Reliability Testing of *in vivo* ^{1}H-MRS-signals and Elimination of Signal Artifacts by Median Filtering

Johannes Slotboom[1] and Dirk van Ormondt[2]
Neuroradiology/DRNN, Inselspital, Bern, CH [1], Applied Physics, TU Delft, NL [2].

Johannes.Slotboom@insel.ch, D.vanOrmondt@tudelft.nl
18th September 2008

***Abstract*—In order to make SV-MRS clinically viable, assessment of spectral *quality* and *reliability* of the data should preferably be handled by the MR-scanner system rather than by medical staff. The reliability of the data is most affected by patient motion during acquisition.**
In this paper, we devise, apply, and test two statistical methods that can be used for automated data reliability- and quality-assessment. Next, we establish that, in order to assess the reliability of SV-MR-data, it is essential to store the data of each acquisition separately, rather than averaging the data of all acquisitions irreversibly prior to storage. We propose a) statistical tests on these separately stored acquisitions that can reveal artifacts, b) application of a special type of order-statistics filtering, namely median filtering, once artifacts have been detected. Furthermore, we develop computer algorithms that provide automated, fully user-independent information on spectral quality and reliability of the acquired spectra. Finally, we conclude that combination of statistical tests with median filtering can provide user-free quality and reliability assessment and improvement of SV-MR-spectra.

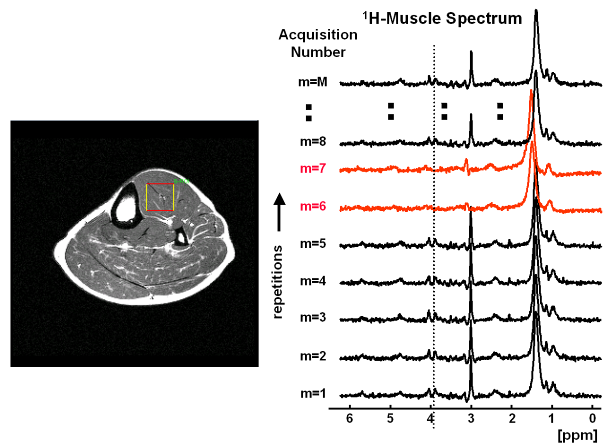

Figure 1. Effect of patient motion - Frequency domain representation of a set of signals recorded during one single measurement. The test person was moving during the signal acquisition resulting in signal artifacts during the acquisition of signal m=6-8.

I. Introduction

In clinical SV-MRS, data *quality* and *reliability* are issues of high importance. Spectral *quality* is determined by line-shape, line-width, presence of spectral artifacts, and signal-to-noise ratio (SNR) [1]. For a first qualitative assessment, an experienced spectroscopist may visually inspect the Fast Fourier Transform (FFT) of acquired data. A quantitative assessment can by obtained from the Cramér-Rao Bounds (CRB's) on parameters estimated from the data [2], [3], [4], [5]. Strictly speaking, CRB's can only be calculated when an *exact* mathematical model of the data is available. CRB's increase with increasing line-width and decreasing SNR.
Another issue, to the best of our knowledge not yet formally addressed in the MRS literature, is that of data *reliability*. Data reliability is affected by time-variance of the patient/scanner system. The largest source of time-variance in such data is patient motion. Note that most often, the data are actually an *average* (sum) of repetitive acquisitions. Fig. 1 shows the spectra of *separately* stored, rather than averaged, acquisitions of data from a tibialis anterior muscle at 3T. The patient moved during acquisitions 6 and 7. As can be observed in Fig. 1, this movement caused changes in frequency, zero-order phase, line-shape, and line-width. Other sources of time-variance are physiologic motion, carrier-frequency drift, radio-frequency (RF) signals of non MR-origin, analogue-to-digital conversion (ADC) problems, and changes in the electronic components of the MR-system. These sources may cause visible artifacts in the FFT (spectrum estimator) of averaged data and affect the quantification. Yet, even for an experienced spectroscopist it is difficult, if not impossible, to judge data reliability by visual inspection only. In fact, data reliability can be truly judged only prior to signal averaging.

II. Theory & Methods

Multiple repetition of an MR pulse sequence and averaging the resulting data (signals) so as to enhance the SNR, is used since the early days of Pulse Response Fourier Transform (PR-FT) NMR [6]. In this well-known approach, averaging of M acquired steady-state signals plus Gaussian noise yields an SNR improvement of a factor of $\sqrt{M}$ compared to that of a single acquisition. A disadvantage of such 'on-the-fly' signal averaging is that testing individual signals for outliers is forfeited. Hence, in order to be able to truly test the textitreliability of the data, all need to be stored individually.

A. Signal model

Let $S_m[n]$ ($1 \leq m \leq M, 1 \leq n \leq N$) represent the m-*th* data acquisition (signal) comprising N samples. Since n is discrete, it is surrounded by square brackets. The signal can be written as a sum of a deterministic part $S_m^{\mathrm{d}}[n]$ which by definition is the *same* for all acquisitions m, and two noise terms:

$$S_m[n] = S_m^{\mathrm{d}}[n] + \delta_m[n] + \varepsilon_m[n] \tag{1}$$

In Eq. 1 the term $\varepsilon_m[n]$ refers to the electric noise picked up by the coil and the noise of the electronic components of the MR-scanner. This noise has a Gaussian probability density function (pdf). Distortions of signals (artifacts) are regarded here as a special type of noise, denoted by $\delta_m[n]$. The latter represents signals of totally different character, like sudden poor water suppression, phase- and carrier-offset, line-shape anomalies, and RF-signals of non-MR origin. Signals 6 and 7, coloured red in Fig. 1, are two examples of signals with a large $\delta_m[n]$component. Their contribution to the pdf is not Gaussian but '*heavily-tailed*', *i.e.*, the tails of the ensuing distribution fall off rather more slowly than that of Gaussian noise [7]. Signals with large $\delta_m[n]$ contributions are *unreliable* because they cause large artifacts that does not average-out in the same fashion as Gaussian noise. If only the term $\varepsilon_m[n]$ is large, the signal has poor SNR but is still deemed reliable.

Finally, we place all data in an $M \times N$ matrix $\mathbf{D}$, with elements $\mathbf{D}_{mn}$, in which each row m, $m = 1, \dots, M$, is filled with a signal $S_m[n], n = 1, \dots, N$ (or its spectrum). Within a column, the deterministic part of the signal, $S_m^{\mathrm{d}}[n]$, is constant by definition. Hence the elements of a column differ only in their noise values $\delta_m[n]$ and $\varepsilon_m[n]$. The next section describes statistical tests on the individual *columns* of $\mathbf{D}$.

B. Estimating Noise Characteristics

A user-independent signal-reliability test must enable to distinguish $\delta_m[n]$-noise from $\varepsilon_m[n]$-noise. A test capable of this is proposed below. It involves the first four *standardized moments about the mean* [8], [9] of the pdf of $\delta_m[n] + \varepsilon_m[n], m = 1, \dots, M$, for each n. In probability theory and statistics, the p-th Standardized Moment $\mathrm{SM}^{(p)}$ of a pdf is defined as

$$\mathrm{SM}^{(p)} \stackrel{\mathrm{def}}{=} \mu_p / \sigma^p \ , \quad p = 1, 2, 3, 4, \dots, \tag{2}$$

where μ_p is the p-th moment about the mean and σ the standard deviation. $\mathrm{SM}^{(1)} = 0$, because the first moment about the mean is zero by definition. $\mathrm{SM}^{(2)} = 1$, because the second moment about the mean equals the variance. Because of symmetry, all odd moments of a Gaussian distribution are zero. Non-Gaussian distributions, like heavily-tailed ones, can have non-zero odd moments.

For use further on in this paper, we introduce the regular statistical names of the standard moments $\mathrm{SM}^{(3)}$ and $\mathrm{SM}^{(4)}$ (see Ref.[8]):

$$\text{skewness} \stackrel{\mathrm{def}}{=} \mathrm{SM}^{(3)} \ , \qquad \text{kurtosis} \stackrel{\mathrm{def}}{=} \mathrm{SM}^{(4)} \ . \tag{3}$$

Skewness and kurtosis are popular entities for characterizing a pdf. Skewness is a measure of asymmetry of a pdf, whereas kurtosis is a measure of its "peakedness". The higher the kurtosis, the more the variance is due to infrequent extreme deviations. For a large population (many samples) with a Gaussian pdf, kurtosis = 3. A heavily-tailed pdf yields a much larger kurtosis than a Gaussian pdf. In this study we use the so-called '<u>excess kurtosis</u>', which equals 'kurtosis - 3'.

Lacking *a priori* knowledge of the pdf of the noise processes in Eq. 1, $\mathrm{SM}^{(p)}$ has to be estimated from the sampled signals. To this end, we applied the following estimators, using the signal (spectrum) matrix notation $\mathbf{D}_{mn}$ introduced above:

1) Sample Mean: The unbiased estimator of the sample mean is (see Ref.[8]) is given by:

$$\mathrm{mean}[n] = \frac{1}{M} \sum_{m=1}^{M} \mathbf{D}_{mn} \tag{4}$$

with variance given by Eq.5.

2) Sample Variance: The unbiased estimator of the sample variance (see Ref.[8]) is given by:

$$\mathrm{variance}[n] = \frac{1}{M-1} \sum_{m=1}^{M} (\mathbf{D}_{mn} - \mathrm{mean}[n])^2 \ , \tag{5}$$

The variance-of-the-sample-variance is given below.

3) Sample Skewness: The, in general, biased estimator of the sample skewness is given by (see Ref.[10]):

$$\mathrm{skewness}[n] = F \cdot \frac{\sqrt{M} \sum_{m=1}^{M} (\mathbf{D}_{mn} - \mathrm{mean}[n])^3}{(\sum_{m=1}^{M} (\mathbf{D}_{mn} - \mathrm{mean}[n])^2)^{3/2}} \ , \tag{6}$$

in which F is defined as:

$$F = \frac{\sqrt{M(M-1)}}{M-2} \ , \tag{7}$$

4) Sample Kurtosis: The, in general, biased estimator of the sample kurtosis is given by (see Ref. [10])

$$\mathrm{kurtosis}[n] = G \cdot \frac{\sum_{m=1}^{M} (\mathbf{D}_{mn} - \mathrm{mean}[n])^4}{\mathrm{variance[n]}^2} - 3 \cdot H \ , \tag{8}$$

in which G is defined as:

$$G = \frac{(M+1)M}{(M-1)(M-2)(M-3)} \ , \tag{9}$$

and H as:

$$H = \frac{(M-1)^2}{(M-2)(M-3)} \ , \tag{10}$$

Since the estimators defined Eqs. 4, 5, 6, and 8 are a *random variables* as well, they have in turn also a variance which should be known, in order to judge their significance. For example: in case of normal distributed signal noise, the estimated skewness and kurtosis should be zero. For any *finite* number of normally distributed sample points

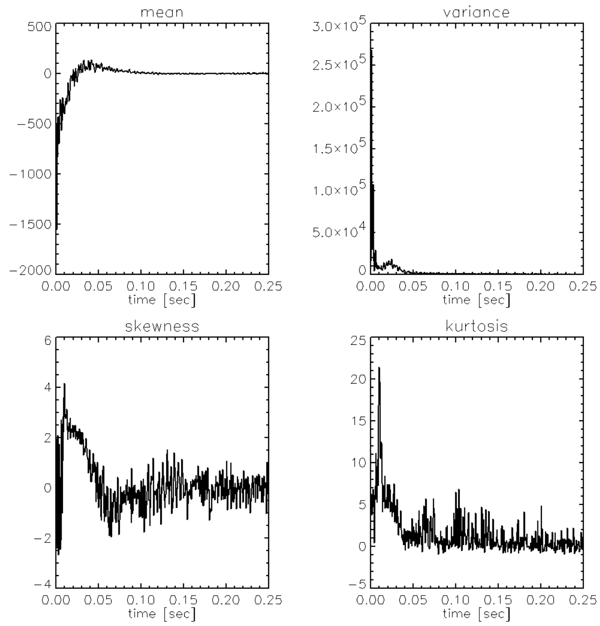

Figure 2. The first four moments of the time domain MRS-signal; test person was asked to make some movements during recording phase. The separately stored signal consists of 32 acquisitions of 8 averages of a complete EXOR-cycle, thus $M = 32$.

however, the estimated values will be in general non-zero. In order to be able to determine whether the estimated values deviate significantly from zero or not, their *variances* of the estimators must be estimated as well.

5) Variance of Sample Mean: The unbiased variance of the sample mean given by Eq.4 is given by Eq. 5.

6) Variance of the Sample Variance: The unbiased estimator of *variance-of-the-sample-variance* parameters denoted by var(variance$[n]$) is given by the following expression (see Ref. [11], [12]):

$$\mathrm{var}(\mathrm{variance}[n]) = a_1\mu_4[n] + a_2\mu_2[n]^2 \ . \tag{11}$$

in which μ_2 is given by:

$$\mu_2[n] = \frac{1}{M}\sum_{m=1}^{M} \mathbf{D}_{mn}{}^2 \ . \tag{12}$$

and μ_4 is given by:

$$\mu_4[n] = \frac{1}{M}\sum_{m=1}^{M} \mathbf{D}_{mn}{}^4 \ . \tag{13}$$

with a_1 given by:

$$a_1 = \frac{M(M-n)(Mn-M-n-1)}{(M-3)(M-2)(M-1)(n-1)n} \ . \tag{14}$$

and a_2 given by:

$$a_2 = \frac{M(M-n)(M^2n-3n-3M^2+6M-3)}{(M-3)(M-2)(M-1)(n-1)n} \ . \tag{15}$$

7) Variance of the Sample Skewness: For normally distributed samples, the variance of the sample skewness parameters var(skewness$[n]$) was given by Cramér in 1946 [13] and can be computed as follows (see Ref. [10]):

$$\mathrm{var(skewness}[n]) = \mathrm{var}(\frac{m_3}{{m_2}^{3/2}}) = \frac{6(M-2)}{(M+1)(M+3)} \ . \tag{16}$$

8) Variance of the Sample Kurtosis: For normally distributed samples, the variance of the sample kurtosis var(kurtosis[n]) is given by (see Ref. [10], [13]):

$$\mathrm{var(kurtosis}[n]) = \mathrm{var}(\frac{m_4}{{m_2}^2}) = \frac{24M(M-2)(M-3)}{(M+1)^2(M+3)(M+5)} \ . \tag{17}$$

Note that the variance of the sample skewness estimator given by Eq.16 and the variance of the sample kurtosis given by 17 above is only valid for normal distributed samples.

C. Signal Reliability Testing

In the previous section we have seen that for each sample point n of the time- or frequency-domain response we can estimate the first four (standardized) moments about the mean. Figure 2 shows the first four moments about the mean of an *in vivo* ^{1}H-MRS dataset ($M = 32$) collected on a 3T MR-scanner from the human cortex of a healthy test person. The test person was asked to move his head occasionally during the recording phase of the MR-signal. In this case, as becomes clear from the first moment about the mean, the *in vivo* MR-signal $S^{\mathrm{d}}_m[n]$ is a superposition of decaying complex sinusoids which decay during the first 150 msec. During this time-interval $\delta_m[n]$ contributes to the signal as becomes clear from analyzing the skewness and kurtosis parameters in the same figure. For acquisition times greater than 150ms we only deal with Gaussian distributed noise originating from the RF-coil and the signal acquisition system. During the first 150 msec this occasional movement causes the skewness and kurtosis parameter to deviate *significantly* from the theoretical value of zero; the variance of the skewness for $M = 32$ is 0.1558 and the variance of the kurtosis for $M = 32$ is equal to 0.4738. By studying Fig.2, intuitively, we can classify the MRS-signal as *unreliable*, if the interval-averaged skewness parameter value systematically exceeds one standard deviation of the skewness (in this case $\sqrt{0.1558} = 0.3947$), and the interval-averaged kurtosis parameter value systematically exceeds one standard deviation of the kurtosis (in this case $\sqrt{0.4738} = 0.6883$). Due to the fact that the skewness- and kurtosis-estimates have positive and negative values, interval-averaging can lead to annihilation, causing an unreliable signal to be classified reliable. Additionally, due to the fact that the signal contribution of $\delta_m[n]$ is only significant in those regions where $S^{\mathrm{d}}_m[n]$ approximately has an SNR>2, since the random signal contribution $\delta_m[n]$ scales with $S^{\mathrm{d}}_m[n]$. Testing should only be done on high SNR part of the signal. Therefore we propose the following two test parameters, $\kappa_{\mathrm{skewness}}$ and

κ_{kurtosis} defined as follows:

$$\kappa_{\text{skewness}} \stackrel{\text{def}}{=} \frac{1}{N_{\text{SNR}}} \sum_{n=1}^{N_{\text{SNR}}} |\text{skewness}[n]| \,, \tag{18}$$

and

$$\kappa_{\text{kurtosis}} \stackrel{\text{def}}{=} \frac{1}{N_{\text{SNR}}} \sum_{n=1}^{N_{\text{SNR}}} |\text{kurtosis}[n]| \,. \tag{19}$$

in which N_{SNR} is the number of points for which the SNR of $S_m[n]$ is larger than a specified value. Analogous κ-values can be defined for the mean, κ_{mean} and the variance parameter, κ_{variance}. Note that the κ-parameters are computed as averages of the *absolute* values of the mean$[n]$, variance$[n]$, skewness$[n]$, and kurtosis$[n]$. If we consider, for simplicity, the values of the skewness$[n]$ and kurtosis$[n]$ to be normally distributed, their corresponding κ-values will have a *folded normal distribution*. In order to judge whether the estimated values κ_{skewness} and κ_{kurtosis} are significant, one needs also the related estimation errors. For this, we need formula's for folded normal distributions: If a stochastic variable X is normally distributed with expectation μ and variance σ, the variable $Y = |X|$ has a folded normal distribution with expectation value $\sigma\sqrt{2/\pi}\exp(-\mu^2/2\sigma^2)+\mu[1-2\Phi(-\mu/\sigma)]$, and variance $\mu^2 + \sigma^2 - (\sigma\sqrt{2/\pi}\exp(-\mu^2/2\sigma^2) + \mu[1 - 2\Phi(-\mu/\sigma)])^2$ in which $\Phi(.)$ is the cumulative distribution function (cdf). For a normal distribution with $\mu = 0$ and $\sigma = 1$ the corresponding folded normal distribution has expectation $\sqrt{(}2/\pi)$ (≈ 0.79) and variance $(1-2/\pi)$ (≈ 0.36). For testing, the individual values of κ_{mean} and κ_{variance} are not of interest, but the following relation between them is. In analogy with the 'coefficient of variation' c_κ (see Ref.[8]), we define

$$c_\kappa \stackrel{\text{def}}{=} 100\frac{\sqrt{\kappa_{\text{variance}}}}{\kappa_{\text{mean}}} \,, \tag{20}$$

which is a dimensionless parameter expressed in units of percent. Due to the fact that Eq.20 has a singularity for $\kappa_{\text{mean}} = 0$, this test parameter is of less general use than κ_{skewness} and κ_{mean}. In order to be able to test for normality, the variance in κ_{skewness} and κ_{mean} must be determined. The experimental unbiased variance of the random variables κ_{skewness} and κ_{kurtosis} can be computed as follows:

$$\text{var}(\kappa_{\text{skewness}}) = K \cdot \sum_{n=N_{\text{SNR}}}^{N} \left(|\text{skewness}[n]| - \kappa^*_{skewness}\right)^2 \,. \tag{21}$$

and

$$\text{var}(\kappa_{\text{kurtosis}}) = K \cdot \sum_{n=N_{\text{SNR}}}^{N} \left(|\text{kurtosis}[n]| - \kappa^*_{kurtosis}\right)^2 \,. \tag{22}$$

in which K is defined as:

$$K = \frac{1}{N - N_{\text{SNR}} - 1} \,. \tag{23}$$

The variances are determined for the part of the MR-signal for which $\delta_m[n]$ is negligible. A MRS measurement will be classified *unreliable* if the intervals $\kappa_{\text{skewness}} \pm \text{st.dev}(\kappa_{\text{skewness}})$ and $\kappa_{\text{kurtosis}} \pm \text{st.dev}(\kappa_{\text{kurtosis}})$ do *not* contain zero. Note that $\kappa^*_{skewness}$ and $\kappa^*_{kurtosis}$ are computed for the interval $[N_{\text{SNR}}, N]$.

D. Non-linear filtering: Median Filtering

In case that signal reliability tests, described above, indicate that the spectrum is unreliable, it is very likely that it contains artifacts. In the literature diverse signal reconstructing filtering techniques are available to eliminate these artifacts. Linear filtering techniques have serious limitations in dealing with signals that have been created or processed by a system exhibiting some degree of non-linearity [14]. Signals poorly handled by linear filters are those with changing levels (*e.g.*, due to patient motion) and corrupting noise that is either heavily-tailed due to outliers, or noise that is signal-dependent. Image processing is the field where non-linear filter techniques have first shown clear superiority over linear filters [15]. The design of non-linear filters can follow many approaches since there is no single underlying theory on non-linear filtering [16]. One non-linear filtering approach that has received considerable attention, and for which much theoretical study has been conducted, is that of so-called rank-order filtering, a method whose filtering effect is obtained by rank-ordering the input data. Rank-order filtering is also known as order-statistics filtering. Much attention has been paid to rank-order filters since the running-median filter was first applied to the smoothing of time series by Tukey in 1974 [17]. Rank-ordering of samples enables the design of filter structures that are (a) robust in environments where the assumed statistics deviate from Gaussian (which is the case in patient movement during data acquisition) and are possibly contaminated with outliers, and (b) track signal discontinuities without introducing transient or blurring artifacts as linear filters do. In this paper we will focus on one special type of order-statistics filter, namely the median filter.

Order-statistics filtering applied to the data matrix $\mathbf{D}$ can be denoted as $\mathbf{R} = \text{Rank-Order}[\mathbf{D}]$. The effect of $\mathbf{R}$ is to rank-order the column elements of $\mathbf{D}$, *i.e.*, for all output columns $r_{1n} \leq r_{2n} \leq \cdots \leq r_{Mn}$ holds. When the number of acquisitions M is odd, the median equals element $m = (M + 1)/2$ of an output column; when M is even it equals the average of output column elements $m = M/2 + 1$ and $m = M/2$.

The median filter does not possess the superposition property and impulse-response analysis is not applicable. In fact, median filtering is non-linear. This condition required development of alternative methods of statistical analysis and characterization of the output. An introduction to these methods is given in [16]. Concretely, the methods enable computation of mean and variance of the median for given random input noise with mean and variance $(\mu_{\text{in}}, \sigma_{\text{in}})$. Important for MRS is the case of noise with Gaussian pdf. For this type of noise, the variance of the median σ_{out} of M samples is given by

[16]

$$\sigma^2_{\text{out}} = \frac{\pi\sigma^2_{\text{in}}}{2M} \ . \tag{24}$$

Eq. 24 implies that, as with conventional signal averaging, the application of median filtering too improves the SNR. For conventional averaging of M input signals the well-known relation

$$\sigma^2_{\text{out}} = \frac{\sigma^2_{\text{in}}}{M} \tag{25}$$

holds, which is consistent with the familiar fact that the SNR is proportional with the square root of the number of averaged acquisitions, *i.e.*, with $\sqrt{M}$. Comparing Eq. 24 with Eq. 25, one can directly see that median filtering performs 2dB less well than averaging. However, median filtering is superior in the presence of outliers. Its slight relative loss in SNR is compensated by superior artifact handling, resulting in more reliable quantification of the spectrum.

III. Results

A. Gaussian noise and Scanner noise

In order to get an impression of the values of κ_{mean}, κ_{variance}, κ_{skewness}, and κ_{kurtosis} estimated from simulated Gaussian noise with $N(0,1)$, using a 32 $\times$ 2048 data matrix $\mathbf{D}$. For this case the value of N_{SNR} was 2048. The same entities were estimated from experimental scanner noise, acquired on a Siemens 3T Trio scanner, also on a signal matrix 32 $\times$ 2048 data matrix $\mathbf{D}$. Figure 3 displays the signals of the first four time domain moments. The values of Table I and II support the conclusion that the experimental scanner noise has a Gaussian probability density function. Also note that the pdf of the kurtosis is clearly not normally distributed (Fig. 3) .

B. Experimental in vivo MRS-signals

For the same test person and same localization whose results were displayed in Fig.2 a measurement with the same acquisition parameters was performed in which the test person was asked to refrain from any motion with the head. Fig.4 displays the first four moments. Compared to Fig. 2 the skewness and kurtosis signals are close to zero, indicating reliable spectral data. Table I and II give an overview of the numeric values obtained with N_{SNR} = 250. The parameters κ_{skewness}, κ_{kurtosis} as well as variance(κ_{skewness}) and variance(κ_{kurtosis}) are sensitive parameters to test the signal reliability.

Table I
κ- and c_κ-values for different signal types

Signal type	κ_{mean}	κ_{variance}	c_κ (%)	κ_{skewness}	κ_{kurtosis} †
Gaussian‡	0.1418	0.9898	702	0.3272	0.6101
Scanner‡	1.384	91.59	693	0.3409	0.6348
rest♯	129.7	823.3	22.1	0.3214	0.6427
moved♯	90.40	6161	86.8	1.079	2.883

† ‘Excess’ kurtosis quoted (= kurtosis - 3); ‡ noise; ♯ *in vivo* MRS.

Table II
Estimated variances of the κ-parameters of Table I

Signal type	κ_{mean}	κ_{variance}	c_κ (%)	κ_{skewness}	κ_{kurtosis} †
Gaussian‡	0.0115	0.0660	N.A.	0.0673	0.4155
Scanner‡	1.204	556.6	N.A.	0.0773	0.4798
rest♯	36031.2	323001	N.A.	0.0635	0.3638
moved♯	29441.0	$3.776.10^8$	N.A.	0.0769	0.5159

† ‘Excess’ kurtosis quoted (= kurtosis - 3); ‡ noise; ♯ *in vivo* MRS.

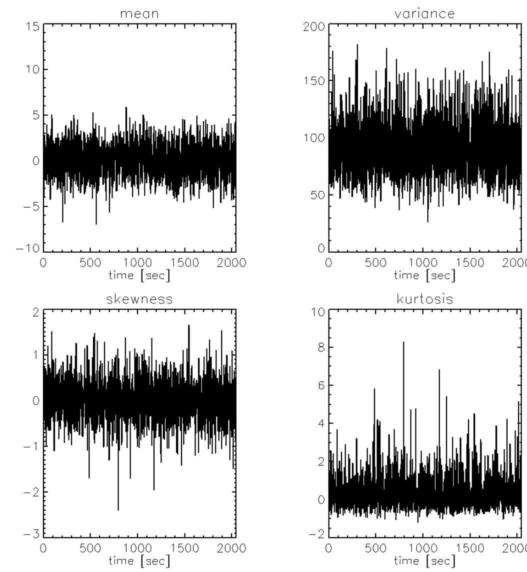

Figure 3. The estimated first four moments about the mean of scanner noise measured on a 3T Siemens Trio scanner. Acquired were 32 signals (consisting of 8 averages each). The observed average values of the mean=1.384, variance=91.59, skewness=0.3409, kurtosis=0.6348. For pure Gaussian noise, the values for the mean, skewness and kurtosis are zero.

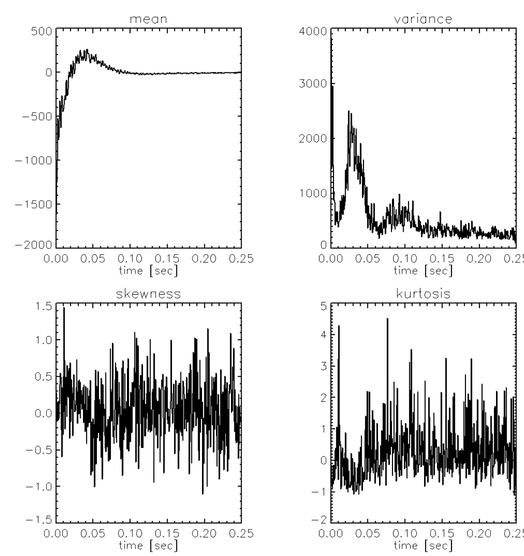

Figure 4. The first four moments of the time domain MRS-signal; test person was asked to refrain from any movements during recording phase. The separately stored signal consists of 32 acquisitions of 8 averages of a complete EXOR-cycle, thus $M = 32$.

C. Median Filtering versus Signal Averaging

Figures 5 and 6 show in vivo ^{1}H muscle spectra (TE 135) of the human tibialis anterior muscle recorded at 3T. The

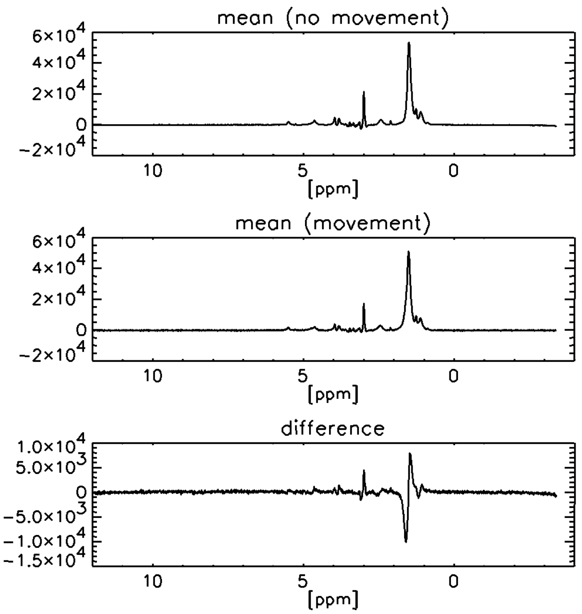

Figure 5. *In vivo*[1]H-muscle spectra: Upper: mean spectrum (rest), Middle: mean spectrum (movement) and Lower: Difference of upper and middle.

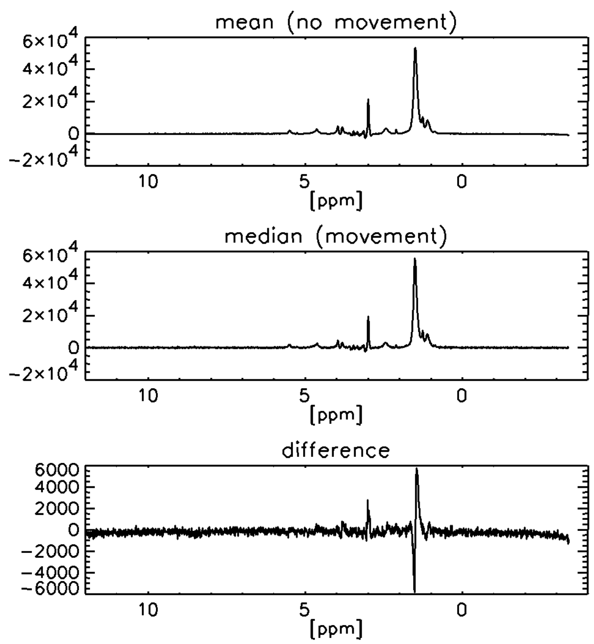

Figure 6. *In vivo*[1]H-muscle spectra: Upper mean spectrum (rest), Middle: Median spectrum (movement) and Lower: Difference of upper and middle.

upper spectra in both figures are the same and are average spectra recorded asking the test person to refrain from any movement. The middle spectra of Fig.5 shows the averaged spectrum and in Fig. 6 the median filtered spectrum. Note that the difference between the median filtered moved- and the average non-movement- spectrum (lower spectrum of Fig. 6) is much smaller than the averaged moved spectrum (lower spectrum Fig. 5) illustrating the ability of the median filter to remove signal artifacts.

IV. Conclusion

Statistical tests are proposed, able to detect the presence of heavily tailed noise components in localized *in vivo* MRS-signals. Most sensitive test parameters are the estimated sample skewness and sample kurtosis parameters determined from separately stored signal responses. Once an MRS-spectrum is classified to be unreliable, non-linear reconstructive filtering techniques can be applied to eliminate these signal artifacts. Median filtering, a non-linear order statistics filtering technique, is shown to be able to remove heavily tailed noise components from *in vivo* MRS-signals.

Acknowledgment

This work is supported by the Swiss National Foundation 3200B0-107499/1.

References

[1] R. Kreis, "Issues of Spectral Quality in clinical 1H-Magnetic Resonance Spectroscopy and a Gallery of Artifacts." *NMR Biomed*, vol. 17, pp. 361–381, 2004.

[2] S. Cavassila, S. Deval, C. Huegen, D. van Ormondt, and D. Graveron-Demilly, "Cramér-Rao bounds: A tool for Quantitation Objectives." *NMR in Biomedicine*, vol. 14, pp. 278–283, 2001.

[3] A. van den Bos, *Parameter Estimation*. John Wiley and Sons, 1982, ch. 8, pp. 331–377, in: Handbook of Measurement Science, Vol. 1, P.H. Sydenham Ed.

[4] ——, "Estimation of Fourier coefficients," *IEEE Trans. Intrumentation and Measurement*, 1989.

[5] ——, *Parameter Estimation for Scientists and Engineers*. Wiley, 2007.

[6] R. Ernst and W. Anderson, "Applications of Fourier Transform Spectroscopy to Magnetic Resonance," *Rev. Sci. Instr.*, vol. 37, pp. 93–102, 1966.

[7] R. Brcic, "Some Aspects of Signal Processing in Heavy Tailed Noise," Curtin University of Technology, Australia, Tech. Rep., 2002.

[8] D. Zwillinger and S. Kokoska, *Standard Probability and Statistics Tables and Formulae*. Chapman & Hall/CRC, 2000.

[9] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in Pascal: The Art of Scientific Computing*. Cambridge University Press, 1990.

[10] D. Joanes and C. Gill, "Comparing measures of sample skewness and kurtosis," *The Statistician*, vol. 47, no. 1, pp. 183–189, 1998.

[11] E. Cho, M. Cho, and J. Eltinge, http://www.bls.gov/ore/abstract/st/st030280.htm, 2003.

[12] E. Cho, "The variance of sample variance for a finite population," in *Section on Survey Research Methods*, 2004, pp. 3345–3350.

[13] H. Cramr, *Mathematical Methods of Statistics*. Princeton: Princeton University Press, 1946.

[14] M. Gabbouj and J. Astola, "Nonlinear order statistic filter design: Methodologies and challenges," in *Proceedings of Eusipco*, Tampere, Finland, September 2000, pp. 377–384.

[15] P. K. J. Astola J, *Fundamentals of Nonlinear Digital Filtering*. CRC Press, 1997.

[16] K. Barner and G. Arce, "Order-statistic filtering and smoothing of time-series: Part II," in *Handbook of Statistics*, N. Balakrishnan and C. Rao, Eds. Elsevier, 1998.

[17] J. Tukey, "Non-linear (non-super-imposable) methods for smoothing data," in *Conf. Rec (Eascon)*, 1974.