y., 1996. Weighted median filters: a tutorial, *IEEE Trans. Circuits Syst.*, **43**, 157–192.

APPENDIX A: THE NLMS ALGORITHM

The filter updating equations for the multichannel NLMS algorithm are given by:

$$w_i(k+1) = w_i(k) + \frac{\beta}{\varepsilon + x_i^T(k)x_i(k)} e(k)x_i(k), \quad (\text{A1})$$

where $i = \{1, 2, 3\}$ represents the three reference coils. β is the step-size parameter ($0 < \beta < 2$) and $x_i(k)$ the signal vector from the specified reference channel. ε is a small number. The output from each filter is given by (Treitel 1970):

$$y_i = w_i^T(k)x_i(k). \quad (\text{A2})$$

The output from the multichannel adaptive filter is given by:

$$e(k) = p(k) - y_1(k) - y_2(k) - y_3(k), \quad (\text{A3})$$

where $p(k)$ denotes the signal in the primary channel. The output is both the noise cancelled MRS signal and the error signal used in the update of the filters.

APPENDIX B: MULTICHANNEL WIENER FILTERING

Multichannel Wiener filtering of a four-channel system consisting of one primary receiver and three reference receivers are almost similar to adaptive filtering, eqs (A2) and (A3). The N filter coefficients for each filter are estimated from a complete noise-only section of data before being used for noise cancelling of the immediately following signal plus noise section. The multichannel Wiener–Hopf equation is :

$$w = \text{ACM}^{-1} \text{CCM}. \quad (\text{B1})$$

ACM is the reference autocorrelation matrix, which is a $3N \times 3N$ Toeplitz block matrix. Each 3×3 block consists of auto- and crosscorrelations of the reference signals. The block corresponding to the k th lag is given by:

$$\text{ACM}(k) = \begin{bmatrix} C_{r1,r1}(k) & C_{r1,r2}(k) & C_{r1,r3}(k) \\ C_{r2,r1}(k) & C_{r2,r2}(k) & C_{r2,r3}(k) \\ C_{r3,r1}(k) & C_{r3,r2}(k) & C_{r3,r3}(k) \end{bmatrix}. \quad (\text{B2})$$

The crosscorrelation matrix CCM is a $3N \times 1$ block vector. Each 3×1 block consists of the crosscorrelation between the primary signal and the three reference signals. The block corresponding to the k th lag is given by:

$$\text{CCM}(k) = \begin{bmatrix} C_{r1,p}(k) \\ C_{r2,p}(k) \\ C_{r3,p}(k) \end{bmatrix}. \quad (\text{B3})$$

The filter coefficients on the left-hand side of eq. (B1) are then given by:

$$w = [w_1(1)w_2(1)w_3(1)w_1(2) \dots w_3(N)]^T. \quad (\text{B4})$$

APPENDIX C: ENVELOPE DETECTION

The desired MRS signal, that is the envelope of the free induction decay is extracted from the noise cancelled and averaged signal with the following four steps.

- (1) Determination of the true Larmor frequency, f_L
- (2) Down-conversion to baseband using f_L
- (3) Determination and correction of the phase offset in the baseband signal
- (4) Lowpass filtering of the phase corrected baseband signal

Step 1 is performed by searching the frequency domain of the signal around the transmitter-generated frequency, f_T , used for MRS excitation. The frequency of the highest amplitude found within a few Hertz from f_T is taken as f_L . The frequency offset is defined as:

$$\Delta f = |f_T - f_L|. \quad (C1)$$

The down-conversion in step 2 is done by:

$$s_{\text{cdc}}(t) = s(t)e^{-i2\pi f_L t}. \quad (C2)$$

Where s is the original signal and s_{cdc} is the down converted signal, which is now complex.

The phase offset of the signal is found as:

$$\varphi = \tan^{-1} \left[\frac{\sum_{t=0}^{t_{\text{end}}} \text{imag}(s_{\text{cdc}}(t))}{\sum_{t=0}^{t_{\text{end}}} \text{real}(s_{\text{cdc}}(t))} \right] \quad (C3)$$

Upon determination of φ the offset correction is done by:

$$s_{\text{cdc,corr}}(t) = s_{\text{cdc}}(t)e^{-i\varphi} \quad (C4)$$

Finally, a lowpass filter is applied for further noise reduction.