LDA Data Filtering

Holger Nobach

September 13, 1999

Abstract

The LDA technique is widely used in the study of turbulent flow fields. The two most important statistical functions with information about the fluid dynamics are the autocorrelation function (ACF) and the power spectral density (PSD). The techniques for calculating the ACF or the PSD of an LDA data set are sophisticated. Nevertheless, the computation becomes difficult for a dominating low-frequency periodicity while the interesting frequency range lies much higher. This report compares filter techniques statistically, derives correction algorithms for systematic deviations and investigates their usefulness. In the end a recommendation is given how to handle this kind of data set to avoid systematic errors with a minimum of estimation's variability.

1 Introduction

The LDA technique is widely used in the study of turbulent flow fields. Two important statistical functions with information about the fluid dynamics are the autocorrelation function (ACF) and the power spectral density (PSD). The techniques for calculating the ACF or the PSD of an LDA data set are sophisticated [2, 4]. The ACF and the PSD correspond through the Fourier transform, hence they contain the same information. Therefore, estimators for ACF or PSD can be compared by transforming one of these into the other. It is nearly the same which domain is used for the comparison. Even the statistics (mean and variance) are comparable in time domain as well as in frequency domain through the linearity of the Fourier transform. Nevertheless, the computation becomes difficult for a dominating low-frequency periodicity, while the interesting frequency range lies much higher, i.e. the investigation of micro-turbulence within cyclic flow fields.

In [8] a filtering technique is used to reduce the estimation's variability. But it can be used to isolate the interesting high-frequency part of the spectrum containing the information about micro-scale turbulence as well. Unfortunately, the results in [8] have a constant value for frequencies below a characteristic cut-off frequency depending on the filter size. The description of the used filter suggests that the PSD should disappear for lower frequencies. That effect was not explained contentedly.

Furthermore, the ACF/PSD of the filtered data are changed through the filter in a characteristic way. That leads to a systematic error in the ACF/PSD estimation. It can be seen in [8] very clearly for the strong high pass filter with (n=1). Therefore, the authors judged this filter to be not suitable. But this influence does not appear for filters with higher order, only the characteristic cut-off frequency is shifted and the errors become acceptable for a given frequency range.

The specific characteristic of the pre-filter technique is completely different to that of the "particle rate filter" for LDA data reconstruction, i.e. for the sample-and-hold reconstruction [1]. Therefore, the results of the pre-filtering technique are more realistic, especially for high frequencies. Nevertheless, an estimation of the systematic deviations would be usefull to judge the reliability of the results and to design a correction filter similar to the procedure given in [7].

Because of the non-regularly sampling the filtering of LDA data is non-linear. Therefore, a description of the system using an impulse response is not suitable. That exclude all known filtering techniques for equidistant sampled data sets, like convolution of correlation functions or spectral amplification. A possible description of non-linear systems, the estimation of statistical functions and their refinement is given in [7] where this technique was applied to the sample-and-hold reconstruction successfully.

2 Filtering Techniques

2.1 Symmetric Filter

2.1.1 Constant Number of Samples

The first filter is similar to that in [8]. The original LDA data samples $x_i = x(t_i)$ at sample times t_i are averaged to a local mean μ_i

$$\mu_i = \frac{1}{2M+1} \sum_{j=-M}^{M} x_{i+j} \tag{1}$$

with a fixed number of samples M, on either side. The sample $y_{\mathrm{HP}i} = y(t_i)$ of the high pass (HP) filtered series is found by

$$y_{\mathrm{HP}i} = x_i - \mu_i \tag{2}$$

leading to an LDA data series with the same sampling scheme like the original data set. Furthermore, the local mean is used as the corresponding low pass (LP) signal.

$$y_{\mathrm{LP}i} = \mu_i \tag{3}$$

The *Fuzzy Slotting Technique* (FST) [6]

$$R_k = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} f_k(t_j - t_i) y_i y_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} f_k(t_j - t_i)} \tag{4}$$

$$f_k(\tau) = \begin{cases} 1 - \left| \frac{\tau}{\Delta\tau} - k \right| & \text{for } \left| \frac{\tau}{\Delta\tau} - k \right| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

with the total number of samples N is used to estimate the ACF of the data set. The *Local Normalization* (LN) [9, 5] and the *Merged Technique* [4] are not used because of the the more difficult derivation of the filter characteristic. Figure 6 shows the filter effect on the PSD for a simulated data set.

2.1.2 Constant Time

The second filter uses all original samples within an constant time window $[t_i - B; t_i + B]$ with a given maximum delay B, symmetric to the time t_i of the sample x_i. The local mean is given through

$$\mu_i = \frac{\sum_{j=1}^{N} b(t_j - t_i) x_j}{\sum_{j=1}^{N} b(t_j - t_i)} \tag{6}$$

with the top head window function

$$b(\Delta t) = \begin{cases} 1 & \text{for } |\Delta t| \le B \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

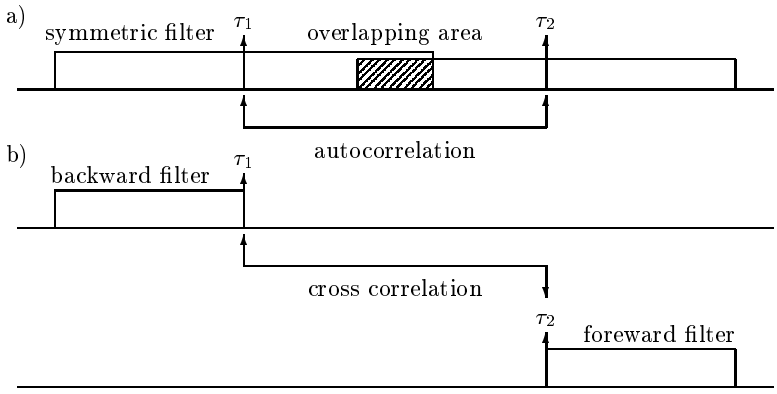

Figure 1: Filter scheme a) symmetric filter; b) asymmetric filter using the cross correlation

and the total number of samples N in the data set. Therefore, the expression $\sum_{j=1}^{N} b(t_j - t_i)$ gives the number of samples within the interval. The relative complicated mathematical description of this filter does not reflect the possibility of an algorithmic implementation correctly. Using the slot correlation a time limited data buffer already exists. Therefore, all routines of the slot correlation can be used for the filtering as well. The filtered samples $y_{\mathrm{HP}i}$ and $y_{\mathrm{LP}i}$ are defined similar to the filter using a constant number of samples (section 2.1.1). Figure 8 shows the filter effect on the PSD for a simulated data set.

2.2 Asymmetric Filter

The mathematical description of the filter characteristics and the derivation of the expected ACF of the filtered data set (section 3) becomes much more easy for asymmetrical filtering. In opposit to the symmetric filters, the asymmetric filters use only samples before or after the processed data sample for the calculation of the local mean (figure 1b). For a constant number of samples the local mean becomes

$$\mu_{\mathrm{F}i} = \frac{1}{M+1} \sum_{j=0}^{M} x_{i+j} \tag{8}$$

$$\mu_{\mathrm{B}i} = \frac{1}{M+1} \sum_{j=-M}^{0} x_{i+j} \tag{9}$$

with the indices F refering to the *foreward* filter and B refering to the *backward* filter. Note, that the backward filter is causal (excepting $j = 0$) and the foreward filter is non-causal. For a constant time window the local mean becomes

$$\mu_{\mathrm{F}i} = \frac{\sum_{j=i}^{N-1} b(t_j - t_i) x_j}{\sum_{j=i}^{N-1} b(t_j - t_i)} \tag{10}$$

$$\mu_{\mathrm{B}i} = \frac{\sum_{j=0}^{i} b(t_j - t_i) x_j}{\sum_{j=0}^{i} b(t_j - t_i)} \tag{11}$$

respectively. The filtered samples $y_{\mathrm{HP}i}$ and $y_{\mathrm{LP}i}$ are defined similar to the filter using the symmetric filter (section 2.1).

To obtain a relatively easy description of the filter characteristics, the overlapping of filter intervals during the convolution should be prevent (figure 1a). Therefore, instead of the ACF, the cross correlation function (CCF) of the foreward and the backward filtered data is used.

Only time lags of the CCF with non-overlapping filter intervals are taken as the ACF estimation (figure 1b).

There are four advantages of the asymmetric filters:

1. a relatively simple mathematical description of the filter characteristics,
2. the filter characteristics function becomes linear (for the ACF, not for the time function),
3. a non-disappearing low-frequency PSD in the high pass path,
4. there are no modifications necessary for the LN.

The first two points lead to a reduced number of calculations. That makes an easy implementation of the algorithm in real applications possible. The third point looks like a disadvantage in the first view, but if the power in the low frequency range is not suppressed completely the correction coefficients in the inverted filter matrix (see section 4) become not as heavy as for the symmetric filter and the corrected estimation becomes more reliable (especially the ACF estimation). With the fourth point the *Merged Technique*, combining the FST and the LN

$$R_k = \frac{\sqrt{\hat{\sigma}^2_{y_\mathrm{F}} \hat{\sigma}^2_{y_\mathrm{B}}} \sum_{i=1}^{N} \sum_{j=1}^{N} f_k(t_j - t_i) y_{\mathrm{B}i} y_{\mathrm{F}j}}{\sqrt{\left(\sum_{i=1}^{N} \sum_{j=1}^{N} f_k(t_j - t_i) y^2_{\mathrm{B}i}\right) \left(\sum_{i=1}^{N} \sum_{j=1}^{N} f_k(t_j - t_i) y^2_{\mathrm{F}j}\right)}} \tag{12}$$

$$f_k(\tau) = \begin{cases} 1 - \left|\frac{\tau}{\Delta\tau} - k\right| & \text{for } \left|\frac{\tau}{\Delta\tau} - k\right| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$$\hat{\sigma}^2_{y_\mathrm{F}} = \frac{1}{N-1} \sum_{i=1}^{N} \left(y_{\mathrm{F}i} - \bar{y}_\mathrm{F}\right)^2 \tag{14}$$

$$\hat{\sigma}^2_{y_\mathrm{B}} = \frac{1}{N-1} \sum_{i=1}^{N} \left(y_{\mathrm{B}i} - \bar{y}_\mathrm{B}\right)^2 \tag{15}$$

$$\bar{y}_\mathrm{F} = \frac{1}{N} \sum_{i=1}^{N} y_{\mathrm{F}i} \tag{16}$$

$$\bar{y}_\mathrm{B} = \frac{1}{N} \sum_{i=1}^{N} y_{\mathrm{B}i} \tag{17}$$

can be used to calculate the CCF/ACF. Figures 9 and 10 show the filter effect on the PSD for a simulated data set using an asymmetric filter with a constant number of samples and with a constant time window respectively.

3 Filter Characteristics

3.1 Symmetric Filter

3.1.1 Constant Number of Samples

The filter characteristic can be described through the expectation of the ACF R' using the filtered data in terms of the original ACF R (see [7]). The derivations are similar for the high pass and the low pass. Only two different coefficients are necessary:

$$\begin{array}{llll} v_s = 1 & \text{and} & v_o = 1 & \text{for the low pass} \\ v_s = 2M & \text{and} & v_o = -1 & \text{for the high pass} \end{array}$$

The filtered samples are

$$y_i = \frac{1}{2M+1}\left[v_s x_i + v_o\left(\sum_{j=-M}^{-1} x_{i+j} + \sum_{j=1}^{M} x_{i+j}\right)\right] \tag{18}$$

To derive the expectation of R'_k for a given time lag $\tau = k\Delta\tau$ with the temporal resolution $\Delta\tau$ the filtered samples at two different points in time $\tau_1 = k_1\Delta\tau$ and $\tau_2 = k_2\Delta\tau$ (see figure 1a) are of interest.

$$y_{k_1} = \frac{1}{2M+1}\left[v_s x_{k_1} + v_o\left(\sum_{j=-M}^{-1} x_{k_1+j} + \sum_{j=1}^{M} x_{k_1+j}\right)\right] \tag{19}$$

$$y_{k_2} = \frac{1}{2M+1}\left[v_s x_{k_2} + v_o\left(\sum_{j=-M}^{-1} x_{k_2+j} + \sum_{j=1}^{M} x_{k_2+j}\right)\right] \tag{20}$$

For the expectation of the ACF follows

$$E\{R'_k\} = E\{y_{k_1} y_{k_2}\} \tag{21}$$

Before the derivation of this expression a look at the probabilities of the different samples is necessary:

1. The product of the filtered samples y_{k_1} and y_{k_2} influences the result of the final ACF estimation only under the condition that the samples at the two time lags τ_1 and τ_2 are present. Therefore, the probability of presence is 1 for these samples.

2. The presence of the other samples at given times have a probability density $p_s(t) = \dot{n}$ depending on the data rate $\dot{n}$. Note, that the integral of this probability density is not equal to 1 because several samples can occur within a time interval. The samples are independent and non-numerated, they can be exchanged without an influence to the results (becomes important later).

Therefore, the products of every two samples have different joint probabilities:

1. The product $x_{k_1} x_{k_2}$ of the both samples at the time lags τ_1 and τ_2 have the joint probability 1.

   (a) Presuming that x_{k_1} is not a member of the sample group that builds the local mean at τ_2 and vice versa the sum of y_{k_1} contains only x_{k_1} and y_{k_2} contains only x_{k_2} of sample type 1. The product $y_{k_1} y_{k_2}$ contains only $v_s^2 x_{k_1} x_{k_2}$ of product type 1.

   (b) If x_{k_1} is a member of this local group for τ_2 and vice versa then both sums contain both samples. The product $y_{k_1} y_{k_2}$ contain $(v_s x_{k_1} + v_o x_{k_2})(v_o x_{k_1} + v_s x_{k_2}) = (v_s^2 + v_o^2) x_{k_1} x_{k_2} + v_s v_o (x_{k_1}^2 + x_{k_2}^2)$.

2. The products $x_{k_1} x_j$ with $j \neq k_1$ and $x_{k_2} x_j$ with $j \neq k_2$ of one sample at τ_1 or τ_2 and another sample at the time t ($t \notin [\tau_1; \tau_2]$) have the joint probability $\dot{n}$.

   (a) Presuming that x_j is a member of the local group at τ_1 and not a member of the local group at τ_2 and x_{k_1} is not a member of the local group at τ_2 and vice versa the sum of y_{k_1} contains only x_j and y_{k_2} contains only x_{k_2} of sample types building the product type 2. The product $y_{k_1} y_{k_2}$ contains only $v_s v_o x_j x_{k_2}$ of product type 2.

   (b) Presuming that x_j is a member of the local group at τ_2 and not a member of the local group at τ_1 and x_{k_1} is not a member of the local group at τ_2 and vice versa the sum of y_{k_1} contains only x_{k_1} and y_{k_2} contains only x_j of sample types building the product type 2. The product $y_{k_1} y_{k_2}$ contains only $v_s v_o x_j x_{k_1}$ of product type 2.

(c) Presuming that x_j is a member of the local group at τ_1 and a member of the local group at τ_2 and x_{k_1} is not a member of the local group at τ_2 and vice versa the sum of y_{k_1} contains x_{k_1} and x_j and y_{k_2} contains x_{k_2} and x_j of sample types building the product type 2. The product $y_{k_1}y_{k_2}$ contains only $v_s v_o x_j(x_{k_1}+x_{k_2})+v_o^2 x_j^2$ of product type 2. The self product of x_j is also of type 2, because if the sample exists also the product exists.

(d) Presuming that x_j is a member of the local group at τ_1 and not a member of the local group at τ_2 and x_{k_1} is a member of the local group at τ_2 and vice versa the sum of y_{k_1} contains only x_j and y_{k_2} contains x_{k_1} and x_{k_2} of sample types building the product type 2. The product $y_{k_1}y_{k_2}$ contains $v_o x_j(v_o x_{k_1}+v_s x_{k_2})$ of product type 2.

(e) Presuming that x_j is a member of the local group at τ_2 and not a member of the local group at τ_1 and x_{k_1} is a member of the local group at τ_2 and vice versa the sum of y_{k_1} contains x_{k_1} and x_{k_2} and y_{k_2} contains only x_j of sample types building the product type 2. The product $y_{k_1}y_{k_2}$ contains $v_o x_j(v_s x_{k_1}+v_o x_{k_2})$ of product type 2.

(f) Presuming that x_j is a member of the local group at τ_1 and a member of the local group at τ_2 and x_{k_1} is a member of the local group at τ_2 and vice versa the sums of y_{k_1} and y_{k_2} contains x_{k_1}, x_{k_2} and x_j of sample types building the product type 2. The product $y_{k_1}y_{k_2}$ contains $v_o x_j(v_s+v_o)(x_{k_1}+x_{k_2})+v_o^2 x_j^2$ of product type 2.

3. The products $x_i x_j$ with $i;j \notin [k_1;k_2]$ of two samples at t_i and t_j ($t_i;t_j \notin [\tau_1;\tau_2]$) have the joint probability $\dot{n}^2$. Because the samples are independent, it can be presumed that $t_i < t_j$. With $\tau_1 < \tau_2$ the following four subcases are left.

(a) If x_i is a member of the local group at τ_1 and not a member of the local group at τ_2 and x_j is a member of the local group at τ_2 and not a member of the local group at τ_1 then the sum of y_{k_1} contains only x_i and the sum of y_{k_2} contains only x_j of samples building a product of type 3. The product $y_{k_1}y_{k_2}$ contains only $v_o^2 x_i x_j$ of type 3.

(b) If x_i is a member of the local group at τ_1 and not a member of the local group at τ_2 and x_j is a member of the local group at τ_2 and a member of the local group at τ_1 then the sum of y_{k_1} contains x_i and x_j and the sum of y_{k_2} contains only x_j of samples building a product of type 3. The product $y_{k_1}y_{k_2}$ contains only $v_o^2 x_i x_j$ of type 3.

(c) If x_i is a member of the local group at τ_1 and a member of the local group at τ_2 and x_j is a member of the local group at τ_2 and not a member of the local group at τ_1 then the sum of y_{k_1} contains only x_i and the sum of y_{k_2} contains x_i and x_j of samples building a product of type 3. The product $y_{k_1}y_{k_2}$ contains only $v_o^2 x_i x_j$ of type 3.

(d) If x_i and x_j are members of both local groups at τ_1 and τ_2 then the sum of y_{k_1} and y_{k_2} contain x_i and x_j of samples building a product of type 3. The product $y_{k_1}y_{k_2}$ contains $2v_o^2 x_i x_j$ of type 3.

Now equation (21) can be solved. It follows

$$
\begin{aligned}
E\{R'_k\} &= E\{y_{k_1}y_{k_2}\} = E\{y(\tau_1)y(\tau_2)\} \\
&= \frac{1}{(2M+1)^2}\left\{p_{1\mathrm{a}}E\{v_s^2 x(\tau_1)x(\tau_2)\}+p_{1\mathrm{b}}E\{(v_s x(\tau_1)+v_o x(\tau_2))(v_o x(\tau_1)+v_s x(\tau_2))\}\right. \\
&\quad + \int_{-\infty}^{\infty} p_s(t)\left[p_{2\mathrm{a}}(t)E\{v_s v_o x(t)x(\tau_2)\}+p_{2\mathrm{b}}(t)E\{v_s v_o x(t)x(\tau_1)\}+p_{2\mathrm{c}}(t)E\{v_s v_o x(t)\right. \\
&\quad (x(\tau_1)+x(\tau_2))+v_o^2 x^2(t)\}+p_{2\mathrm{d}}(t)E\{v_o x(t)(v_o x(\tau_1)+v_s x(\tau_2))\}+p_{2\mathrm{e}}(t)E\{v_o x(t) \\
&\quad \left.\left.(v_s x(\tau_1)+v_o x(\tau_2))\}+p_{2\mathrm{f}}(t)E\{v_o x(t)(v_s+v_o)(x(\tau_1)+x(\tau_2))+v_o^2 x^2(t)\}\right]\,\mathrm{d}t
\end{aligned}
$$

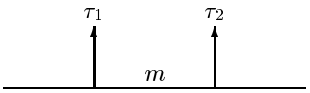

Figure 2: Time and sample case for product type 1

$$
\begin{aligned}
&+\int_{-\infty}^{\infty}\int_{t_1}^{\infty} p_s(t_1)p_s(t_2)(p_{3\mathrm{a}}(t_1;t_2)+p_{3\mathrm{b}}(t_1;t_2)+p_{3\mathrm{c}}(t_1;t_2)\\
&+2p_{3\mathrm{d}}(t_1;t_2))E\{v_o^2x(t_1)x(t_2)\}\,\mathrm{d}t_2\mathrm{d}t_1\Big\} \qquad (22)\\
=\;&\frac{1}{(2M+1)^2}\left\{p_{1\mathrm{a}}v_s^2R(\tau)+p_{1\mathrm{b}}((v_s^2+v_o^2)R(\tau)+2v_sv_oR(0))\right.\\
&+\int_{-\infty}^{\infty} p_s(t)\left[p_{2\mathrm{a}}(t)v_sv_oR(\tau-t)+p_{2\mathrm{b}}(t)v_sv_oR(-t)+p_{2\mathrm{c}}(t)(v_sv_o(R(-t)\right.\\
&+R(\tau-t))+v_o^2R(0))+p_{2\mathrm{d}}(t)(v_o^2R(-t)+v_sv_oR(\tau-t))+p_{2\mathrm{e}}(t)(v_sv_oR(-t)\\
&\left.+v_o^2R(\tau-t))+p_{2\mathrm{f}}(t)(v_o(v_s+v_o)(R(-t)+R(\tau-t))+v_o^2R(0))\right]\,\mathrm{d}t\\
&+\int_{-\infty}^{\infty}\int_{t_1}^{\infty} p_s(t_1)p_s(t_2)(p_{3\mathrm{a}}(t_1;t_2)+p_{3\mathrm{b}}(t_1;t_2)+p_{3\mathrm{c}}(t_1;t_2)\\
&\left.+2p_{3\mathrm{d}}(t_1;t_2))v_o^2R(t_2-t_1)\,\mathrm{d}t_2\mathrm{d}t_1\right\} \qquad (23)
\end{aligned}
$$

with the probability density $p_s(t)$ of the existence of a sample at the time t and the probability p_{xy} of the cases described above. The integration way of the second integral in the double integral is a result of the independent (non-numbered) samples in the data set. Note, that p_{xy} depends on the order of τ_1, τ_2, t or t_1 and t_2 respectively and on the number of samples between these points in time.

For the product types (1, 2 and 3) there are different cases for the membership of τ_1, τ_2 and t or t_1 and t_2 respectively to the local groups at τ_1 and τ_2.

1. product type 1: There are m samples between τ_1 and τ_2 (figure 2). The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m < M$. This can be written like a matrix
$$O = \begin{pmatrix} v_s & v_ov_p(m<M) \\ v_ov_p(m<M) & v_s \end{pmatrix} \qquad (24)$$
with the binary function
$$v_p(x) = \begin{cases} 1 & \text{if } x \text{ is true} \\ 0 & \text{otherwise} \end{cases} \qquad (25)$$
giving the contribution of x_{k_1} (first column) and x_{k_2} (second column) to the local group of τ_1 (first raw) and τ_2 (second raw). The products $O_{11}O_{21}$, $O_{11}O_{22}$, $O_{12}O_{21}$ and $O_{12}O_{22}$ are of type 1.

2. product type 2:

   (a) The sample time t lies before τ_1. There are m_1 samples between t and τ_1 and m_2 samples between τ_1 and τ_2 (figure 3a). The sample at t is a member of the local group at τ_1 if $m_1 < M$. The sample at t is a member of the local group at τ_2 if $m_1 + m_2 < M - 1$. Note, that the sample at τ_1 is also a member in this case. Therefore, the number of independent samples between t and τ_2 must be less than

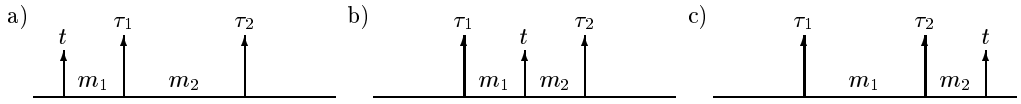

Figure 3: Time and sample case for product type 2

$M-1$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_2 < M$. This can be written like a matrix

$$P_A = \begin{pmatrix} v_s & v_o v_p(m_2 < M) & v_o v_p(m_1 < M) \\ v_o v_p(m_2 < M) & v_s & v_o v_p(m_1 + m_2 < M-1) \end{pmatrix} \tag{26}$$

giving the contribution of x_{k_1} (first column), x_{k_2} (second column) and x_j (third column) to the local group of τ_1 (first raw) and τ_2 (second raw). The products $P_{11}P_{23}$, $P_{12}P_{23}$, $P_{13}P_{21}$, $P_{13}P_{22}$ and $P_{13}P_{23}$ are of type 2.

(b) The sample time t lies between τ_1 and τ_2. There are m_1 samples between τ_1 and t and m_2 samples between t and τ_2 (figure 3b). The sample at t is a member of the local group at τ_1 if $m_1 < M$. The sample at t is a member of the local group at τ_2 if $m_2 < M$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_1 + m_2 < M-1$. Note, that the sample at t is also a member of both groups in this case. Therefore, the number of independent samples between τ_1 and τ_2 must be less than $M-1$. This can be written like a matrix

$$P_B = \begin{pmatrix} v_s & v_o v_p(m_1 + m_2 < M-1) & v_o v_p(m_1 < M) \\ v_o v_p(m_1 + m_2 < M-1) & v_s & v_o v_p(m_2 < M) \end{pmatrix} \tag{27}$$

(c) The sample time t lies after τ_2. There are m_1 samples between τ_1 and τ_2 and m_2 samples between τ_2 and t (figure 3c). The sample at t is a member of the local group at τ_2 if $m_2 < M$. The sample at t is a member of the local group at τ_1 if $m_1 + m_2 < M-1$. Note, that the sample at τ_2 is also a member in this case. Therefore, the number of independent samples between τ_1 and t must be less than $M-1$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_1 < M$. This can be written like a matrix

$$P_C = \begin{pmatrix} v_s & v_o v_p(m_1 < M) & v_o v_p(m_1 + m_2 < M-1) \\ v_o v_p(m_1 < M) & v_s & v_o v_p(m_2 < M) \end{pmatrix} \tag{28}$$

3. product type 3:

(a) The sample times t_1 and t_2 lie before τ_1. There are m_1 samples between t_1 and t_2, m_2 samples between t_2 and τ_1 and m_3 samples between τ_1 and τ_2 (figure 4a). The sample at t_1 is a member of the local group at τ_1 if $m_1 + m_2 < M-1$. The sample at t_2 is a member of the local group at τ_1 if $m_2 < M$. The sample at t_1 is a member of the local group at τ_2 if $m_1 + m_2 + m_3 < M-2$. The sample at t_2 is a member of the local group at τ_2 if $m_2 + m_3 < M-1$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_3 < M$. This can be written like a matrix

$$Q_A = \begin{pmatrix} v_o v_p(m_1 + m_2 < M-1) & v_o v_p(m_2 < M) \\ v_o v_p(m_1 + m_2 + m_3 < M-2) & v_o v_p(m_2 + m_3 < M-1) \end{pmatrix} \tag{29}$$

giving the contribution of $x(t_1)$ (first column) and $x(t_2)$ (second column) to the local group of τ_1 (first raw) and τ_2 (second raw). The samples at τ_1 and τ_2 do not have a contribution of product type 3. Only the products $Q_{11}Q_{22}$ and $Q_{12}Q_{21}$ are of type 3.

(b) The sample time t_1 lies before τ_1 and t_2 lies between τ_1 and τ_2. There are m_1 samples between t_1 and τ_1, m_2 samples between τ_1 and t_2 and m_3 samples between t_2 and τ_2 (figure 4b). The sample at t_1 is a member of the local group at τ_1 if $m_1 < M$. The sample at t_2 is a member of the local group at τ_1 if $m_2 < M$. The sample at t_1 is a member of the local group at τ_2 if $m_1 + m_2 + m_3 < M - 2$. The sample at t_2 is a member of the local group at τ_2 if $m_3 < M$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_2 + m_3 < M - 1$. This can be written like a matrix

$$Q_B = \begin{pmatrix} v_o v_p(m_1 < M) & v_o v_p(m_2 < M) \\ v_o v_p(m_1 + m_2 + m_3 < M - 2) & v_o v_p(m_3 < M) \end{pmatrix} \tag{30}$$

(c) The sample time t_1 lies before τ_1 and t_2 lies after τ_2. There are m_1 samples between t_1 and τ_1, m_2 samples between τ_1 and τ_2 and m_3 samples between τ_2 and t_2 (figure 4c). The sample at t_1 is a member of the local group at τ_1 if $m_1 < M$. The sample at t_2 is a member of the local group at τ_1 if $m_2 + m_3 < M - 1$. The sample at t_1 is a member of the local group at τ_2 if $m_1 + m_2 < M - 1$. The sample at t_2 is a member of the local group at τ_2 if $m_3 < M$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_2 < M$. This can be written like a matrix

$$Q_C = \begin{pmatrix} v_o v_p(m_1 < M) & v_o v_p(m_2 + m_3 < M - 1) \\ v_o v_p(m_1 + m_2 < M - 1) & v_o v_p(m_3 < M) \end{pmatrix} \tag{31}$$

(d) The sample times t_1 and t_2 lie between τ_1 and τ_2. There are m_1 samples between τ_1 and t_1, m_2 samples between t_1 and t_2 and m_3 samples between t_2 and τ_2 (figure 4d). The sample at t_1 is a member of the local group at τ_1 if $m_1 < M$. The sample at t_2 is a member of the local group at τ_1 if $m_1 + m_2 < M - 1$. The sample at t_1 is a member of the local group at τ_2 if $m_2 + m_3 < M - 1$. The sample at t_2 is a member of the local group at τ_2 if $m_3 < M$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_1 + m_2 + m_3 < M - 2$. This can be written like a matrix

$$Q_D = \begin{pmatrix} v_o v_p(m_1 < M) & v_o v_p(m_1 + m_2 < M - 1) \\ v_o v_p(m_2 + m_3 < M - 1) & v_o v_p(m_3 < M) \end{pmatrix} \tag{32}$$

(e) The sample time t_1 lies between τ_1 and τ_2 and t_2 lies after τ_2. There are m_1 samples between τ_1 and t_1, m_2 samples between t_1 and τ_2 and m_3 samples between τ_2 and t_2 (figure 4e). The sample at t_1 is a member of the local group at τ_1 if $m_1 < M$. The sample at t_2 is a member of the local group at τ_1 if $m_1 + m_2 + m_3 < M - 2$. The sample at t_1 is a member of the local group at τ_2 if $m_2 < M$. The sample at t_2 is a member of the local group at τ_2 if $m_3 < M$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_1 + m_2 < M - 1$. This can be written like a matrix

$$Q_E = \begin{pmatrix} v_o v_p(m_1 < M) & v_o v_p(m_1 + m_2 + m_3 < M - 2) \\ v_o v_p(m_2 < M) & v_o v_p(m_3 < M) \end{pmatrix} \tag{33}$$

(f) The sample times t_1 and t_2 lie after τ_2. There are m_1 samples between τ_1 and τ_2, m_2 samples between τ_2 and t_1 and m_3 samples between t_1 and t_2 (figure 4f). The sample at t_1 is a member of the local group at τ_1 if $m_1 + m_2 < M - 1$. The sample at t_2 is a member of the local group at τ_1 if $m_1 + m_2 + m_3 < M - 2$. The sample at t_1 is a member of the local group at τ_2 if $m_2 < M$. The sample at t_2 is a member of the local group at τ_2 if $m_2 + m_3 < M - 1$. The sample at τ_1 is a member of the local group at τ_2 and vica versa if $m_1 < M$. This can be written like a matrix

$$Q_F = \begin{pmatrix} v_o v_p(m_1 + m_2 < M - 1) & v_o v_p(m_1 + m_2 + m_3 < M - 2) \\ v_o v_p(m_2 < M) & v_o v_p(m_2 + m_3 < M - 1) \end{pmatrix} \tag{34}$$

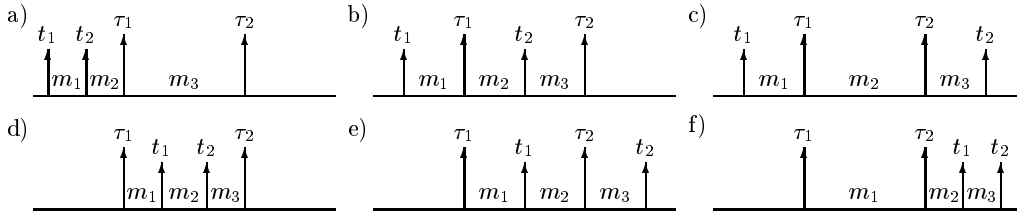

Figure 4: Time and sample case for product type 3

To make the joint probabilities of the several time order and product cases not too complicated, a summation of possible numbers of m, m_1, m_2 and m_3 for each time order case is done. The products are divided into the three types. The described product subtypes are represented by pre-factors depending on the number of independent samples m, m_1, m_2 and m_3 between the times τ_1, τ_2, t_1 and t_2. They are given through the matrices P above. Therefore, equation (23) becomes

$$
\begin{aligned}
E\{R'_k\} \;=\; & \frac{1}{(2M+1)^2}\sum_{m=0}^{\infty} p_m(m;\tau)\,[(O_{11}O_{22}+O_{12}O_{21})R(\tau)+(O_{11}O_{21}+O_{12}O_{22})R(0)] \\
& +\frac{\dot n}{(2M+1)^2}\int_{-\infty}^{0}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} p_m(m_1;-t)p_m(m_2;\tau)\,[(P_{A11}P_{A23} \\
& +P_{A13}P_{A21})R(-t)+(P_{A12}P_{A23}+P_{A13}P_{A22})R(\tau-t)+P_{A13}P_{A23}R(0)]\Bigg\}\,\mathrm{d}t \\
& +\frac{\dot n}{(2M+1)^2}\int_{0}^{\tau}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} p_m(m_1;t)p_m(m_2;\tau-t)\,[(P_{B11}P_{B23} \\
& +P_{B13}P_{B21})R(t)+(P_{B12}P_{B23}+P_{B13}P_{B22})R(\tau-t)+P_{B13}P_{B23}R(0)]\Bigg\}\,\mathrm{d}t \\
& +\frac{\dot n}{(2M+1)^2}\int_{\tau}^{\infty}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} p_m(m_1;\tau)p_m(m_2;t-\tau)\,[(P_{C11}P_{C23} \\
& +P_{C13}P_{C21})R(t)+(P_{C12}P_{C23}+P_{C13}P_{C22})R(t-\tau)+P_{C13}P_{C23}R(0)]\Bigg\}\,\mathrm{d}t \\
& +\frac{\dot n^2}{(2M+1)^2}\int_{-\infty}^{0}\int_{t_1}^{0}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty} p_m(m_1;t_2-t_1)p_m(m_2;-t_2)p_m(m_3;\tau) \\
& [(Q_{A11}Q_{A22}+Q_{A12}Q_{A21})R(t_2-t_1)]\Bigg\}\,\mathrm{d}t2\,\mathrm{d}t1 \\
& +\frac{\dot n^2}{(2M+1)^2}\int_{-\infty}^{0}\int_{0}^{\tau}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty} p_m(m_1;-t_1)p_m(m_2;t_2)p_m(m_3;\tau-t_2) \\
& [(Q_{B11}Q_{B22}+Q_{B12}Q_{B21})R(t_2-t_1)]\Bigg\}\,\mathrm{d}t2\,\mathrm{d}t1 \\
& +\frac{\dot n^2}{(2M+1)^2}\int_{-\infty}^{0}\int_{\tau}^{\infty}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty} p_m(m_1;-t_1)p_m(m_2;\tau)p_m(m_3;t_2-\tau) \\
& [(Q_{C11}Q_{C22}+Q_{C12}Q_{C21})R(t_2-t_1)]\Bigg\}\,\mathrm{d}t2\,\mathrm{d}t1 \\
& +\frac{\dot n^2}{(2M+1)^2}\int_{0}^{\tau}\int_{t_1}^{\tau}\Bigg\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty} p_m(m_1;t_1)p_m(m_2;t_2-t_1)p_m(m_3;\tau-t_2) \\
& [(Q_{D11}Q_{D22}+Q_{D12}Q_{D21})R(t_2-t_1)]\Bigg\}\,\mathrm{d}t2\,\mathrm{d}t1
\end{aligned}
$$

$$+\frac{\dot{n}^2}{(2M+1)^2}\int_0^{\tau}\int_{\tau}^{\infty}\left\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty}p_m(m_1;t_1)p_m(m_2;\tau-t_1)p_m(m_3;t_2-\tau)\right.$$
$$\left.[(Q_{E11}Q_{E22}+Q_{E12}Q_{E21})R(t_2-t_1)]\right\}\,\mathrm{d}t2\,\mathrm{d}t1$$
$$+\frac{\dot{n}^2}{(2M+1)^2}\int_{\tau}^{\infty}\int_{t_1}^{\infty}\left\{\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_3=0}^{\infty}p_m(m_1;\tau)p_m(m_2;t_1-\tau)p_m(m_3;t_2-t_1)\right.$$
$$\left.[(Q_{F11}Q_{F22}+Q_{F12}Q_{F21})R(t_2-t_1)]\right\}\,\mathrm{d}t2\,\mathrm{d}t1 \tag{35}$$

with the probability

$$p_m(m,\Delta t)=\frac{(\dot{n}\Delta t)^m}{m!}e^{-\dot{n}\Delta t} \tag{36}$$

of m independent samples within the time interval Δt. Because of the certainty that no relation exists over an interval with at least M samples, the summations can be reduced by using the probability

$$p'_m(m,\Delta t)=\begin{cases} p_m(m,\Delta t) & \text{for } m<M \\ 1-\sum_{i=0}^{M-1}p_m(i,\Delta t) & \text{for } m=M \end{cases} \tag{37}$$

instead and building the sum over 0 to M. The integrals in equation (35) are calculated numerically with the time resolution $\Delta\tau$ of the ACF through a substitution by sums. Therefore, equation (35) becomes

$$\begin{aligned}
E\{R'_k\} \;=\;& \frac{1}{(2M+1)^2}\sum_{m=0}^{M}p'_m(m;k\Delta\tau)\,[(O_{11}O_{22}+O_{12}O_{21})R_k+(O_{11}O_{21}+O_{12}O_{22})R_0] \\
&+\frac{\dot{n}\Delta\tau}{(2M+1)^2}\sum_{i=-\infty}^{0}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}v_i p'_m(m_1;-i\Delta\tau)p'_m(m_2;k\Delta\tau)\,[(P_{A11}P_{A23} \\
&+P_{A13}P_{A21})R_{-i}+(P_{A12}P_{A23}+P_{A13}P_{A22})R_{k-i}+P_{A13}P_{A23}R_0] \\
&+\frac{\dot{n}\Delta\tau}{(2M+1)^2}\sum_{i=0}^{k}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}v_i p'_m(m_1;i\Delta\tau)p'_m(m_2;(k-i)\Delta\tau)\,[(P_{B11}P_{B23} \\
&+P_{B13}P_{B21})R_i+(P_{B12}P_{B23}+P_{B13}P_{B22})R_{k-i}+P_{B13}P_{B23}R_0] \\
&+\frac{\dot{n}\Delta\tau}{(2M+1)^2}\sum_{i=k}^{\infty}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}v_i p'_m(m_1;k\Delta\tau)p'_m(m_2;(i-k)\Delta\tau)\,[(P_{C11}P_{C23} \\
&+P_{C13}P_{C21})R_i+(P_{C12}P_{C23}+P_{C13}P_{C22})R_{i-k}+P_{C13}P_{C23}R_0] \\
&+\left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2\sum_{i=-\infty}^{0}\sum_{j=i}^{0}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M}v_i v_j v_= p'_m(m_1;(j-i)\Delta\tau)p'_m(m_2;-j\Delta\tau) \\
&p'_m(m_3;k\Delta\tau)\,[(Q_{A11}Q_{A22}+Q_{A12}Q_{A21})R_{j-i}] \\
&+\left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2\sum_{i=-\infty}^{0}\sum_{j=0}^{k}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M}v_i v_j v_= p'_m(m_1;-i\Delta\tau)p'_m(m_2;j\Delta\tau) \\
&p'_m(m_3;(k-j)\Delta\tau)\,[(Q_{B11}Q_{B22}+Q_{B12}Q_{B21})R_{j-i}] \\
&+\left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2\sum_{i=-\infty}^{0}\sum_{j=k}^{\infty}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M}v_i v_j v_= p'_m(m_1;-i\Delta\tau)p'_m(m_2;k\Delta\tau) \\
&p'_m(m_3;(j-k)\Delta\tau)\,[(Q_{C11}Q_{C22}+Q_{C12}Q_{C21})R_{j-i}] \\
&+\left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2\sum_{i=0}^{k}\sum_{j=i}^{k}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M}v_i v_j v_= p'_m(m_1;i\Delta\tau)p'_m(m_2;(j-i)\Delta\tau)
\end{aligned}$$

parameter	unit	value
model	–	AR2
		$(y_n = \phi_1 y_{n-1} + \phi_2 y_{n-2} + a_n)$
model parameters	–	$\phi_1 = 1.8$, $\phi_2 = -0.9$
sampling frequency (primary series)	kHz	10
mean velocity	ms^{-1}	0.0
variance	$\mathrm{m}^2\mathrm{s}^{-2}$	0.3
observation time	s	1 000
mean data rate (secondary series)	kHz	1.0
noise power	$\mathrm{m}^2\mathrm{s}^{-2}$	0.0
velocity bias	–	no

Table 1: Simulation parameters for single realizations

$$
\begin{aligned}
& p'_m(m_3;(k-j)\Delta\tau)\,[(Q_{D11}Q_{D22}+Q_{D12}Q_{D21})R_{j-i}] \\
& + \left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2 \sum_{i=0}^{k}\sum_{j=k}^{\infty}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M} v_i v_j v_= p'_m(m_1;i\Delta\tau)p'_m(m_2;(k-i)\Delta\tau) \\
& p'_m(m_3;(j-k)\Delta\tau)\,[(Q_{E11}Q_{E22}+Q_{E12}Q_{E21})R_{j-i}] \\
& + \left(\frac{\dot{n}\Delta\tau}{2M+1}\right)^2 \sum_{i=k}^{\infty}\sum_{j=i}^{\infty}\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\sum_{m_3=0}^{M} v_i v_j v_= p'_m(m_1;k\Delta\tau)p'_m(m_2;(i-k)\Delta\tau) \\
& p'_m(m_3;(j-i)\Delta\tau)\,[(Q_{F11}Q_{F22}+Q_{F12}Q_{F21})R_{j-i}] \qquad (38)
\end{aligned}
$$

with the pre-factors

$$
v_i = \begin{cases} 0.5 & \text{for } i \in [0;k] \\ 1 & \text{otherwise} \end{cases} \qquad (39)
$$

$$
v_j = \begin{cases} 0.5 & \text{for } j \in [0;k] \\ 1 & \text{otherwise} \end{cases} \qquad (40)
$$

$$
v_= = \begin{cases} 0.5 & \text{for } i = j \\ 1 & \text{otherwise} \end{cases} \qquad (41)
$$

to minimize the edge artefacts.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 6 shows the results of a computer simulation [3]. It shows a good similarity of the calculated PSD expectation and the PSD estimation from filtered data. The simulation parameters are listed in table 1. A single realization (reduced observation time) of the simulated data set is shown in figure 5.

3.1.2 Constant Time

For the derivation of the ACF expectation $E\{R'_k\}$ of filtered data sets the interaction of two points $\tau_1 = k_1\Delta\tau$ and $\tau_2 = k_2\Delta\tau$ (see figure 1a) is of interest. There are four basic time regimes (see figure 7) depending on the time lag.

1. The time windows of τ_1 and τ_2 are non-overlapping ($\tau \geq 2B$, figure 7a). There are m_1 independent samples within the time window of τ_1 and m_2 independent samples within

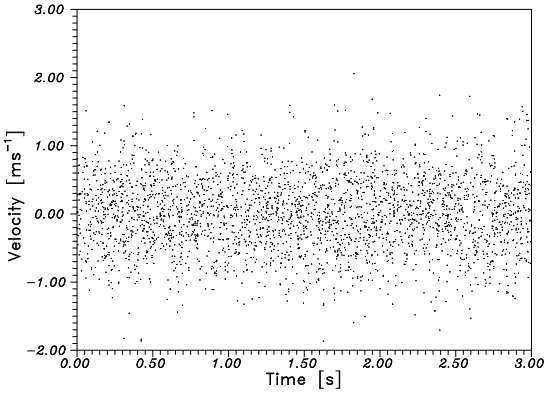

Figure 5: Simulated LDA data set

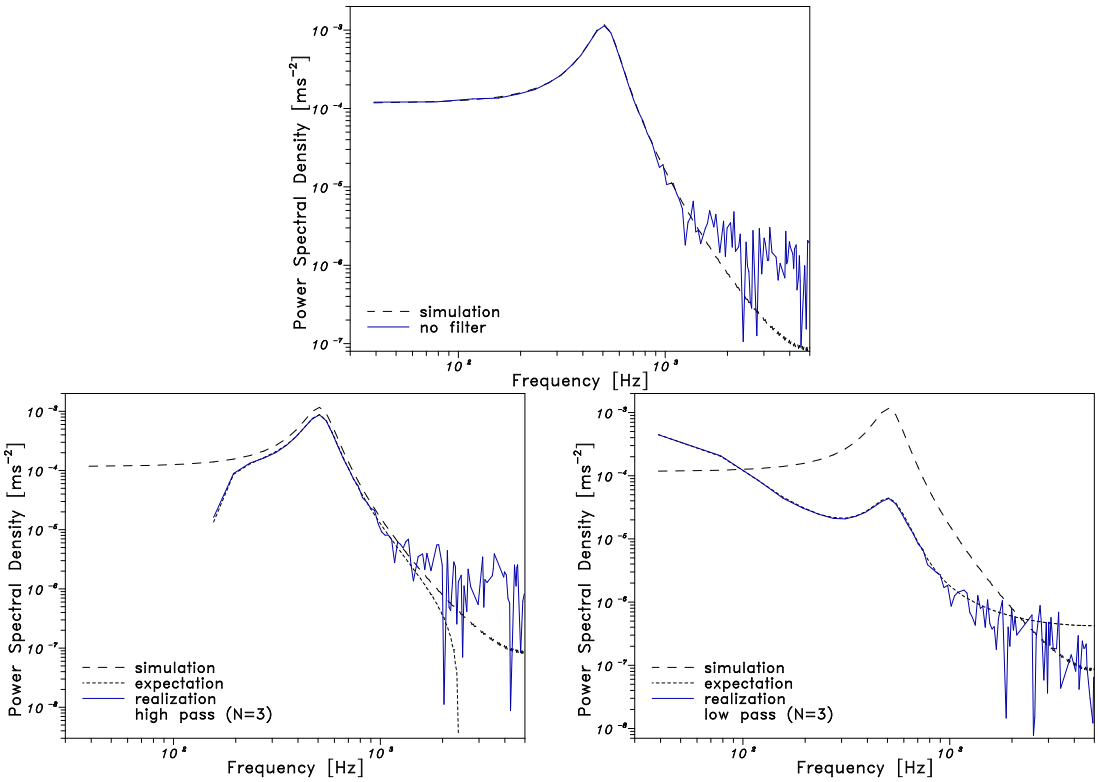

Figure 6: The PSD of a simulated LDA data set, filtered symmetrically with a constant number of samples compared to the PSD of the unfiltered data set (FST and LN)

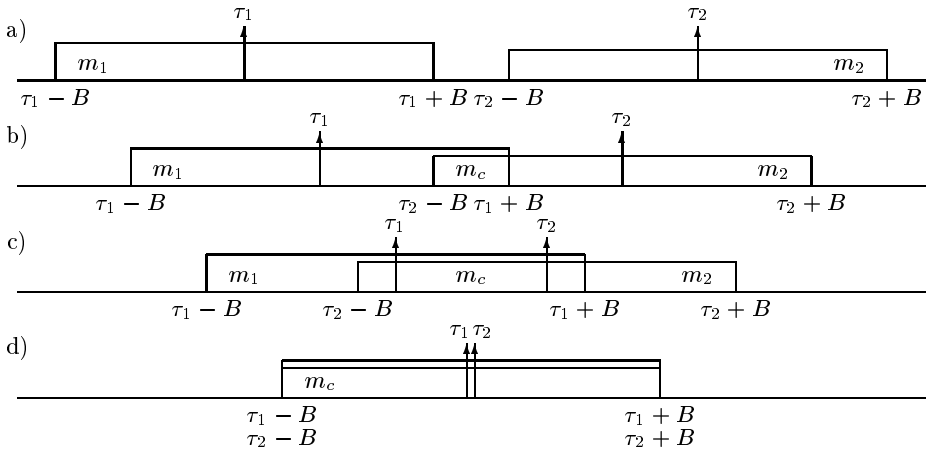

Figure 7: Time regimes for the symmetric filter with a constant time window

the time window of τ_2. The filtered samples are

$$y_{k_1} = \frac{1}{m_1+1}\left(v_{s1}x_{k_1} + v_o\sum_{i=1}^{m_1} x_{s_1+i}\right) \tag{42}$$

$$y_{k_2} = \frac{1}{m_2+1}\left(v_{s2}x_{k_2} + v_o\sum_{i=1}^{m_2} x_{s_2+i}\right) \tag{43}$$

with

$$\begin{array}{llllll} v_{s1}=1 & , & v_{s2}=1 & \text{and} & v_o=1 & \text{for the low pass} \\ v_{s1}=m_1 & , & v_{s2}=m_2 & \text{and} & v_o=-1 & \text{for the high pass} \end{array}$$

The samples within the sums are shifted by the indices s_1 and s_2 respectively. They are separated in that way that all samples are independent.

For the ACF expectation follows

$$E\{R'_k\} = E\{y_{k_1}y_{k_2}\} \tag{44}$$

$$= E\left\{\frac{1}{(m_1+1)(m_2+1)}\left(v_{s1}x_{k_1} + v_o\sum_{i=1}^{m_1} x_{s_1+i}\right)\left(v_{s2}x_{k_2} + v_o\sum_{i=1}^{m_2} x_{s_2+i}\right)\right\}$$

$$= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\frac{p_m(m_1;2B)p_m(m_2;2B)}{(m_1+1)(m_2+1)}E\left\{v_{s1}v_{s2}x_{k_1}x_{k_2} + v_{s1}v_o x_{k_1}\sum_{i=1}^{m_2} x_{s_2+i}\right.$$

$$\left. + v_{s2}v_o x_{k_2}\sum_{i=1}^{m_1} x_{s_1+i} + v_o^2\left(\sum_{i=1}^{m_1} x_{s_1+i}\right)\left(\sum_{i=1}^{m_2} x_{s_2+i}\right)\right\} \tag{45}$$

with the probability $p_m(m;\Delta t)$ of m independent samples within the interval Δt (equation 36). The m_1 and m_2 independent samples are equally distributed in the time interval $[\tau_1 - B;\tau_1 + B]$ and $[\tau_2 - B;\tau_2 + B]$ with the probability density $\frac{m_1}{2B}$ and $\frac{m_2}{2B}$ respectively. Therefore, from equation (45) follows

$$E\{R'_k\} = \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\frac{p_m(m_1;2B)p_m(m_2;2B)}{(m_1+1)(m_2+1)}\left(v_{s1}v_{s2}R(\tau) + \frac{v_{s1}v_o m_2}{2B}\int_{\tau-B}^{\tau+B} R(t)\,\mathrm{d}t\right.$$

$$\left. + \frac{v_{s2}v_o m_1}{2B}\int_{\tau-B}^{\tau+B} R(t)\,\mathrm{d}t + \frac{v_o^2 m_1 m_2}{4B^2}\int_{-B}^{B}\int_{\tau-B}^{\tau+B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\right)$$

$$= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\frac{p_m(m_1;2B)p_m(m_2;2B)}{(m_1+1)(m_2+1)}\left(v_{s1}v_{s2}R(\tau)+\frac{v_o(v_{s1}m_2+v_{s2}m_1)}{2B}\right.$$
$$\left.\int_{\tau-B}^{\tau+B}R(t)\,\mathrm{d}t+\frac{v_o^2m_1m_2}{4B^2}\int_{-B}^{B}\int_{\tau-B}^{\tau+B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\right) \quad (46)$$

2. The time windows of τ_1 and τ_2 are overlapping particularly, x_{k_1} is not a member of the local group at τ_2 and vice versa ($B<\tau<2B$, figure 7b). There are m_1 independent samples within the interval $[\tau_1-B;\tau_2-B]$, m_2 independent samples within the interval $[\tau_1+B;\tau_2+B]$ and m_c independent samples within the interval $[\tau_2-B;\tau_1+B]$. The filtered samples are

$$y_{k_1} = \frac{1}{m_1+m_c+1}\left(v_{s1}x_{k_1}+v_o\sum_{i=1}^{m_1}x_{s_1+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right) \quad (47)$$

$$y_{k_2} = \frac{1}{m_2+m_c+1}\left(v_{s2}x_{k_2}+v_o\sum_{i=1}^{m_2}x_{s_2+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right) \quad (48)$$

with

$v_{s1}=1$	,	$v_{s2}=1$	and $v_o=1$	for the low pass
$v_{s1}=m_1+m_c$	,	$v_{s2}=m_2+m_c$	and $v_o=-1$	for the high pass

The samples within the sums are shifted by the indices s_1, s_2 and s_c respectively. They are separated in that way that the samples of different sums are independent.

For the ACF expectation follows

$$E\{R'_k\} = E\{y_{k_1}y_{k_2}\} \quad (49)$$

$$= E\left\{\frac{1}{(m_1+m_c+1)(m_2+m_c+1)}\left(v_{s1}x_{k_1}+v_o\sum_{i=1}^{m_1}x_{s_1+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right)\right.$$
$$\left.\left(v_{s2}x_{k_2}+v_o\sum_{i=1}^{m_2}x_{s_2+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right)\right\}$$
$$= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_c=0}^{\infty}\frac{p_m(m_1;\tau)p_m(m_2;\tau)p_m(m_c;2B-\tau)}{(m_1+m_c+1)(m_2+m_c+1)}E\left\{v_{s1}v_{s2}x_{k_1}x_{k_2}\right.$$
$$+v_{s1}v_ox_{k_1}\sum_{i=1}^{m_2}x_{s_2+i}+v_{s2}v_ox_{k_2}\sum_{i=1}^{m_1}x_{s_1+i}+v_o(v_{s1}x_{k_1}+v_{s2}x_{k_2})\sum_{i=1}^{m_c}x_{s_c+i}$$
$$+v_o^2\left(\sum_{i=1}^{m_1}x_{s_1+i}\right)\left(\sum_{i=1}^{m_2}x_{s_2+i}\right)+v_o^2\left(\sum_{i=1}^{m_1}x_{s_1+i}+\sum_{i=1}^{m_2}x_{s_2+i}\right)\left(\sum_{i=1}^{m_c}x_{s_c+i}\right)$$
$$\left.+v_o^2\left(\sum_{i=1}^{m_c}x_{s_c+i}\right)^2\right\} \quad (50)$$

The m_1, m_2 and m_c independent samples are equally distributed in the time interval $[\tau_1-B;\tau_2-B]$, $[\tau_1+B;\tau_2+B]$ and $[\tau_2-B;\tau_1+B]$ with the probability density $\frac{m_1}{\tau}$, $\frac{m_2}{\tau}$ and $\frac{m_c}{2B-\tau}$ respectively. The products of the sums are independent, excepting the last term. The square of the sum has to be written as

$$\left(\sum_{i=1}^{m_c}x_{s_c+i}\right)^2=\sum_{i=1}^{m_c}x_{s_c+i}^2+\sum_{i=1}^{m_c}\sum_{\substack{j=1\\j\neq i}}^{m_c}x_{s_c+i}x_{s_c+j} \quad (51)$$

with the products $x_{s_c+i}x_{s_c+j}$ of independent samples.

Therefore, from equation (50) follows

$$
\begin{aligned}
E\{R'_k\} &= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_c=0}^{\infty}\frac{p_m(m_1;\tau)p_m(m_2;\tau)p_m(m_c;2B-\tau)}{(m_1+m_c+1)(m_2+m_c+1)} \\
&\quad\Bigg(v_{s1}v_{s2}R(\tau)+\frac{v_{s1}v_o m_2}{\tau}\int_B^{\tau+B}R(t)\,\mathrm{d}t+\frac{v_{s2}v_o m_1}{\tau}\int_B^{\tau+B}R(t)\,\mathrm{d}t \\
&\quad+\frac{v_{s1}v_o m_c}{2B-\tau}\int_{\tau-B}^{B}R(t)\,\mathrm{d}t+\frac{v_{s2}v_o m_c}{2B-\tau}\int_{\tau-B}^{B}R(t)\,\mathrm{d}t \\
&\quad+\frac{v_o^2 m_1 m_2}{\tau^2}\int_{-B}^{\tau-B}\int_{B}^{\tau+B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1+\frac{v_o^2 m_c m_1}{\tau(2B-\tau)}\int_{-B}^{\tau-B}\int_{\tau-B}^{B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\
&\quad+\frac{v_o^2 m_c m_2}{\tau(2B-\tau)}\int_{\tau-B}^{B}\int_{B}^{\tau+B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1+v_o^2 m_c R(0) \\
&\quad+\frac{v_o^2 m_c(m_c-1)}{(2B-\tau)^2}\int_{\tau-B}^{B}\int_{\tau-B}^{B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\Bigg) \\
&= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_c=0}^{\infty}\frac{p_m(m_1;\tau)p_m(m_2;\tau)p_m(m_c;2B-\tau)}{(m_1+m_c+1)(m_2+m_c+1)} \\
&\quad\Bigg(v_{s1}v_{s2}R(\tau)+\frac{v_o(v_{s1}m_2+v_{s2}m_1)}{\tau}\int_B^{\tau+B}R(t)\,\mathrm{d}t \\
&\quad+\frac{v_o m_c(v_{s1}+v_{s2})}{2B-\tau}\int_{\tau-B}^{B}R(t)\,\mathrm{d}t+\frac{v_o^2 m_1 m_2}{\tau^2}\int_{-B}^{\tau-B}\int_{B}^{\tau+B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\
&\quad+\frac{v_o^2 m_c(m_1+m_2)}{\tau(2B-\tau)}\int_{-B}^{\tau-B}\int_{\tau-B}^{B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1+v_o^2 m_c R(0) \\
&\quad+\frac{v_o^2 m_c(m_c-1)}{(2B-\tau)^2}\int_{\tau-B}^{B}\int_{\tau-B}^{B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\Bigg) \qquad (52)
\end{aligned}
$$

3. The time windows of τ_1 and τ_2 are overlapping particularly, x_{k_1} is a member of the local group at τ_2 and vice versa ($0<\tau\le B$, figure 7c). There are m_1 independent samples within the interval $[\tau_1-B;\tau_2-B]$, m_2 independent samples within the interval $[\tau_1+B;\tau_2+B]$ and m_c independent samples within the interval $[\tau_2-B;\tau_1+B]$. The filtered samples are

$$
y_{k_1} = \frac{1}{m_1+m_c+2}\left(v_{s1}x_{k_1}+v_o x_{k_2}+v_o\sum_{i=1}^{m_1}x_{s_1+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right) \qquad (53)
$$

$$
y_{k_2} = \frac{1}{m_2+m_c+2}\left(v_o x_{k_1}+v_{s2}x_{k_2}+v_o\sum_{i=1}^{m_2}x_{s_2+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right) \qquad (54)
$$

with

$$\begin{array}{llllll} v_{s1}=1 & , & v_{s2}=1 & \text{and} & v_o=1 & \text{for the low pass} \\ v_{s1}=m_1+m_c+1 & , & v_{s2}=m_2+m_c+1 & \text{and} & v_o=-1 & \text{for the high pass} \end{array}$$

The samples within the sums are shifted by the indices s_1, s_2 and s_c respectively. They are separated in that way that the samples of different sums are independent.

For the ACF expectation follows

$$\begin{aligned} E\{R'_k\} &= E\{y_{k_1}y_{k_2}\} \qquad (55) \\ &= E\left\{\frac{1}{(m_1+m_c+2)(m_2+m_c+2)}\left(v_{s1}x_{k_1}+v_o x_{k_2}+v_o\sum_{i=1}^{m_1}x_{s_1+i}\right.\right. \\ &\quad \left.\left.+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right)\left(v_o x_{k_1}+v_{s2}x_{k_2}+v_o\sum_{i=1}^{m_2}x_{s_2+i}+v_o\sum_{i=1}^{m_c}x_{s_c+i}\right)\right\} \\ &= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_c=0}^{\infty}\frac{p_m(m_1;\tau)p_m(m_2;\tau)p_m(m_c;2B-\tau)}{(m_1+m_c+2)(m_2+m_c+2)}E\left\{(v_{s1}v_{s2}+v_o^2)x_{k_1}x_{k_2}\right. \\ &\quad +v_o(v_{s1}x_{k_1}^2+v_{s2}x_{k_2}^2)+v_o(v_{s1}x_{k_1}+v_o x_{k_2})\sum_{i=1}^{m_2}x_{s_2+i}+v_o(v_{s2}x_{k_2} \\ &\quad +v_o x_{k_1})\sum_{i=1}^{m_1}x_{s_1+i}+v_o((v_{s1}+v_o)x_{k_1}+(v_{s2}+v_o)x_{k_2})\sum_{i=1}^{m_c}x_{s_c+i} \\ &\quad +v_o^2\left(\sum_{i=1}^{m_1}x_{s_1+i}\right)\left(\sum_{i=1}^{m_2}x_{s_2+i}\right)+v_o^2\left(\sum_{i=1}^{m_1}x_{s_1+i}+\sum_{i=1}^{m_2}x_{s_2+i}\right)\left(\sum_{i=1}^{m_c}x_{s_c+i}\right) \\ &\quad \left.+v_o^2\left(\sum_{i=1}^{m_c}x_{s_c+i}\right)^2\right\} \qquad (56) \end{aligned}$$

The m_1, m_2 and m_c independent samples are equally distributed in the time interval $[\tau_1-B;\tau_2-B]$, $[\tau_1+B;\tau_2+B]$ and $[\tau_2-B;\tau_1+B]$ with the probability density $\frac{m_1}{\tau}$, $\frac{m_2}{\tau}$ and $\frac{m_c}{2B-\tau}$ respectively. The products of the sums are independent, excepting the last term. The square of the sum has to be written like in equation (51) with products of independent samples.

Therefore, from equation (56) follows

$$\begin{aligned} E\{R'_k\} &= \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\sum_{m_c=0}^{\infty}\frac{p_m(m_1;\tau)p_m(m_2;\tau)p_m(m_c;2B-\tau)}{(m_1+m_c+1)(m_2+m_c+1)} \\ &\quad \left((v_{s1}v_{s2}+v_o^2)R(\tau)+v_o(v_{s1}+v_{s2})R(0)+\frac{v_{s1}v_o m_2}{\tau}\int_B^{\tau+B}R(t)\,\mathrm{d}t\right. \\ &\quad +\frac{v_o^2 m_2}{\tau}\int_{B-\tau}^{B}R(t)\,\mathrm{d}t+\frac{v_{s2}v_o m_1}{\tau}\int_B^{\tau+B}R(t)\,\mathrm{d}t+\frac{v_o^2 m_1}{\tau}\int_{-B}^{\tau-B}R(t)\,\mathrm{d}t \\ &\quad +\frac{v_o(v_{s1}+v_o)m_c}{2B-\tau}\int_{\tau-B}^{B}R(t)\,\mathrm{d}t+\frac{v_o(v_{s2}+v_o)m_c}{2B-\tau}\int_{\tau-B}^{B}R(t)\,\mathrm{d}t \\ &\quad \left.+\frac{v_o^2 m_1 m_2}{\tau^2}\int_{-B}^{\tau-B}\int_B^{\tau+B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1+\frac{v_o^2 m_c m_1}{\tau(2B-\tau)}\int_{-B}^{\tau-B}\int_{\tau-B}^{B}R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\right. \end{aligned}$$

$$
\begin{aligned}
&+\frac{v_o^2 m_c m_2}{\tau(2B-\tau)} \int_{\tau-B}^{B} \int_{B}^{\tau+B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 + v_o^2 m_c R(0) \\
&+\frac{v_o^2 m_c (m_c-1)}{(2B-\tau)^2} \int_{\tau-B}^{B} \int_{\tau-B}^{B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \Bigg) \\
= \; & \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_c=0}^{\infty} \frac{p_m(m_1;\tau) p_m(m_2;\tau) p_m(m_c;2B-\tau)}{(m_1+m_c+1)(m_2+m_c+1)} \\
&\big((v_{s1}v_{s2}+v_o^2)R(\tau) + v_o(v_{s1}+v_{s2}+v_o m_c)R(0) \\
&+\frac{v_o(v_{s1}m_2+v_{s2}m_1)}{\tau} \int_{B}^{\tau+B} R(t)\,\mathrm{d}t + \frac{v_o^2(m_1+m_2)}{\tau} \int_{-B}^{\tau-B} R(t)\,\mathrm{d}t \\
&+\frac{v_o m_c (v_{s1}+v_{s2}+2v_o)}{2B-\tau} \int_{\tau-B}^{B} R(t)\,\mathrm{d}t + \frac{v_o^2 m_1 m_2}{\tau^2} \int_{-B}^{\tau-B} \int_{B}^{\tau+B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\
&+\frac{v_o^2 m_c (m_1+m_2)}{\tau(2B-\tau)} \int_{-B}^{\tau-B} \int_{\tau-B}^{B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\
&+\frac{v_o^2 m_c (m_c-1)}{(2B-\tau)^2} \int_{\tau-B}^{B} \int_{\tau-B}^{B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \Bigg)
\end{aligned} \tag{57}
$$

4. The time windows of τ_1 and τ_2 are overlapping completely. Nevertheless, using the slot-correlation without selfproducts the times τ_1 and τ_2 and hence the corresponding samples x_{k_1} and x_{k_2} have a small deviation ($\tau = 0$, figure 7d). The sample x_{k_1} is a member of the local group at τ_2 and vice versa. There are m_c independent samples within the interval $[\tau_1 - B = \tau_2 - B; \tau_1 + B = \tau_2 + B]$. The filtered samples are

$$
y_{k_1} = \frac{1}{m_c+2} \left(v_s x_{k_1} + v_o x_{k_2} + v_o \sum_{i=1}^{m_c} x_{s_c+i} \right) \tag{58}
$$

$$
y_{k_2} = \frac{1}{m_c+2} \left(v_o x_{k_1} + v_s x_{k_2} + v_o \sum_{i=1}^{m_c} x_{s_c+i} \right) \tag{59}
$$

with

$v_s = 1$	and	$v_o = 1$	for the low pass
$v_s = m_c + 1$	and	$v_o = -1$	for the high pass

For the ACF expectation follows

$$
\begin{aligned}
E\{R'_k\} &= E\{y_{k_1} y_{k_2}\} \\
&= E\left\{ \frac{1}{(m_c+2)^2} \left(v_s x_{k_1} + v_o x_{k_2} + v_o \sum_{i=1}^{m_c} x_{s_c+i} \right) \right. \\
&\quad \left. \left(v_o x_{k_1} + v_s x_{k_2} + v_o \sum_{i=1}^{m_c} x_{s_c+i} \right) \right\} \\
&= \sum_{m_c=0}^{\infty} \frac{p_m(m_c;2B)}{(m_c+2)^2} E\left\{ (v_s^2+v_o^2) x_{k_1} x_{k_2} + v_o v_s (x_{k_1}^2 + x_{k_2}^2) \right.
\end{aligned} \tag{60}
$$

$$+v_o(v_s+v_o)(x_{k_1}+x_{k_2})\sum_{i=1}^{m_c} x_{s_c+i} + v_o^2\left(\sum_{i=1}^{m_c} x_{s_c+i}\right)^2\Bigg\} \tag{61}$$

The m_c independent samples are equally distributed in the time interval $[\tau_1 - B = \tau_2 - B; \tau_1 + B = \tau_2 + B]$ with the probability density $\frac{m_c}{2B}$. The square of the sum has to be written like in equation (51) with products of independent samples.

Therefore, from equation (61) follows

$$\begin{aligned}
E\{R'_k\} &= \sum_{m_c=0}^{\infty} \frac{p_m(m_c; 2B)}{(m_c+1)^2}\left((v_s^2+v_o^2)R(0) + 2v_o v_s R(0) + \frac{2v_o(v_s+v_o)m_c}{2B}\int_{-B}^{B} R(t)\,\mathrm{d}t\right.\\
&\quad \left. + v_o^2 m_c R(0) + \frac{v_o^2 m_c(m_c-1)}{(2B)^2}\int_{-B}^{B}\int_{-B}^{B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\right)\\
&= \sum_{m_c=0}^{\infty} \frac{p_m(m_c; 2B)}{(m_c+1)^2}\left(((v_s+v_o)^2 + v_o^2 m_c)R(0) + \frac{2v_o m_c(v_s+v_o)}{2B}\int_{-B}^{B} R(t)\,\mathrm{d}t\right.\\
&\quad \left. + \frac{v_o^2 m_c(m_c-1)}{(2B)^2}\int_{-B}^{B}\int_{-B}^{B} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1\right)
\end{aligned} \tag{62}$$

There are two types of integrals:

$$\int_a^b R(t)\,\mathrm{d}t \qquad \text{and} \qquad \int_{a_1}^{b_1}\int_{a_2}^{b_2} R(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \tag{63}$$

Presuming that $b_1 > a_1$ and $b_2 > a_2$ the second integral can be calculated through

$$\int_{a_c}^{b_c} v_c R(t)\,\mathrm{d}t \tag{64}$$

with

$$\begin{aligned}
a_c &= a_2 - b_1 && (65)\\
b_c &= b_2 - a_1 && (66)\\
v_c &= \min(v_m; v_t) && (67)\\
v_m &= \min(b_1 - a_1; b_2 - a_2) && (68)\\
v_t &= \min(t - a_c; b_c - t) && (69)
\end{aligned}$$

Both can be calculated using numerical summation. Presuming that τ and B and hence also a and b have the temporal resolution $\Delta\tau$ with $b > a$, $b_1 > a_1$ and $b_2 > a_2$, the integrals can be approximated through

$$\int_a^b R(t)\,\mathrm{d}t \approx \Delta\tau \sum_{i=a/\Delta\tau}^{b/\Delta\tau} v_{Ai} R_{|i|} \tag{70}$$

$$\int_{a_c}^{b_c} v_c R(t)\,\mathrm{d}t \approx \Delta\tau \sum_{i=a_c/\Delta\tau}^{b_c/\Delta\tau} v_{Bi} R_{|i|} \tag{71}$$

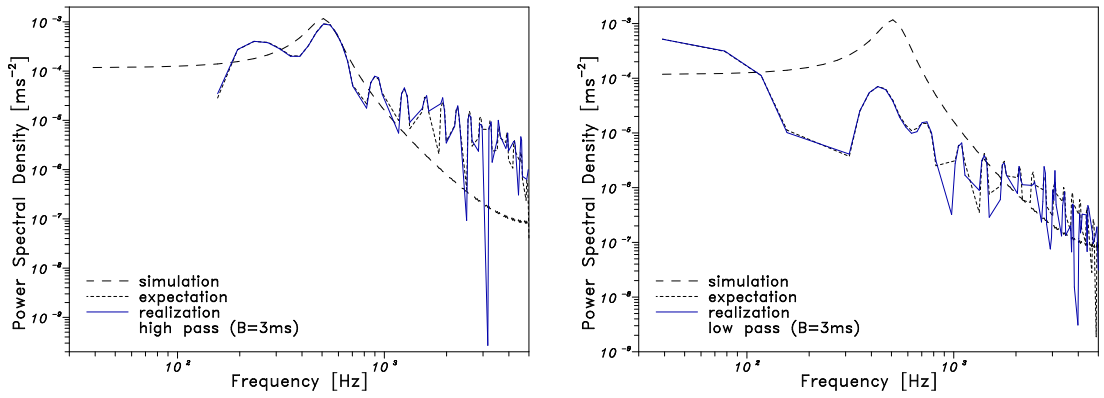

Figure 8: The PSD of a simulated LDA data set, filtered symmetrically with a constant time window

To minimize edge artefacts the pre-factors

$$v_{Ai} = \begin{cases} 1/2 & \text{for } i \in [a/\Delta\tau; b/\Delta\tau] \\ 1 & \text{otherwise} \end{cases} \tag{72}$$

$$v_{Bi} = \begin{cases} \Delta\tau/8 & \text{for } i \in [a_c/\Delta\tau; b_c/\Delta\tau] \\ v_m - \Delta\tau/8 & \text{for } i \in [(a_c+v_m)/\Delta\tau; (b_c-v_m)/\Delta\tau] \text{ and } (a_c+v_m) \neq (b_c-v_m) \\ v_m - \Delta\tau/4 & \text{for } i \in [(a_c+v_m)/\Delta\tau; (b_c-v_m)/\Delta\tau] \text{ and } (a_c+v_m) = (b_c-v_m) \\ v_t & \text{otherwise} \end{cases} \tag{73}$$

Further edge artefacts occure for $\tau = B$. This is the border between the time regimes 2 and 3. Using the slotcorrelation for the ACF estimation, both time regimes occure in this slot. Therefore the contributions of both time regimes have to be calculated and averaged to a final result. The other transitions are uncomplicated.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 8 shows the results of the computer simulation. It shows a good similarity of the calculated PSD expectation and the PSD estimation from filtered data.

3.2 Asymmetric Filter

3.2.1 Constant Number of Samples

The derivation of the filter characteristic for asymmetric filtering with a constant number of samples correspond to that of the symmetric filter with only one time case. Furthermore, the number of independent samples between two points in time is not needed. The derivations are similar for the high pass and the low pass. Only two different coefficients are necessary:

$$\begin{array}{llll} v_s = 1 & \text{and} & v_o = 1 & \text{for the low pass} \\ v_s = M & \text{and} & v_o = -1 & \text{for the high pass} \end{array}$$

The filtered samples are

$$y_i = \frac{1}{2M+1}\left[v_s x_i + v_o\left(\sum_{j=-M}^{-1} x_{i+j} + \sum_{j=1}^{M} x_{i+j}\right)\right] \tag{74}$$

To derive the expectation of R'_k for a given time lag $\tau = k\Delta\tau$ with the temporal resolution $\Delta\tau$ the filtered samples at two different points at the time $\tau_1 = k_1\Delta\tau$ and $\tau_2 = k_2\Delta\tau$ (see figure 1b) are of interest.

$$y_{k_1} = \frac{1}{M+1}\left(v_s x_{k_1} + v_o \sum_{j=-M}^{-1} x_{k_1+j}\right) \tag{75}$$

$$y_{k_2} = \frac{1}{M+1}\left(v_s x_{k_2} + v_o \sum_{j=1}^{M} x_{k_2+j}\right) \tag{76}$$

Because the samples in both sums are completely independent, for the expectation of the ACF follows

$$\begin{aligned} E\{R'_k\} &= E\{y_{k_1} y_{k_2}\} \qquad (77)\\ &= \frac{1}{(M+1)^2} E\left\{\left(v_s x_{k_1} + v_o \sum_{j=-M}^{-1} x_{k_1+j}\right)\left(v_s x_{k_2} + v_o \sum_{j=1}^{M} x_{k_2+j}\right)\right\} \\ &= \frac{1}{(M+1)^2} E\left\{v_s^2 x_{k_1} x_{k_2} + v_s v_o x_{k_1} \sum_{j=1}^{M} x_{k_2+j} + v_s v_o x_{k_2} \sum_{j=-M}^{-1} x_{k_1+j}\right. \\ &\quad \left. + v_o^2 \left(\sum_{j=1}^{M} x_{k_2+j}\right)\left(\sum_{j=-M}^{-1} x_{k_1+j}\right)\right\} \\ &= \frac{1}{(M+1)^2}\left[v_s^2 R(\tau) + 2 v_s v_o \int_0^\infty p_g(t) R(\tau+t)\,\mathrm{d}t\right. \\ &\quad \left. + v_o^2 \int_0^\infty \int_0^\infty p_g(t_1) p_g(t_2) R(\tau + t_1 + t_2)\,\mathrm{d}t_2\,\mathrm{d}t_1\right] \qquad (78) \end{aligned}$$

with the probability density

$$p_g(t) = \begin{cases} \dot{n} & \text{for } t = 0 \\ \dot{n} \sum_{i=0}^{M-1} \frac{(t\dot{n})^i}{i!} e^{-t\dot{n}} & \text{otherwise} \end{cases} \tag{79}$$

of the presence of a sample at the time t away from τ_1 or τ_2 and being a member of the corresponding local group.

The integrals have to be calculated numerically with the time resolution $\Delta\tau$ of the ACF. Equation (78) becomes

$$\begin{aligned} E\{R'_k\} &= \frac{1}{(M+1)^2}\left[v_s^2 R_k + 2 v_s v_o \Delta\tau \sum_{i=0}^{\infty} v_i p_g(i\Delta\tau) R_{k+i}\right. \\ &\quad \left. + v_o^2 (\Delta\tau)^2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} v_i v_j p_g(i\Delta\tau) p_g(j\Delta\tau) R_{k+i+j}\right] \qquad (80) \end{aligned}$$

with the pre-factors

$$v_i = \begin{cases} 0.5 & \text{for } i = 0 \\ 1 & \text{otherwise} \end{cases} \tag{81}$$

$$v_j = \begin{cases} 0.5 & \text{for } j = 0 \\ 1 & \text{otherwise} \end{cases} \tag{82}$$

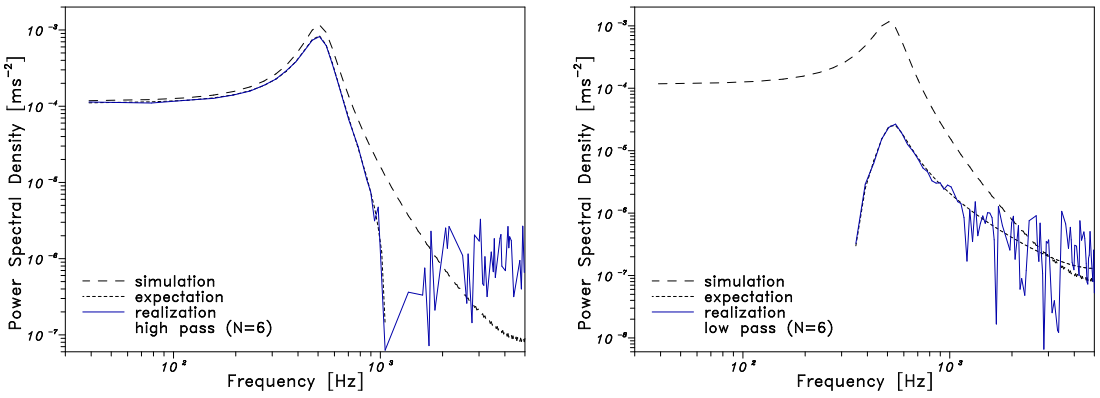

Figure 9: The PSD of a simulated LDA data set, filtered asymmetrically with a constant number of samples

to reduce the edge artefacts.

An important advantage of this filter type is the linearity of the transformation from the true ACF R to the expected ACF of the filtered data set R'. Therefore, the equation (80) has to be calculated only once for $k = 0$ leading to a discrete filter vector f_k. The expectation of the ACF for the filtered data becomes

$$E\{R'_k\} = f_k * R_k \tag{83}$$

with the convolution operator $*$.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 9 shows the results of the computer simulation. It shows a good similarity of the calculated PSD expectation and the PSD estimation from filtered data.

3.2.2 Constant Time

The derivation of the filter characteristic for asymmetric filtering with a constant time window correspond to that of the symmetric filter with only one time regime. For the derivation of the ACF expectation $E\{R'_k\}$ of filtered data sets the interaction of two points $\tau_1 = k_1\Delta\tau$ and $\tau_2 = k_2\Delta\tau$ (see figure 1b) is of interest.

The time windows of τ_1 and τ_2 are non-overlapping. There are m_1 independent samples within the time window of τ_1 and m_2 independent samples within the time window of τ_2. The filtered samples are

$$y_{k_1} = \frac{1}{m_1+1}\left(v_{s1}x_{k_1} + v_o\sum_{i=1}^{m_1} x_{k_1-i}\right) \tag{84}$$

$$y_{k_2} = \frac{1}{m_2+1}\left(v_{s2}x_{k_2} + v_o\sum_{i=1}^{m_2} x_{k_2+i}\right) \tag{85}$$

with

$$\begin{array}{llllll} v_{s1} = 1 & , & v_{s2} = 1 & \text{and} & v_o = 1 & \text{for the low pass} \\ v_{s1} = m_1 & , & v_{s2} = m_2 & \text{and} & v_o = -1 & \text{for the high pass} \end{array}$$

The samples within the sums are independent.

For the ACF expectation follows

$$
\begin{aligned}
E\{R'_k\} &= E\{y_{k_1} y_{k_2}\} && (86)\\
&= E\left\{ \frac{1}{(m_1+1)(m_2+1)} \left(v_{s1} x_{k_1} + v_o \sum_{i=1}^{m_1} x_{k_1-i} \right) \left(v_{s2} x_{k_2} + v_o \sum_{i=1}^{m_2} x_{k_2+i} \right) \right\} \\
&= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{p_m(m_1;B) p_m(m_2;B)}{(m_1+1)(m_2+1)} E\left\{ v_{s1} v_{s2} x_{k_1} x_{k_2} + v_{s1} v_o x_{k_1} \sum_{i=1}^{m_2} x_{k_2+i} \right. \\
&\quad \left. + v_{s2} v_o x_{k_2} \sum_{i=1}^{m_1} x_{k_1-i} + v_o^2 \left(\sum_{i=1}^{m_1} x_{k_1-i} \right) \left(\sum_{i=1}^{m_2} x_{k_2+i} \right) \right\} && (87)
\end{aligned}
$$

with the probability $p_m(m;\Delta t)$ of m independent samples within the interval Δt (equation 36). The m_1 and m_2 independent samples are equally distributed in the time interval $[\tau_1 - B; \tau_1]$ and $[\tau_2; \tau_2 + B]$ with the probability density $\frac{m_1}{B}$ and $\frac{m_2}{B}$ respectively.

Therefore, from equation (87) follows

$$
\begin{aligned}
E\{R'_k\} &= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{p_m(m_1;B) p_m(m_2;B)}{(m_1+1)(m_2+1)} \left(v_{s1} v_{s2} R(\tau) + \frac{v_{s1} v_o m_2}{B} \int_0^B R(\tau+t)\,\mathrm{d}t \right. \\
&\quad \left. + \frac{v_{s2} v_o m_1}{B} \int_0^B R(\tau+t)\,\mathrm{d}t + \frac{v_o^2 m_1 m_2}{B^2} \int_{-B}^{0} \int_0^B R(\tau + t_2 - t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \right) \\
&= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{p_m(m_1;B) p_m(m_2;B)}{(m_1+1)(m_2+1)} \left(v_{s1} v_{s2} R(\tau) + \frac{v_o (v_{s1} m_2 + v_{s2} m_1)}{B} \right. \\
&\quad \left. \int_0^B R(\tau+t)\,\mathrm{d}t + \frac{v_o^2 m_1 m_2}{B^2} \int_0^{2B} (B - |B-t|) R(\tau+t)\,\mathrm{d}t \right) && (88)
\end{aligned}
$$

The integrals can be calculated using numerical summation. Presuming that τ and B has the temporal resolution $\Delta\tau$ with $B > 0$ the integrals can be approximated through

$$
\int_0^B R(\tau+t)\,\mathrm{d}t \approx \Delta\tau \sum_{i=0}^{B/\Delta\tau} v_{Ai} R_i \qquad (89)
$$

$$
\int_0^{2B} (B - |B-t|) R(\tau+t)\,\mathrm{d}t \approx \Delta\tau \sum_{i=0}^{2B/\Delta\tau} v_{Bi} R_i \qquad (90)
$$

To minimize edge artefacts the pre-factors

$$
v_{Ai} = \begin{cases} 1/2 & \text{for } i \in [0; B/\Delta\tau] \\ 1 & \text{otherwise} \end{cases} \qquad (91)
$$

$$
v_{Bi} = \begin{cases} \Delta\tau/8 & \text{for } i \in [0/\Delta\tau; 2B/\Delta\tau] \\ B - \Delta\tau/4 & \text{for } i = B/\Delta\tau \\ B - |B - i\Delta\tau| & \text{otherwise} \end{cases} \qquad (92)
$$

are used.

This filter can also be described as a linear transform from the true ACF R to the expected ACF of the filtered data set R'. Therefore, the equation (88) has to be calculated only once for

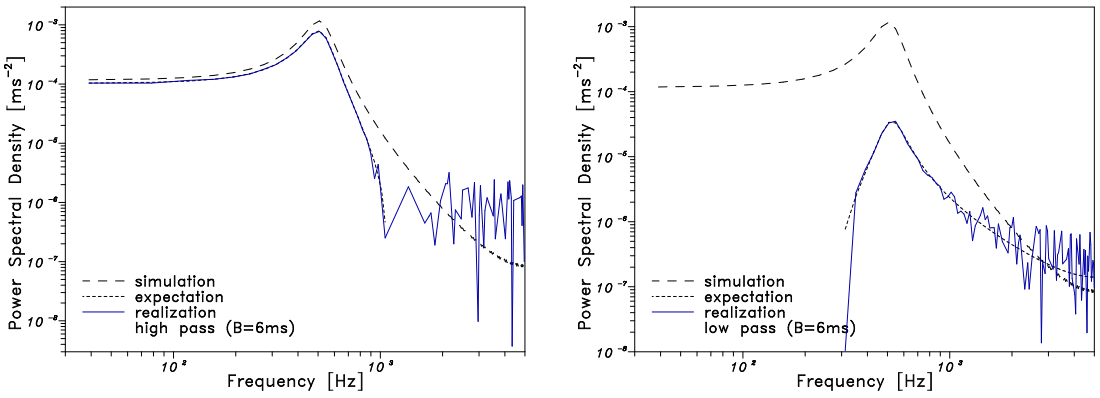

Figure 10: The PSD of a simulated LDA data set, filtered asymmetrically with a constant time window

$k = 0$ leading to a discrete filter vector f_k. The expectation of the ACF for the filtered data becomes

$$E\{R'_k\} = f_k * R_k \tag{93}$$

with the convolution operator $*$.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 10 shows the results of the computer simulation. It shows a good similarity of the calculated PSD expectation and the PSD estimation from filtered data.

4 Filter Correction

4.1 Symmetric Filter

The expectation of the ACF calculated from the filtered data $E\{R'\}$ (equations 38, 46, 52, 57 and 62) are linear functions of the true ACF R. The sums can be sorted by the arguments of the true ACF. That leads to a matrix F that transforms the true ACF into the expectation of the ACF calculated from the filtered data.

$$E\{R'\} = FR \tag{94}$$

The inverse matrix F^{-1} can be used to calculate a modified ACF R^* from the ACF R' estimated from measured and filtered data.

$$R^* = F^{-1}R' \tag{95}$$

The modified ACF is corrected for systematic errors connected to the filter.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (constant number of samples and constant time window, high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 11 shows the results of the computer simulation. It shows a good similarity of the corrected PSD from filtered data and the theoretic PSD derived from the simulation parameters.

4.2 Asymmetric Filter

Like described in section 3.2.1 and section 3.2.2 the asymmetric filters can be written as a linear transform from the true ACF R to the expected ACF of the filtered data set R'. Therefore, the

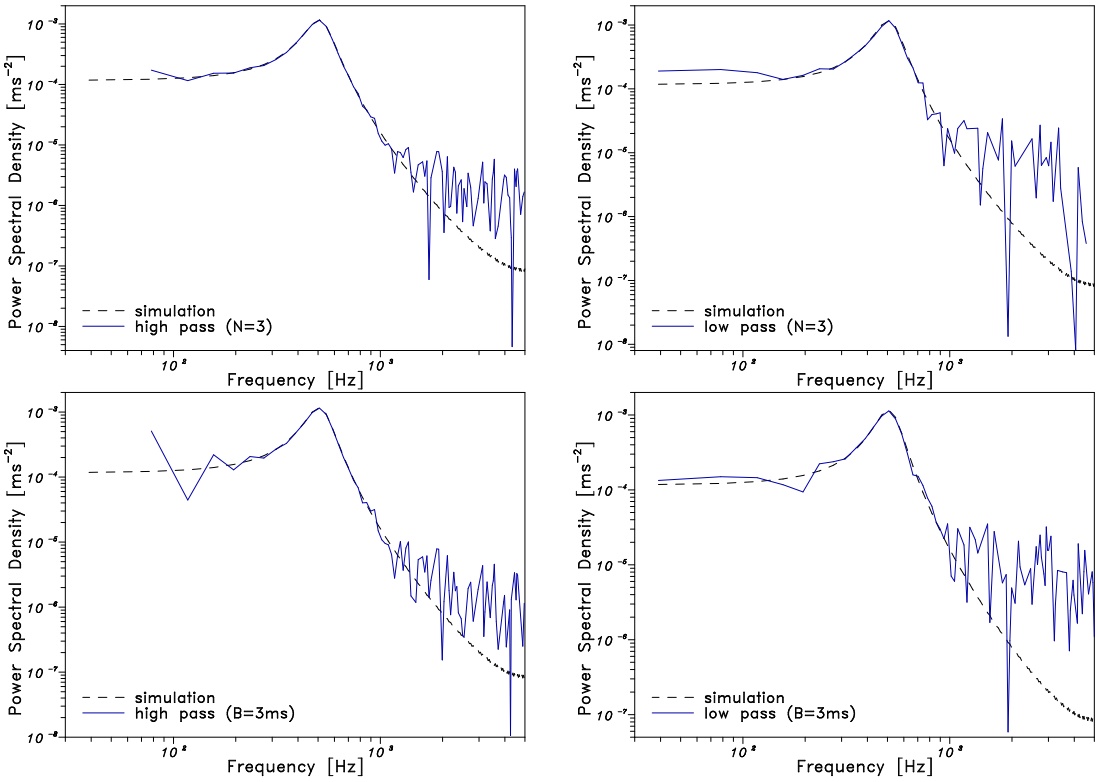

Figure 11: The corrected PSD of a simulated LDA data set, filtered symmetrically

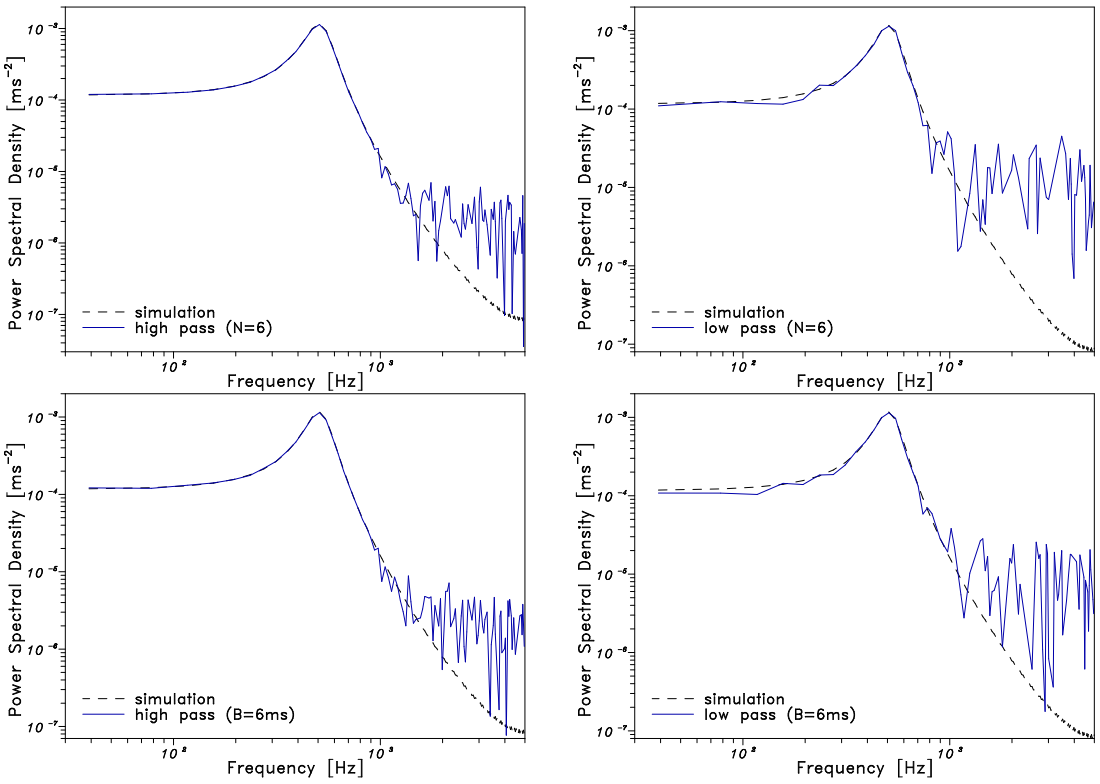

Figure 12: The corrected PSD of a simulated LDA data set, filtered asymmetrically

equation (88) has to be calculated only once for $k = 0$ leading to a discrete filter vector f_k. The expectation of the ACF for the filtered data becomes

$$E\{R'_k\} = f_k * R_k \tag{96}$$

with the convolution operator $*$.

An inverse Filter and hence a modified ACF estimation R^* with corrected filter errors can be found from the filter coefficients by the discrete and implicite deconvolution

$$R^*_k = \frac{1}{f_0}\left(R'_k - \sum_{i=1}^{K-k-1} f_j R^*_{k+j}\right) \tag{97}$$

with the maximum time lag $(K-1)\Delta\tau$. Because of the use of R^*_{k+j} for the derivation of R^*_k the calculation has to start with the maximum time lag R^*_{K-1}.

The theoretic ACF of the simulation has been used to calculate the expectation of the ACF from filtered (constant number of samples and constant time window, high pass and low pass) data sets. The ACF are transformed to the frequency domain using the Fourier transform. Figure 12 shows the results of the computer simulation. It shows a good similarity of the corrected PSD from filtered data and the theoretic PSD derived from the simulation parameters.

parameter	unit	value
number of realizations	–	1 000
model	–	AR2
		$(y_n = \phi_1 y_{n-1} + \phi_2 y_{n-2} + a_n)$
model parameters	–	$\phi_1 = 1.8$, $\phi_2 = -0.9$
sampling frequency (primary series)	kHz	10
mean velocity	ms^{-1}	0.0
variance	m^2s^{-2}	0.3
observation time	s	10
mean data rate (secondary series)	kHz	1.0
noise power	m^2s^{-2}	0.0
velocity bias	–	no
harmonic amplitude	ms^{-1}	0.0/1.0
harmonic frequency	Hz	1.0

Table 2: Simulation parameters for statistical investigations

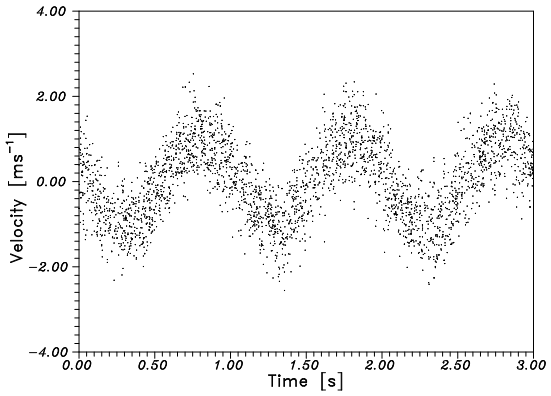

Figure 13: Simulated LDA data set with periodic component

5 Statistical Investigations

To get information about the expectation and the variability of ACF and PSD estimators with filter and correction techniques a series of 1 000 realizations of LDA data sets were simulated (table 2). To simulate the usual task of LDA data filtering, additionally to the normal data sets (figure 5) another type of data was generated by adding a harmonic component of given frequency and amplitude and randomly chosen phase (figure 13). The 1 000 series of each type were processed by the different filter types (section 2) and corrected in the described way (section 4). The ACF of each data set was transformed into the PSD using the Fourier transform. The sets of PSDs were analysed statistically. For each data and filter type the expectation was estimated through the mean PSD and the variability through the estimation's variance.

Figure 14 shows the results for the symmetric filters (high pass and low pass; constant number of samples and constant time window) without a periodic component. Figure 15 shows the corresponding results for the data sets with the periodic component and figures 16 and 17 the results for the asymmetric filters.

For the data sets without a periodic component (figures 14 and 16) the systematic errors of the data filtering can be corrected completely for all filter types. But the estimation's variability is larger than the variability of the procedure without any filter. Therefore, it doesn't make

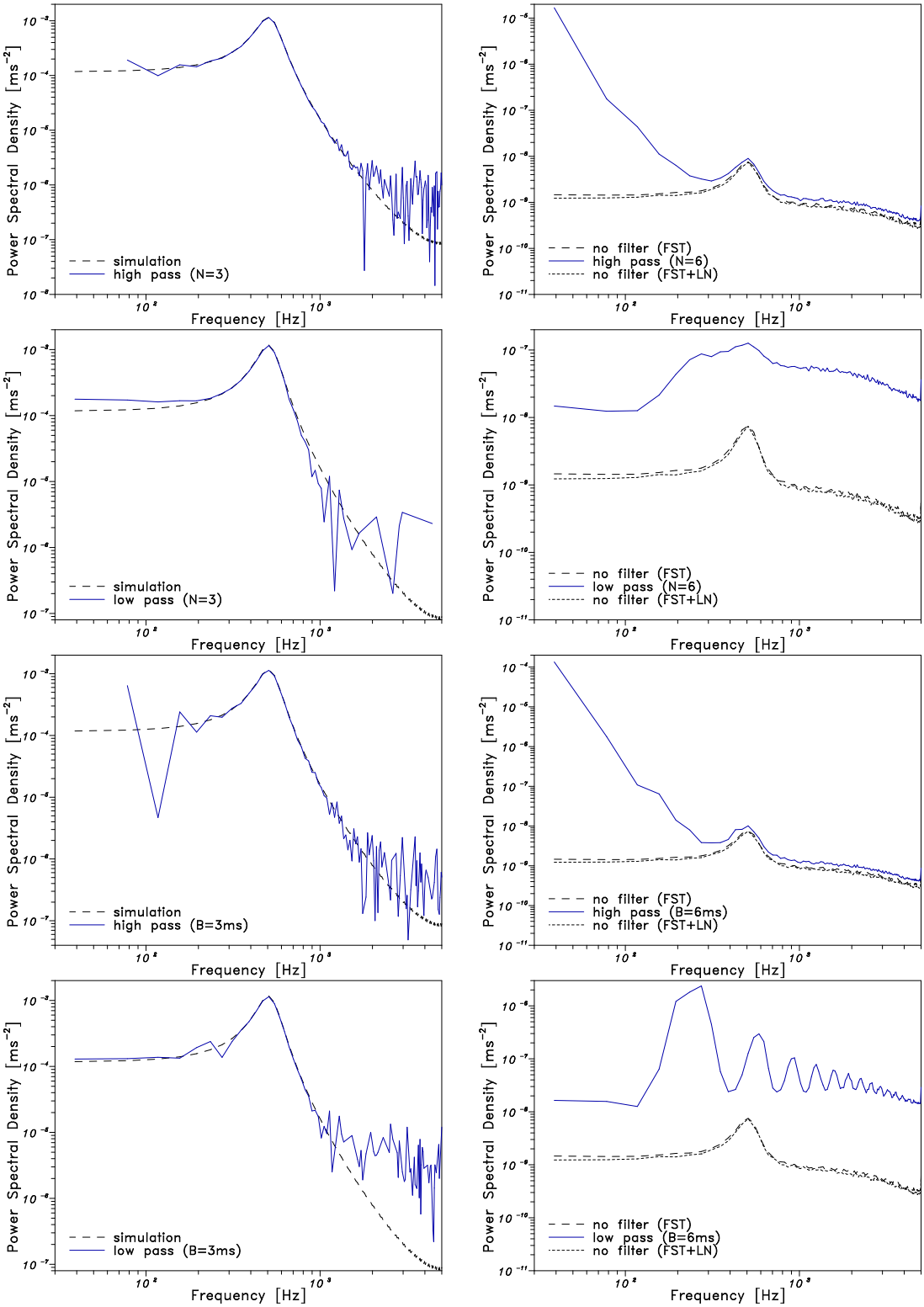

Figure 14: Empirical Expectation and estimation's variability for the symmetric filters without a periodic component

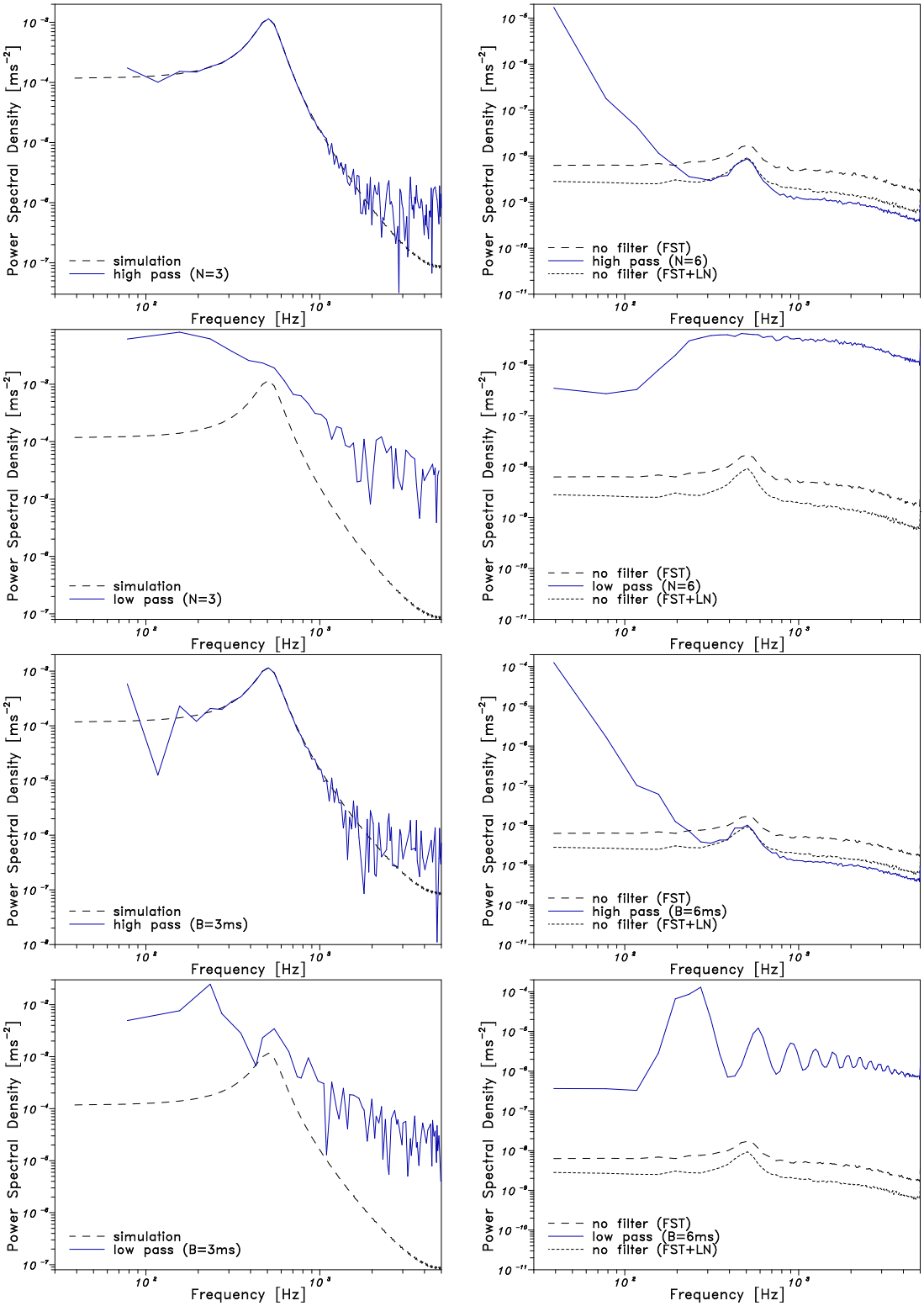

Figure 15: Empirical Expectation and estimation's variability for the symmetric filters with a periodic component

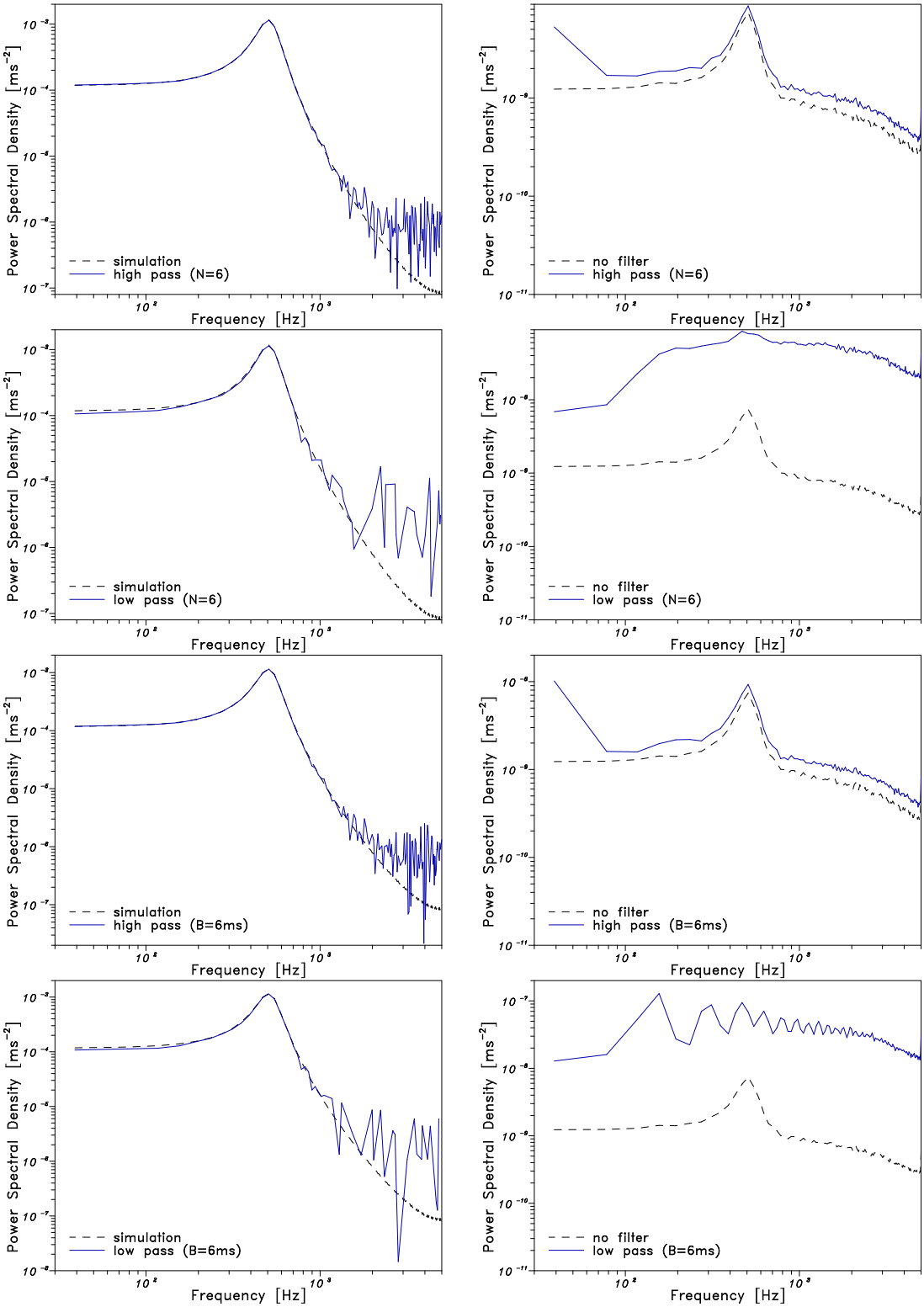

Figure 16: Empirical Expectation and estimation's variability for the asymmetric filters without a periodic component

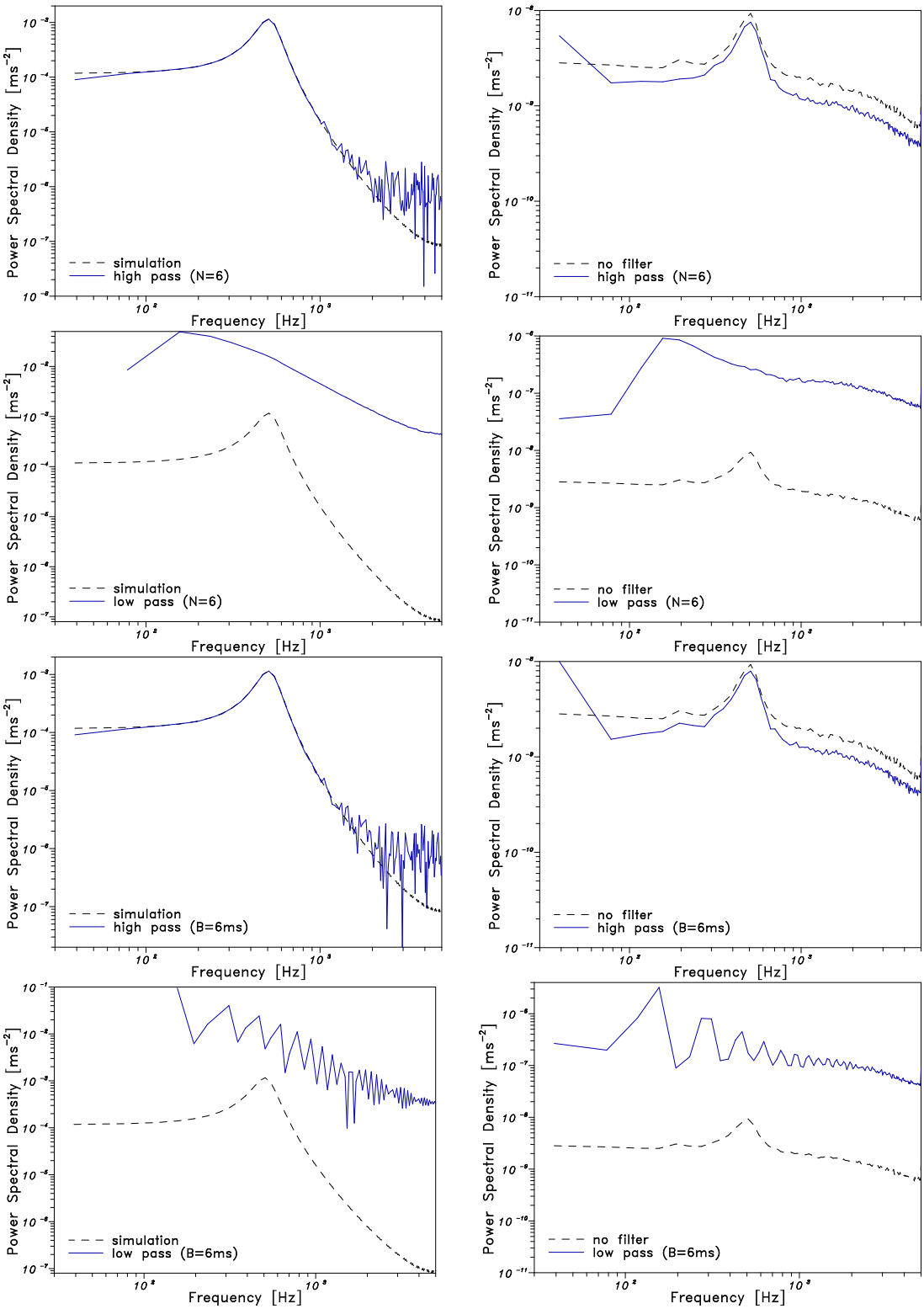

Figure 17: Empirical Expectation and estimation's variability for the asymmetric filters with a periodic component

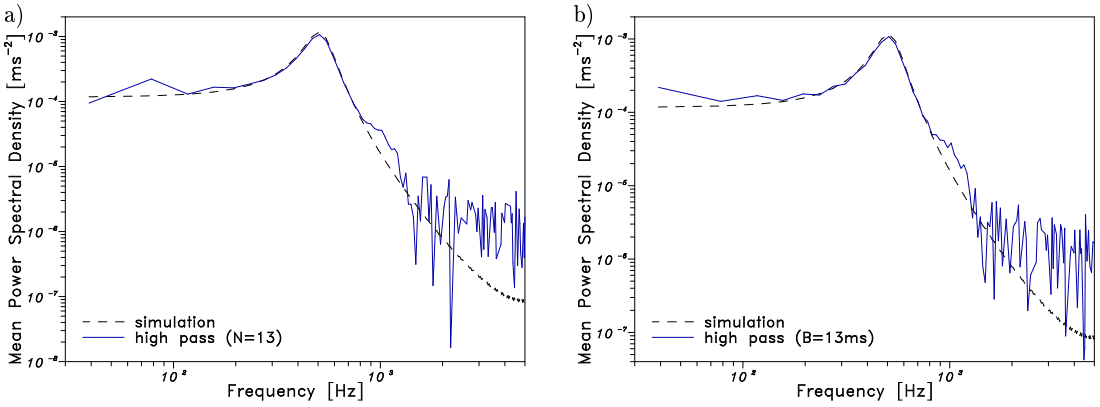

Figure 18: Systematic errors (the step at 1 kHz) for symmetric high pass filtering a) with a large constant number of samples and b) with a large constant time length (data set without a periodic component)

sense to use the filtering technique for data sets without significant power in the non-interesting low-frequency range.

For the data sets with a periodic component (figures 15 and 17) the systematic errors of the data filtering can be corrected only for the high pass filters. The estimates using the low pass filters are completely wrong. The course of that effect is the high power in large time scales that was not used in the derivation of the filter characteristics. The high pass filters don't show this effect, because they suppress just this frequency range. Furthermore, it can be seen that the variability of estimation is reduced for the high pass filters, but only for higher frequencies. The low frequencies are suppressed by the filters hence the coefficients for the filter correction are heavy in this range. This is more distinct for the symmetric filters because they suppress the power in the low-frequency range more heavy than the asymmetric filters. Therefore, the correction coefficients are more heavy for the symmetric filters and hence the variability is higher, especially for low frequencies. That makes the asymmetric high pass filtering with the corresponding correction preferable for this problematic type of LDA data. Besides, the derivation of the filter characteristics and the correction is significantly more simple for this filter type and the LN as well as the FST can be used to reduce the estimation's variability once more.

6 Recommendations

The filtering technique should be used only in the case of significant power in a non-interesting frequency range. To separate the high frequency part an asymmetric high pass filter should be used with the corresponding correction. It is equal whether a constant number of samples or a constant time window has been used. The ACF should be calculated with the slot correlation without self products. The asymmetric filtering with it's correction can be combined with the FST and the LN to reduce the estimation's variability.

The filtering techniques should not be used without an adequate correction because of systematic errors, even in the case of symmetric filtering with large numbers of samples or large time windows to calculate the local mean (figure 18).

To calculate the statistics of the low frequency range a reduced temporal resolution can be used. There are no aliasing errors (figure 19). Note that the variability of the spectral estimate is proportional to the number of slots and the time step $\Delta\tau$. A reduced temporal resolution with the same number of slots leads to an increased variability (figure 20). Furthermore, this leads to

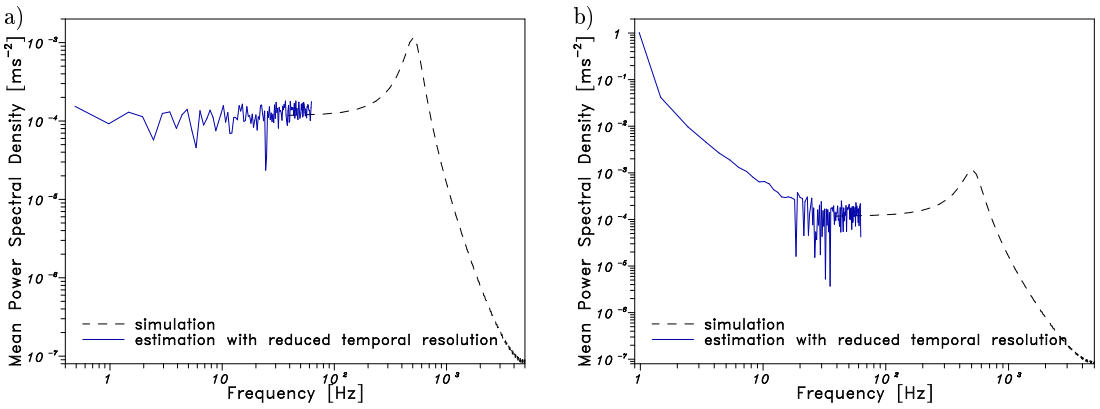

Figure 19: PSD with reduced temporal resolution a) without and b) with a periodic component of 1 Hz

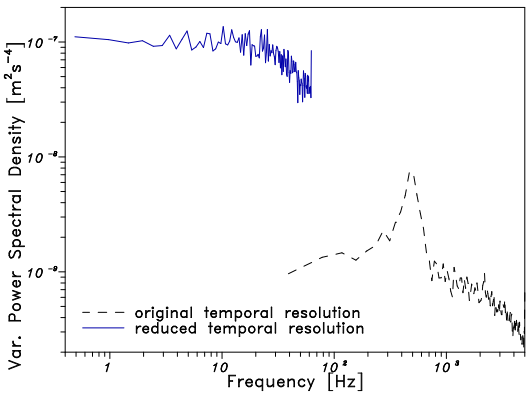

Figure 20: Variability of the PSD estimate with reduced temporal resolution (data set without a periodic component)

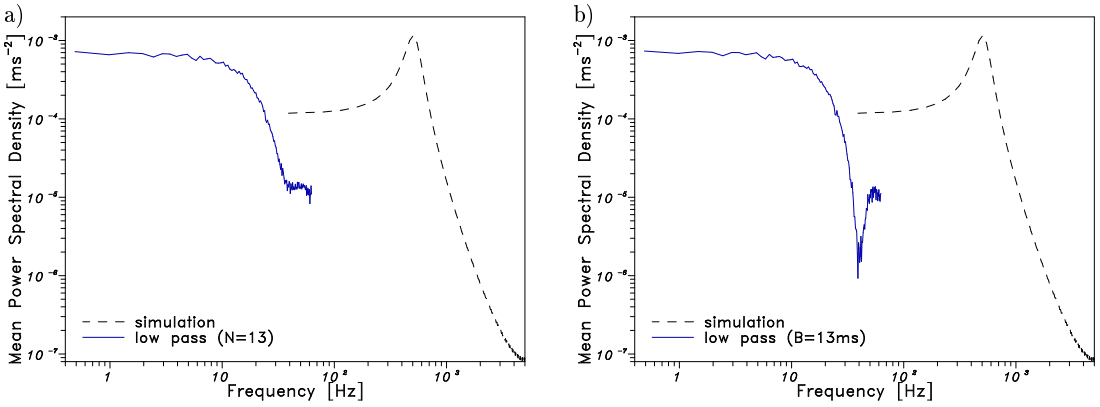

Figure 21: Systematic errors for symmetric low pass filtering a) with a constant number of samples and b) with a constant time length (data set without a periodic component)

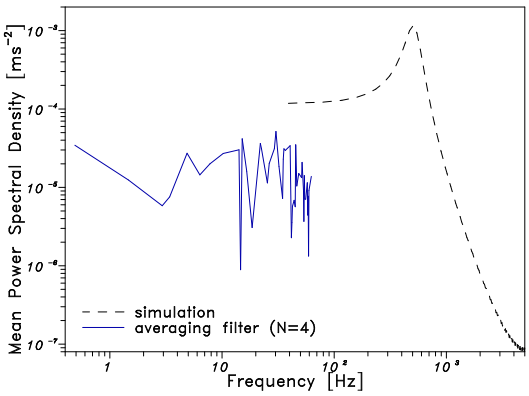

Figure 22: Systematic error for an averaging filter (data set without a periodic component)

an increased (to the power of two) number of calculations. Therefore it is necessary to reduce the number of samples.

The low pass filters should not be used for this. Without a filter correction there is a systematic error in the ACF or PSD of the filtered data (figure 21). It is not an aliasing error, but it is a complex dependence of the result on the true data statistics and the filter. The use of the low pass filters together with their correction leads to completely wrong results in the presence of a periodicity with a low frequency (figures 15 and 17). Furthermore, these filters do not lead to a reduced data set.

An averaging filter with

$$t_{AVi} = \frac{1}{M} \sum_{j=M(i-1)+1}^{Mi} t_j \tag{98}$$

$$y_{AVi} = \frac{1}{M} \sum_{j=M(i-1)+1}^{Mi} x_j \tag{99}$$

should never be used to reduce the number of data samples in an LDA data set, because of systematic errors (figure 22). Furthermore, the sampling statistics of the filtered data set is not LDA data like. That leads to problems with the slot correlation.

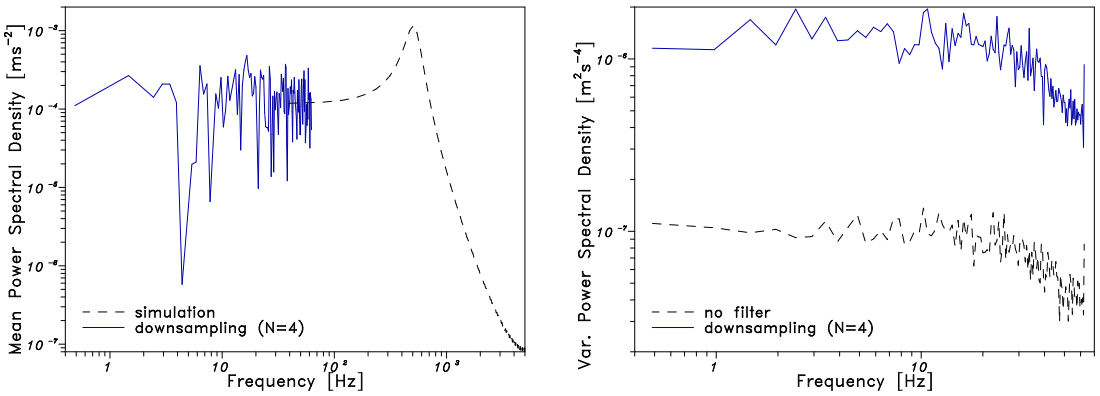

Figure 23: Expectation and variability of the downsampling procedure (data set without a periodic component)

A solution of that problem is the downsampling. For each original data sample an independent random number is taken from the interval $[0; 1)$. Only these samples are written to the new data set where this random number is less than $1/M$. That leads to a reduced number of samples but some information are lost like with a reduced data rate. The variability of following ACF estimations is increased (figure 23). But the new data set has an LDA conform sampling statistic, there are no problems with the slot correlation, even not with the FST or the LN, and there is the possibility of cascading several downsampling filters to expand the frequency range repeated.

7 Summery

The characteristics of LDA data filtering techniques, especially of the filter described in [8], could be derived. The systematic errors of this technique could be predicted and removed successfully. New, asymmetric filters were developed which lead to better results together with easier prediction and correction expressions. The proof could be furnished that the asymmetric high pass filtering and the corresponding correction are useful to reduce the variability of ACF or PSD estimations in the presence of a significant periodic component with a low frequency. For the statistical estimations with lower temporal resolution a bias free procedure was developed to reduce the number of data samples and hence the computational expenditure.

8 Postscriptum

In connection with investigations how processor delays influence the results of ACF and PSD estimations using the new filter algorithms the refined expression for the probability density (equation 79)

$$p_g(t) = \begin{cases} \dot{n} & \text{for } t = 0 \\ \dot{n} \sum_{i=0}^{M-1} \frac{((t-(i+1)t_0)\dot{n})^i}{i!} e^{-(t-(i+1)t_0)\dot{n}} & \text{otherwise} \end{cases} \tag{100}$$

was used with respect to the processor delay t_0. It presumes that a group of i samples needs at least an interarrival time of $(i+1)t_0$ between the preceding and the following sample. Because of the limited number of samples within a given time interval an additional norm factor is necessary in the case of using a constant time window (equations 36 and 86–88).

$$p_m(m, \Delta t) = \frac{\frac{(\dot{n}(\Delta t-(m+1)t_0))^m}{m!} e^{-\dot{n}(\Delta t-(m+1)t_0)}}{\sum_{i=0}^{\lfloor \Delta t/t_0 \rfloor} \frac{(\dot{n}(\Delta t-(i+1)t_0))^i}{i!} e^{-\dot{n}(\Delta t-(i+1)t_0)}} \tag{101}$$

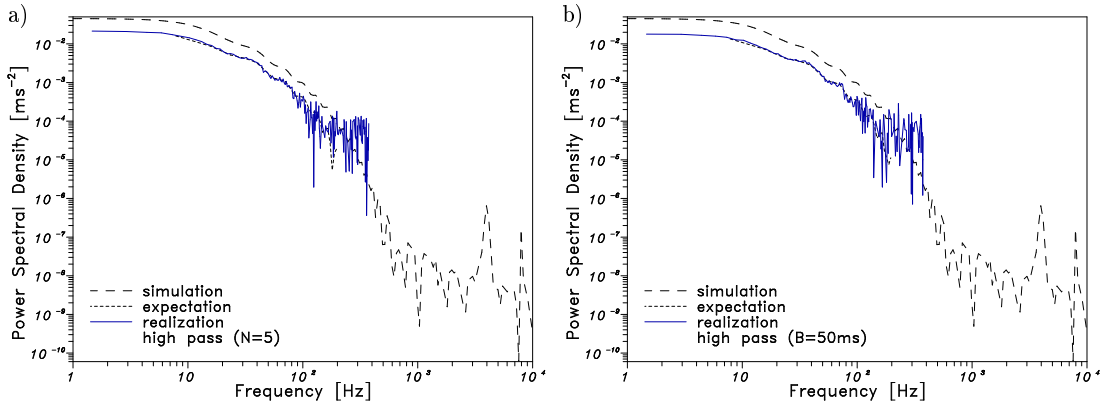

Figure 24: Prediction of the PSD of an asymmetricly filtered data set with processor delay a) with a constant number of samples and b) with a constant time window

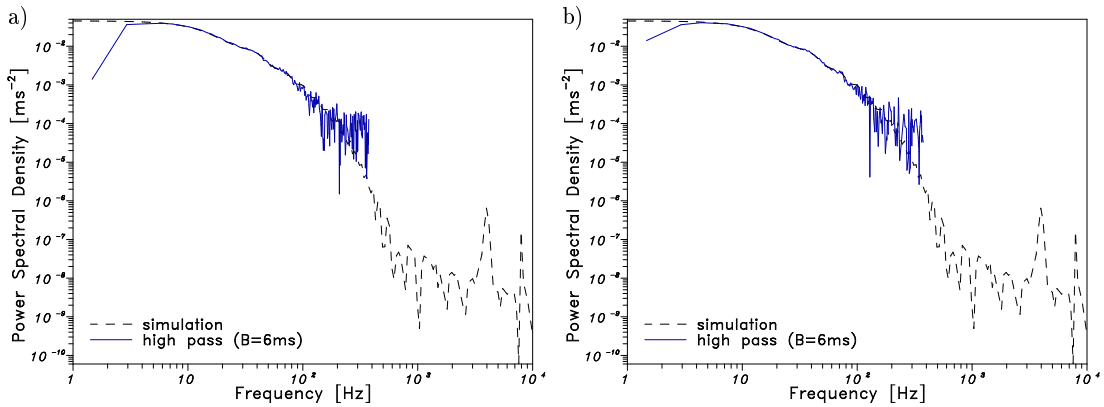

Figure 25: Mean of the corrected PSD of an asymmetricly filtered data set with processor delay a) with a constant number of samples and b) with a constant time window

The prediction of the PSD of a filtered data set (figure 24) shows a good correspondence with the estimation.

Figures 25 and 26 show the mean and the variance of the PSD estimation using the correction method. The corrected estimation has no significant deviations to the simulation excepting the lowest frequencies, where the estimation variance is very large.

This is a basic problem of the filter and correction algorithms. The aim of the filtering technique is to reduce the power in the low frequency range yielding to deviations. The aim of the correction algorithm is to correct these deviations yielding to a higher estimation variability. The combination of both algorithms brings out a better result compared to the unfiltered data only if the filter constants (M or B) are chosen in that way that the final variability becomes smaller. While the lowest frequencies of the PSD are not that important, the procedures are quite robust. But the estimation of ACF is very sensitive to the low frequency range of the PSD. The PSD in figure 25 yield to significant deviations with large scales in the ACF (figure 27). Note that this is not an effect of the processor delay, but it is a basic problem of the filter and correction technique.

The conclusions from these results are:

- The new filter and correction technique can be used to reduce the variability of the PSD estimation in the high frequency range for LDA data sets with a periodicity of significant

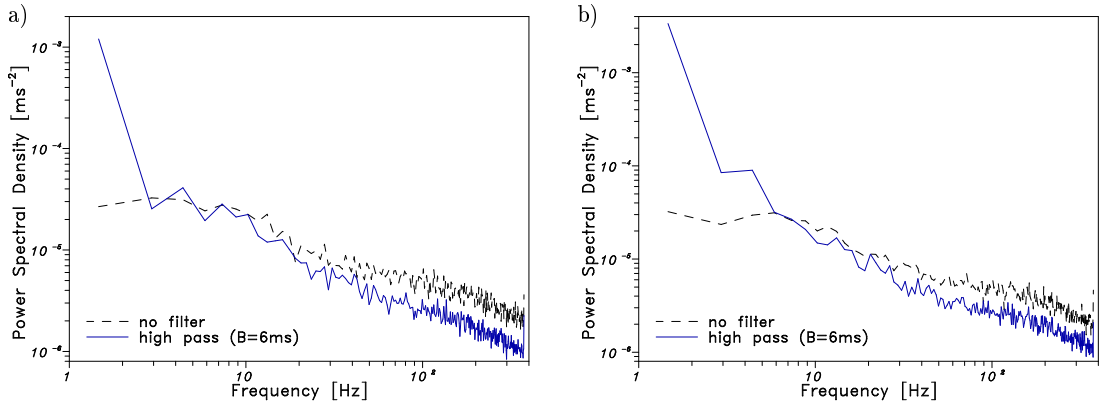

Figure 26: Variance of the corrected PSD of an asymmetricly filtered data set with processor delay a) with a constant number of samples and b) with a constant time window

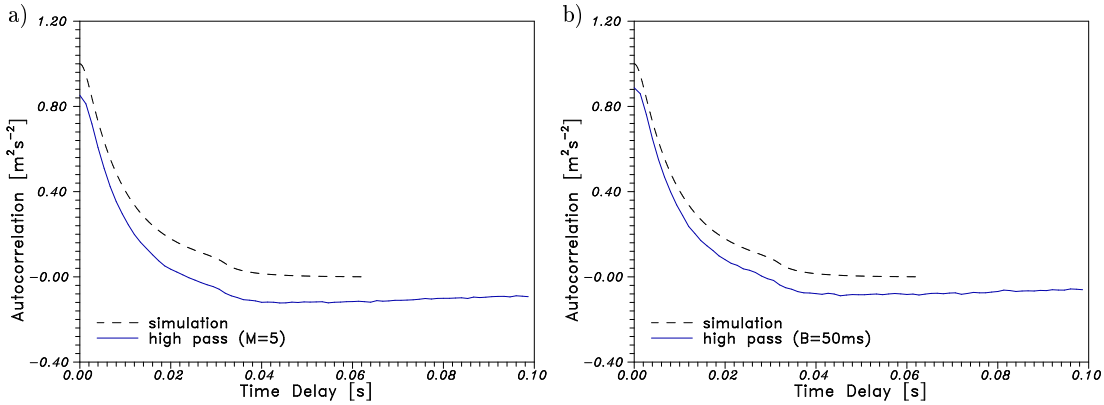

Figure 27: Mean of the corrected ACF of an asymmetricly filtered data set with processor delay a) with a constant number of samples and b) with a constant time window

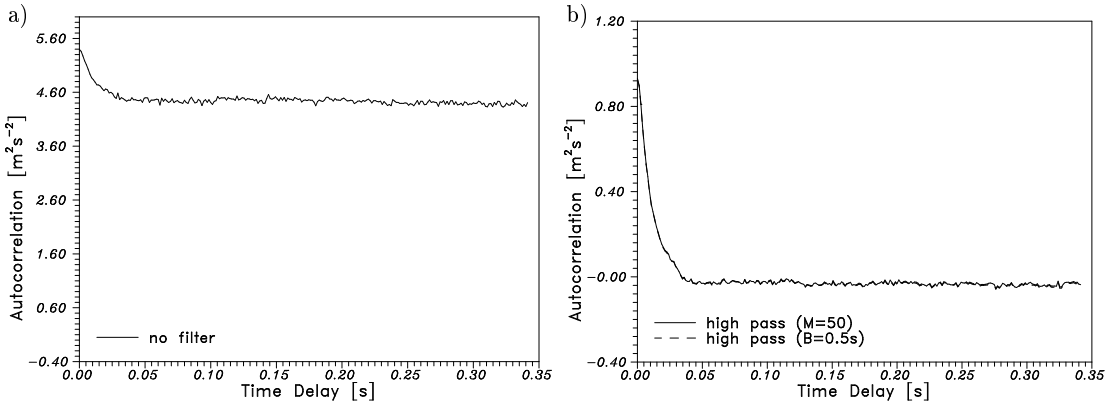

Figure 28: ACF estimation a) with unfiltered data and b) with large filter constants

power.

- For the estimation of the ACF the algorithm should not be used because of possible deviations with large scales. In that case the unfiltered data set (figure 28a) or a symmetric filter with large filter constants without a correction (figure 28b) should be prefered. The visible bend corresponds directly to the large scale periodicity.

Acknowledgment

The author wishes to acknowledge the financial support of the Deutsche Forschungsgemeinschaft through the grant No 373/1-1.

References

[1] R J Adrian and C S Yao. Power spectra of fluid velocities measured by laser Doppler velocimetry. *Exp. in Fluids*, 5:17–28, 1987.

[2] L H Benedict, H Nobach, and C Tropea. Benchmark tests for the estimation of power spectra from LDA signals. In *Proc. 9th Int. Symp. on Appl. of Laser Techn. to Fluid Mechanics*, Lisbon, Portugal, 1998. paper 32.6.

[3] W Fuchs, H Nobach, and C Tropea. Laser Doppler anemometry data simulation: Application to investigate the accuracy of statistical estimators. *AIAA Journal*, 32:1883–1889, 1994.

[4] H R E van Maanen, H Nobach, and L H Benedict. Improved estimator for the slotted autocorrelation function of randomly sampled LDA data. *Meas. Sci. Technol.*, 10(1):L4–L7, 1999.

[5] H R E van Maanen and M J Tummers. Estimation of the autocorrelation function of turbulent velocity fluctuations using the slotting technique with local normalization. In *Proc. 8th Int. Symp. on Appl. of Laser Techn. to Fluid Mechanics*, Lisbon, Portugal, 1996. paper 36.4.

[6] H Nobach, E Müller, and C Tropea. Correlation estimator for two-channel, non-coincidence laser-Doppler-anemometer. In *Proc. 9th Int. Symp. on Appl. of Laser Techn. to Fluid Mechanics*, Lisbon, Portugal, 1998. paper 32.1.

[7] H Nobach, E Müller, and C Tropea. Efficient estimation of power spectral density from laser Doppler anemometer data. *Experiments in Fluids*, 24:499–509, 1998.

[8] D Sree, S O Kjelgaard, and W L Sellers III. Spectral enhancement of randomly sampled signals by pre-filtering techniques. *Laser Anemometry — 1994: Advances and Applications*, pages 680–685, ASME 1994. FED-Vol. 191.

[9] M J Tummers, D M Passchier, and P A Aswatha Narayana. LDA measurements of time- and spatial correlation functions in an adverse pressure gradient wake. In *Proc. ASME/JSME Fluids Engineering and Laser Anemometry Conf.*, pages 347–355, Hilton Head Island, South Carolina, USA, 1995. FED-Vol. 229.