An interpolation method for autocorrelation and spectral density estimation from laser Doppler anemometry data

Holger Nobach

December 3, 2016

Abstract

The interpolation method for calculating the autocorrelation function and the spectrum from laser Doppler data is revisited and extended by recently developed processing steps.

1 Introduction

For calculating the correlation function or the power spectral density from randomly sampled data from laser Doppler velocity measurements, estimation procedures, which consider the specific characteristics of LDV data are required, namely the sampling of the flow velocity at random arrival times, the data noise and the correlation of the sampling rate and the instantaneous velocity. Much effort has been put onto autocorrelation and autospectral estimators following three different estimator classes, slot correlation, estimating a correlation function (correlogram) from the data [3, 27, 28, 9, 10, 13, 15, 19, 21, 22, 23, 24, 26], direct spectral estimators, estimating a spectrum (periodogram) directly from the randomly sampled data [3, 4, 5, 6, 7, 8, 16, 17, 29] and interpolation methods of the randomly sampled LDV data set yielding a continuous velocity over time, which then is re-sampled equidistantly with a given sampling frequency and processed by usual signal processing tools for equidistantly sampled data, including corrections of systematic errors [2, 11, 20, 25] and noise removal [18, 20]. A fourth processing method, using quantitized arrival times, has not achieved much attention so far. It has been used before, in [12], however, it is broadly available only after publication of [4], where it has been used to accelerate the direct estimation of auto-spectra in combination with a normalization of the appropriate correlation function.

2 The data set

In the following processing steps a set of irregularly sampled velocity data $u_i = u(t_i)$ at sampling times $t_i, i = 0 \dots N-1$ is assumed. Individual weights, e.g.

the particle's transit times, cannot be considered by the present interpolation method.

3 Interpolation Method

The interpolation method principally follows the procedure in [20], except for a few extensions/modifications, which are

1. An additional weighting factor is introduced, which can be set zero, if a large gap between measurements occurs, as has been seen in experimental data. The removal of data gaps by weighting is similar to the procedure introduced for the inter-arrival time or forwar-backward inter-arrival time weighting for the other estimation procedures. Note that this particular weighting factor for the interpolation method cannot be used as freely as for the other two estimation methods, only factors of one or zero are allowed here.

2. A model-free noise removal procedure as introduced in [18] has been implemented.

3. A normalization of the correlation function of the velocities by the correlation function of the sampling function, as it is inherently used with the slot correlation and as it has adapted to the direct estimation [4, 16], has been adapted also to the interpolation method as a means of correction of the influence of dead times of the measurement system.

4. Bessel's correction of the correlation estimate is added, which suppresses systematic deviations due the under-estimation of the velocity variance for short data sets, if the mean is estimated and removed from the data sets following [16].

Since these little modifications influence the entire estimation procedure, it is summarized here including the modifications and extensions.

3.1 Data pre-processing

The available data may be subdivided into blocks of a certain time duration T_B or the data may be obtained in blocks of a given record length. Due to the combination of Bessel's correction and the temporal limitation of the correlation function, both given below, the block duration can be chosen very flexible (compare [16]). It should be larger than the expected correlation interval of the flow and can be as large as the full data set. Since for the interpolation method, the computational costs increase with the square of the block length, too large a block duration will be computational costly.

The assumed data set $u_i = u(t_i)$ of the block duration T_B is interpolated using the sample-and-hold interpolation and re-sampled equidistantly with the frequency $F = {}^{1}\!/_{\Delta\tau}$, which defines the fundamental frequency of all derived statistical functions hereof.

To avoid the wrap-around error of the derived statistical functions, the interpolation is done for the duration of $2T_\mathrm{B}$, where only the duration $1T_\mathrm{B}$ gets measured data, where the above mentioned weighting factor is set to one. For the other duration of $1T_\mathrm{B}$ the above weighting factor is set to zero to identify the interpolated and re-sampled data as invalid or unknown. This is the pendent to zero padding of equidistantly sampled data for the case of randomly sampled data. An important detail to avoid systematic deviations is to interpolate the valid data for exactly the duration of $1T_\mathrm{B}$. For this purpose, the arrival time of the first data point t_0 is translated to $T_\mathrm{B} + t_0$ and the value u_{N-1} of the last sample in the data record is hold between the occurence at t_{N-1} until this points in time. This yields an interpolated data set of exactly the time duration of $1T_\mathrm{B}$.

Formally, the interpolation looks like

$$u_i' = u'(t_i) = u'(i\Delta\tau) = u_k \left\{ \begin{array}{l} \forall i : t_k \le i\Delta\tau < t_{k+1} \text{ for } k = 0 \ldots N-2 \\ \forall i : t_{N-1} \le i\Delta\tau < T_\mathrm{B} + t_0 \text{ for } k = N-1 \end{array} \right.$$

with $\Delta\tau = {}^{1}\!/_{F}$. Outside the interval $t_0 \le i\Delta\tau < T_\mathrm{B} + t_0$, all values are zero. With experimentally obtained data, gaps in the data stream have been identified, which significantly affect the derived statistical functions. Therefore, interpolated weighting factors are defined similar to the interpolated velocities as

$$w_i' = w'(t_i) = w'(i\Delta\tau) = w_k \left\{ \begin{array}{l} \forall i : t_k \le i\Delta\tau < t_{k+1} \text{ for } k = 0 \ldots N-2 \\ \forall i : t_{N-1} \le i\Delta\tau < T_\mathrm{B} + t_0 \text{ for } k = N-1 \end{array} \right.$$

For the interpolation method, the weights w_k are usually set to one. However, these weights can be used to suppress the gaps in the data stream by setting the weights to zero, if the inter-arrival time between two samples exceeds a certain limit. Good experience has been obtained with a maximum value of five times the mean inter-arrival time. Due to this one looses about 0.7 % of useful data, while the outliers of large inter-arrival times are suppressed effectively. The weights then read

$$w_k = \left\{ \begin{array}{l} \left. \left\{ \begin{array}{ll} 1 & \text{for } t_{k+1} - t_k < 5n \\ 0 & \text{otherwise} \end{array} \right. \right\} \text{ for } k = 0 \ldots N-2 \\ \left. \left\{ \begin{array}{ll} 1 & \text{for } T_\mathrm{B} + t_0 - t_{N-1} < 5n \\ 0 & \text{otherwise} \end{array} \right. \right\} \text{ for } k = N-1 \end{array} \right.$$

with the mean data rates

$$n = \frac{t_{N-1} - t_0}{N-1}$$

Other individual weighting, e.g. transit-time weighting, has not been realized for the interpolation method yet. Since the interpolation holds the values between samples, inherently a kind of an inter-arrival time weighting is realized. The introduction of another weighting scheme then would over-weight the samples. However, the weights introduced here, for the interpolation method, cannot be used as freely as for the other processing methods, while the use of the weighting values to suppress parts of the data stream is possible.

From the interpolated and equidistantly re-sampled data series $u_i', i = 0 \ldots [2T_\mathrm{B}F] - 1$ and the appropriate interpolated weigths w_i' one can calculate the block mean value as

$$\bar{u} = \frac{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} w_i' u_i'}{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} w_i'}$$

and remove the mean from the interpolated data to generate a mean free data block for the following calculations of the correlation function and the appropriate power spectral density.

3.2 Estimation of the initial correlation functions

From the interpolated data one can obtain the correlation functions of the weighted velocity and that of the weights either directly as

$$\begin{aligned}
R_u'(\tau_k) &= \frac{1}{T_\mathrm{B}F} \sum_{i=0}^{[2T_\mathrm{B}F]-1} \sum_{j=0}^{[2T_\mathrm{B}F]-1} w_i' u_i' w_j' u_j' \\
R_w'(\tau_k) &= \frac{1}{T_\mathrm{B}F} \sum_{i=0}^{[2T_\mathrm{B}F]-1} \sum_{j=0}^{[2T_\mathrm{B}F]-1} w_i' w_j'
\end{aligned}$$

or, with less computational costs, via the spectrum using the discrete Fouier transform (DFT)

$$\begin{aligned}
U'(f_j) &= \mathrm{DFT}\left\{w_i' u_i'\right\} = \sum_{i=0}^{[2T_\mathrm{B}F]-1} w_i' u_i' \mathrm{e}^{-2\pi \mathbb{i} f_j i \Delta\tau} \\
W'(f_j) &= \mathrm{DFT}\left\{w_i'\right\} = \sum_{i=0}^{[2T_\mathrm{B}F]-1} w_i' \mathrm{e}^{-2\pi \mathbb{i} f_j i \Delta\tau}
\end{aligned}$$

with the imaginary unit $\mathbb{i}$ yielding the complex Fourier tranforms $U'(f_j)$ and $W'(f_j), f_j = {}^j/_{2T_\mathrm{B}}, j = 0 \ldots [2T_\mathrm{B}F] - 1$. The energy spectra of the interpolated signals then is

$$\begin{aligned}
E_u'(f_j) &= \frac{1}{F^2} U'^*(f_j) U'(f_j) \\
E_w'(f_j) &= \frac{1}{F^2} W'^*(f_j) W'(f_j)
\end{aligned}$$

with the conjugate complex * and the correlation functions of the interpolated signals can be derived using the inverse DFT (IDFT)

$$R'_u(\tau_k) = \frac{F}{T_\mathrm{B}}\mathrm{IDFT}\left\{E'_u(f_j)\right\} = \frac{1}{2T_\mathrm{B}^2}\sum_{j=0}^{[2T_\mathrm{B}F]-1} E'_u(f_j)\mathrm{e}^{2\pi\mathrm{i}\tau_k f_j}$$

$$R'_w(\tau_k) = \frac{F}{T_\mathrm{B}}\mathrm{IDFT}\left\{E'_w(f_j)\right\} = \frac{1}{2T_\mathrm{B}^2}\sum_{j=0}^{[2T_\mathrm{B}F]-1} E'_w(f_j)\mathrm{e}^{2\pi\mathrm{i}\tau_k f_j}$$

At this point the cross-correlation functions are $2T_\mathrm{B}$ long. The maximum time lag of the correlation function is typically chosen much smaller than the duration of the measurement. This reduces the estimation variance of the final spectral estimate [14, 16]. With a given temporal resolution of the correlation function of $\Delta\tau$ the length of the correlation functions can be reduced by choosing a number of samples K with $K\Delta\tau < 2T_\mathrm{B}$. By rearranging the values obtained by the IDFT, the correlation function can be estimated for $\tau_k = k\Delta\tau, k = -\lfloor K/2\rfloor \ldots \lfloor (K-1)/2\rfloor$.

At the same time the low-pass filter of the interpolation can be corrected following the procedure given in [20].

$$R''_u(\tau_k) = \begin{cases} R'_u(0) & \text{for } \tau_k = 0 \\ (1+2c)R'_u(\tau_k) - cR'_u(\tau_{k+1}) - cR'_u(\tau_{k-1}) & \text{otherwise} \end{cases}$$

$$R''_w(\tau_k) = \begin{cases} R'_w(0) & \text{for } \tau_k = 0 \\ (1+2c)R'_w(\tau_k) - cR'_w(\tau_{k+1}) - cR'_w(\tau_{k-1}) & \text{otherwise} \end{cases}$$

with the constant

$$c = \frac{\mathrm{e}^{-n\Delta\tau}}{(1-\mathrm{e}^{-n\Delta\tau})^2}$$

and with the mean data rate n.

3.3 Normalization, noise removal, Bessel's correction and final transformation

The final estimate of the correlation function (reduced to the total length of T_C) is obtained by normalization

$$R(\tau_k) = \begin{cases} \frac{R''_u(\tau_k)}{R''_w(\tau_k)} - \sigma_n^2 + \sigma_{\bar{u}}^2 & \text{for } \tau_k = 0 \\ \frac{R''_u(\tau_k)}{R''_w(\tau_k)} + \sigma_{\bar{u}}^2 & \text{otherwise} \end{cases}$$

including Bessel's correction and noise removal, where $\sigma_{\bar{u}}^2$ is the estimated variance of the mean estimator above and σ_n^2 is the estimated variance of the noise. In the case of the interpolation method $\sigma_{\bar{u}}^2$ is obtained similar to the procedure given in [16] for the direct spectral estimator and for the slotting technique adpated to the interpolation method as

$$\sigma_{\bar{u}}^2 = \frac{\sum\limits_{k=-\lfloor K/2 \rfloor}^{\lfloor (K-1)/2 \rfloor} R_u''(\tau_k)}{T_\mathrm{B} F - \sum\limits_{k=-\lfloor K/2 \rfloor}^{\lfloor (K-1)/2 \rfloor} R_w''(\tau_k)}$$

Data noise contributes to the estimated correlation function at zero time lag only as an added value, whereas for all other time lags noise contributes as an increased estimation variance only, while no systematic deviation can be found there. The spectrum instaed shows a constant offset for all frequencies. The idea of the noise removal in [20] was to model the correlation function towards time lag zero and replace the estimated (and noise affected) value $R(0)$ by the value obtained by the correlation model. Such a procedure always relays on the validity of the model applied. A model-independent, and therefore more objective method to suppress the influence of the data noise can be found with the slotting technique as well as with the direct spectral estimation, where, in both cases, all self-products are removed from the respective sums of velocity products. Doing so, at time lag zero, only cross-products of velocity samples with independent noise values and short inter-arrival times contribute to the value of $R(0)$, while the noise-affected self-products are not counted. For the interpolation method, a similar method is aspired, where for each re-sampling interval the cross- and self-products of all original samples are counted. Unfortunately, after the interpolation and re-sampling process, only the last samples within each re-sampling interval are available, which makes it necessary to re-process the original, irregularly sampled data. Therefore, three additional re-sampled data series are derived, namely $s_{0,i}' = s_0'(t_i) = s_0'(i\Delta\tau)$, counting the number of original, irregular samples within each interval $t_{i-1} \le t < t_i$, $s_{1,i}' = s_1'(t_i) = s_1'(i\Delta\tau)$ giving the sum of all velocities within this interval and $s_{2,i}' = s_2'(t_i) = s_2'(i\Delta\tau)$ giving the sum of all velocity squared within this interval. All velocity values are meant with the mean velocity removed, therefore, these “interpolated” and re-sampled series can be obtained in a second step only, after the mean value has been derived as given above.

The velocity variance estimated including the self-products then reads

$$s_u^2 = \frac{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{1,i}'^2}{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'^2}$$

and the variance estimated without the self-products then reads

$$s_u^2 = \frac{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{1,i}'^2 - \sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{2,i}'^2}{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'^2 - \sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'}$$

Since both values are biased due to the correlation between the instantaneous data rate and the velocity (statistical bias), the correlation value at time lag zero is not simply replaced by the variance estimate without the self-products. Instead the difference between the two estimates, with and without the self-products, is taken as an estimate of the noise variance

$$\sigma_n^2 = \frac{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{1,i}'^2}{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'^2} - \frac{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{1,i}'^2 - \sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{2,i}'^2}{\sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'^2 - \sum\limits_{i=0}^{[2T_\mathrm{B}F]-1} s_{0,i}'}$$

and finally subtracted from the correlation value $R(0)$.

The final correlation estimate is then transformed by means of the discrete Fourier transform (DFT) to a power spectral density

$$S(f_j) = \Delta\tau \cdot \mathrm{DFT}\left\{R(\tau_k)\right\} = \Delta\tau \sum_{k=-\lfloor K/2 \rfloor}^{\lfloor (K-1)/2 \rfloor} R(\tau_k) \mathrm{e}^{-2\pi \mathrm{i} f_j \tau_k}$$

with $f_j = j\Delta f, j = -\lfloor K/2 \rfloor \ldots \lfloor (K-1)/2 \rfloor$ giving a frequency resolution of $\Delta f = 1/K\Delta\tau$.

3.4 Remarks

Local normalization and fuzzy slotting [27, 28, 19, 26, 17] have not yet been adapted to the interpolation method. The noise removal procedure suppresses the influence of noise components in the data on the derived statistical functions. The statistical bias is suppressed due to the inherent weighting of the sample-and-hold interpolation, which holds values longer if the local data rate decreases. An example program can be found at [1].

References

[1] URL `http://ldvproc.nambis.de/programs/pyLDV.html`.

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

[3] P Buchhave, W K George, Jr, and J L Lumley. The measurement of turbulence with the laser Doppler anemometer. In *Annual Review of Fluid Mechanics*, volume 11, pages 443–503. Annual Reviews, Inc., Palo Alto, CA, 1979.

[4] P Buchhave and C M Velte. Reduction of noise and bias in randomly sampled power spectra. *Experiments in Fluids*, 56:79, 2015.

[5] M Gaster and J B Roberts. Spectral analysis of randomly sampled signals. *J. Inst. Maths. Applics.*, 15:195–216, 1975.

[6] M Gaster and J B Roberts. The spectral analysis of randomly sampled records by a direct transform. *Proc. R. Soc. Lond. A.*, 354:27–58, 1977.

[7] W K George. Quantitative measurement with the burst-mode laser Doppler anemometer. *Exp. Therm. and Fluid Sci.*, 1:29–40, 1988.

[8] W K George, P D Beuther, and J L Lumley. Processing of random signals. pages 757–800, 1978. Proceedings of the dynamic flow conference, Skovlunde, Denmark, (PO Box 121, DK-2740 Skovlunde).

[9] W T Mayo, Jr. A discussion of the limitations and extensions of power spectrum estimation with burst counter LDV systems. pages 90–104, 1974. Proceedings of the second international workshop on laser velocimetry, (Bull. No. 144), Purdue Univ., West Lafayette, Ind.

[10] W T Mayo, Jr, M T Shay, and S Ritter. The development of new digital data processing techniques for turbulence measurements with a laser velocimeter, 1974.

[11] S Moreau, G Plantier, J-C Valière, H Bailliet, and L Simon. Estimation of power spectral density from laser Doppler data via linear interpolation and deconvolution. *Experiments in Fluids*, 50:179–188, 2011.

[12] H Nobach. *Verarbeitung stochastisch abgetasteter Signale: Anwendung in der Laser-Doppler-Anemometrie.* Shaker, Aachen, 1998. ISBN 3-8265-3332-1, Zugl.: Rostock, Univ., Diss., 1997.

[13] H Nobach. Processing of stochastic sampled data in laser Doppler anemometry. In *Proc. 3rd International Workshop on Sampling Theory and Applications, Loen, Norway*, pages 149–154, 1999.

[14] H Nobach. A global concept of autocorrelation and power spectral density estimation from LDA data sets. In *Proc. 10th Int. Symp. on Appl. of Laser Techn. to Fluid Mechanics*, Lisbon, Portugal, 2000. paper 3.5.

[15] H Nobach. Local time estimation for the slotted correlation function of randomly sampled LDA data. *Exp. in Fluids*, 32:337–345, 2002.

[16] H Nobach. Corrections to the direct spectral estimation for laser Doppler data. *Exp. in Fluids*, 56:109, 2015.

[17] H Nobach. Fuzzy time quantization and local normalization for the direct spectral estimation from laser Doppler velocimetry data. *Exp. in Fluids*, 56:182, 2015.

[18] H Nobach. A model-free noise removal for the interpolation method of correlation and spectral estimation from laser Doppler data. *Exp. in Fluids*, 56:100, 2015.

[19] 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.

[20] 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.

[21] J B Roberts and D B S Ajmani. Spectral analysis of randomly sampled signals using a correlation-based slotting technique. *IEE Proc.*, 133(2):153–162, 1986.

[22] P F Scott. Random sampling theory and its application to laser velocimeter turbulent spectral measurements. Technical report, General Electric Co., Corporate Res. and Development, 1974. Report No. 74CRD216, Tech, Info Series.

[23] P F Scott. *Distortion and Estimation of the Autocorrelation Function and Spectrum of a Randomly Sampled Signal*. PhD thesis, Rensselear Polytechnic Institute, Troy, NY, 1976.

[24] M T Shay. *Digital Estimation of Autocovariance Functions and Power Spectra from Randomly Sampled Data Using a Lag Product Technique*. PhD thesis, Texas A&M University, College Station, TX, 1976.

[25] L Simon and J Fitzpatrick. An improved sample-and-hold reconstruction procedure for estimation of power spectra from LDA data. *Exp. in Fluids*, 37:272–280, 2004.

[26] 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 Eng. and Laser Anemometry Conf.*, pages 347–355, Hilton Head Island, South Carolina, USA, 1995. FED-Vol. 229.

[27] 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.

[28] 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.

[29] C M Velte, W K George, and P Buchhave. Estimation of burst-mode LDA power spectra. *Exp. in Fluids*, 55:1674, 2014.