A direct estimator for autocorrelation and spectral density estimation from laser Doppler anemometry data including a fuzzy-slotting like time quantization and local normalization

Holger Nobach

December 3, 2016

Abstract
The direct estimation 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 \[3, 26, 27, 9, 10, 14, 16, 20, 21, 22, 23, 25\], direct spectral estimators, estimating a spectrum (periodogram) directly from the randomly sampled data \[3, 4, 5, 6, 7, 8, 15, 16, 28\] 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 \[2, 11, 19, 24\] and noise removal \[17, 19\].

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 sets of irregularly sampled velocity data \( u_i = u(t_i) \) at sampling times \( t_i, i = 0 \ldots N - 1 \) is assumed together with individual weights \( w_i \) according to the velocity samples \( u_i \), e.g. the particle’s transit times. If individual weights are not available, the inter-arrival times can be used for weighting, where both, the forward and the backward inter-arrival times are necessary for the correlation and spectral estimations.

\[
\begin{align*}
w_{bw,i} &= t_i - t_{i-1} \\
w_{fw,i} &= t_{i+1} - t_i
\end{align*}
\]

To avoid that gaps in the data stream of experimental data lead to improperly large weights, as has been observed in experiments, all inter-arrival time weights derived from inter-arrival times larger than five times the mean inter-arrival time are set to zero. Due to this one looses only about 0.7% of useful data, while the outliers of large inter-arrival times are suppressed effectively.

3 Direct Spectral Estimator

The direct spectral estimation as given in [15] is reviewed including individual data weighting (e.g. transit-time weighting) or forward-backward inter-arrival-time weighting as an alternative, if reliable estimates of the particle transit times are not available and several corrections to the algorithm from [28]. Local normalization [27, 25] and fuzzy slotting [18] are adapted to the direct spectral estimation method and introduced as optional extensions of the base algorithm. The Python code of the estimator is available at [1].

3.1 Base Algorithm

A corrected direct spectral estimator with individual data weights or forward-backward interarrival-time weighting, correction of influences of the random sampling and processor dead times and Bessel’s correction is given in [13]. The procedure consists of the following steps:

1. calculation of the primary spectra

\[
S_u(f) = \frac{T_B [U_{FB}(f) + U_{FB}^*(f)]}{W_{FB}(0) + W_{FB}^*(0)}
\]  

(1)

and

\[
S_w(f) = \frac{T_B [W_{FB}(f) + W_{FB}^*(f)]}{W_{FB}(0) + W_{FB}^*(0)}
\]

(2)
(Note, that the denominators of these equations have been modified compared to [15].) with

\[
U_{FB}(f) = \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} w_{bw,i} w_{fw,j} u_i u_j e^{-2\pi i f (t_j - t_i)} \quad (3a)
\]

\[
= \sum_{j=1}^{N-1} w_{fw,j} u_j e^{-2\pi i f t_j} \left( \sum_{i=0}^{j-1} w_{bw,i} u_i e^{-2\pi i f t_i} \right)^* \quad (3b)
\]

\[
W_{FB}(f) = \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} w_{bw,i} w_{fw,j} e^{-2\pi i f (t_j - t_i)} \quad (4a)
\]

\[
= \sum_{j=1}^{N-1} w_{fw,j} e^{-2\pi i f t_j} \left( \sum_{i=0}^{j-1} w_{bw,i} e^{-2\pi i f t_i} \right)^* \quad (4b)
\]

where \(T_B\) is the total time of the data set or that of the data block and the asterix means the conjugate complex. Note, that the second sum is part of the addends of the first, outer sum. Note further, that the sums are not independent of each other, which means that the Fourier transform is not complete and direct calculations of the sums are necessary. Luckily, the sums can be calculated parallel, such that only the individual samples must be processed instead of all pairs of samples.

2. transformation into correlation functions and limiting the correlation function (alternative to block averaging)

\[
R_u(\tau_k) = F \cdot \text{IDFT} \{ S_u(f_j) \} = \frac{1}{J \Delta \tau} \sum_{j=\left\lfloor -\frac{J}{2} \right\rfloor}^{\left\lfloor \frac{J-1}{2} \right\rfloor} S_u(f_j) e^{2\pi i f_j \tau_k} \quad (5)
\]

and

\[
R_w(\tau_k) = F \cdot \text{IDFT} \{ S_w(f_j) \} = \frac{1}{J \Delta \tau} \sum_{j=\left\lfloor -\frac{J}{2} \right\rfloor}^{\left\lfloor \frac{J-1}{2} \right\rfloor} S_w(f_j) e^{2\pi i f_j \tau_k} \quad (6)
\]

by means of the inverse discrete Fourier transform IDFT with the fundamental frequency \(F\), which defines the temporal resolution \(\Delta \tau = 1/F\) of the obtained correlation functions. The primary spectra are calculated for \(J\) different frequencies \(f_j = j \Delta f, j = -\left\lfloor \frac{J}{2} \right\rfloor \ldots \left\lfloor \frac{J-1}{2} \right\rfloor\) with \(J = 2T_B F\) and \(\Delta f = F/J = 1/2T_B\). The correlation function is calculated for \(K\) different time lags, where \(K\) is chosen significantly smaller than \(J\) according to the considered correlation interval \([-T_C/2 : T_C/2], T_C = K/F\), for \(\tau_k = k \Delta \tau, k = -\left\lfloor K/2 \right\rfloor \ldots \left\lfloor (K-1)/2 \right\rfloor\).
3. normalization (correction of dead time influences) and Bessel’s correction

\[ R(\tau_k) = \frac{R_u(\tau_k)}{R_w(\tau_k)} + \sigma_u^2, \]  

(7)

for \( \tau_k = k\Delta \tau, k = -[K/2] \ldots ([K-1]/2) \) with \( \sigma_u^2 \) an estimate of the variance of the mean estimation. An appropriate estimator follows.

4. back-transformation into the final spectrum by means of the discrete Fourier transform DFT

\[ S(f_j) = \Delta \tau \cdot \text{DFT} \{ R(\tau_k) \} = \Delta \tau \sum_{k=-[K/2]}^{([K-1]/2)} R(\tau_k) e^{-2\pi j f_j \tau_k} \]  

(8)

with the (reduced) spectral resolution \( \Delta f = F/K = 1/T_\ell \) for \( f_j = j\Delta f, j = -[K/2] \ldots ([K-1]/2) \)

3.2 Local Normalization

A further reduction of random errors can be achieved with the local normalization \([26, 27, 25]\). The main idea of this correction is to normalize the correlation estimate at every time lag by a variance estimate, which corresponds to exactly the data samples used for the correlation estimate. The result is a correlation coefficient \( \rho \) in the range \([-1 : 1]\). The method has been developed originally for the slotting technique. However, it can be adapted also to the direct spectral estimator as one step of the various corrections acting on the correlation function. The new estimate of the correlation coefficient then becomes

\[ \rho(\tau_k) = \frac{R_u(\tau_k) + \sigma_u^2 R_w(\tau_k)}{\sqrt{(R_{u1}(\tau_k) + \sigma_u^2 R_{w1}(\tau_k)) (R_{u2}(\tau_k) + \sigma_u^2 R_{w2}(\tau_k))}}. \]  

(9)

with \( f_j = j\Delta f, j = -[J/2] \ldots ([J-1]/2) \) and \( \tau_k = k\Delta \tau, k = -[K/2] \ldots ([K-1]/2) \)

(The inverse discrete Fourier transform originally yields \( \tau_k = k\Delta \tau, k = -[J/2] \ldots ([J-1]/2) \) where only the range \( k = -[K/2] \ldots ([K-1]/2) \) is processed further.) with \( R_u(\tau_k) \) and \( R_w(\tau_k) \) as above and

\[ R_{u1}(\tau_k) = F \cdot \text{IDFT} \{ S_{u1}(f_j) \} = \frac{1}{J\Delta \tau} \sum_{j=-[J/2]}^{([J-1]/2)} S_{u1}(f_j) e^{2\pi i j f_j \tau_k} \]  

(10)

and

\[ R_{u2}(\tau_k) = F \cdot \text{IDFT} \{ S_{u2}(f_j) \} = \frac{1}{J\Delta \tau} \sum_{j=-[J/2]}^{([J-1]/2)} S_{u2}(f_j) e^{2\pi i j f_j \tau_k} \]  

(11)

with

\[ S_{u1}(f) = \frac{T_B [U_{FB,1}(f) + U_{FB,2}(f)]}{W_{FB}(0) + W_{FB}^*(0)} \]  

(12)
and

$$S_{u_2}(f) = \frac{T_B [U_{FB,2}(f) + U^*_{FB,1}(f)]}{W_{FB}(0) + W^*_{FB}(0)}$$  (13)

(Note, that the denominators of these equations have been modified compared to [15].) with

$$U_{FB,1}(f) = \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} w_{bw,i} w_{fw,j} u^2_i e^{-2\pi i f(t_j - t_i)}$$  (14a)

$$= \sum_{j=1}^{N-1} w_{fw,j} e^{-2\pi i f t_j} \left( \sum_{i=0}^{j-1} w_{bw,i} u^2_i e^{-2\pi i f t_i} \right)^*$$  (14b)

and

$$U_{FB,2}(f) = \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} w_{bw,i} w_{fw,j} u^2_j e^{-2\pi i f(t_j - t_i)}$$  (15a)

$$= \sum_{j=1}^{N-1} w_{fw,j} u^2_j e^{-2\pi i f t_j} \left( \sum_{i=0}^{j-1} w_{bw,i} e^{-2\pi i f t_i} \right)^*$$  (15b)

Note that $J$ must be chosen odd to ensure that $R_{u_1}(\tau_k)$ and $R_{u_2}(\tau_k)$ are real. This is required only for the case of local normalization applied.

To obtain a correlation function $R(\tau_k)$ and finally a power spectral density, the correlation coefficient function $\rho(\tau_k)$ is expanded by an estimate of the velocities variance $\sigma^2_u$ yielding

$$R(\tau_k) = \sigma^2_u \rho(\tau_k)$$  (16)

The velocity variance is estimated as

$$\sigma^2_u = \frac{\sum_{i=0}^{N-1} w_i u^2_i}{\sum_{i=0}^{N-1} w_i} + \sigma^2_u$$  (17)

The weights $w_i$ are as before for the mean estimate and $\sigma^2_u$ again is the estimate of the variance of the mean estimator. Note that in the variance estimate the data are assumed to be mean-free in both cases (naturally or by estimating and removing the mean from the data), therefore $u_i$ is given in both cases instead of $u_i - \bar{u}$.

The final spectral estimate is again obtained from the correlation by a discrete Fourier transform (DFT) as given above.
3.3 Fuzzy Time Quantization

The adaptation of the fuzzy slotting \[26, 18\] to the direct estimator is not that straightforward, because the sums \(U_{FB}(f), W_{FB}(f), U_{FB,1}(f)\) and \(U_{FB,2}(f)\) above allow access to the inter-arrival times only within the double sum. Separated into two sums, access is given to arrival times only. To obtain still the same weighting of cross-products depending on their inter-arrival time as it is with the window function \(b_k(\Delta t)\) for the slotting technique, all Dirac pulses of the data sequence \(u_i\delta(t - t_i)\) are replaced by finite pulses of the duration \(\Delta \tau\) (symmetric around the original arrival time) and constant amplitude of \(u_i/\Delta \tau\) (Fig. 1). The correlation of these finite pulses then has the triangular shape as the fuzzy slot window functions. If these are sampled at integer numbers of \(\Delta \tau\), the contribution of each data sample to the neighboring time lags corresponds to the values of the slotting window function at the inter-arrival time between the original pulses. Overlapping pulses superimpose linearly. The sums \(U_{FB}(f)\)
and \( W_{FB}(f) \) then become

\[
U_{FB}(f) = \sum_{j=1}^{N-1} w_{bw,j} u_j \left( \frac{1}{\Delta \tau} \int_{t_j - \Delta \tau}^{t_j + \Delta \tau} e^{-2\pi i ft} \, dt \right) \left[ \sum_{i=0}^{j-1} w_{bw,i} u_i \left( \frac{1}{\Delta \tau} \int_{t_i - \Delta \tau}^{t_i + \Delta \tau} e^{-2\pi i ft} \, dt \right) \right]^*
\]

\[
= \begin{cases} 
\sum_{j=1}^{N-1} w_{bw,j} u_j \left[ \frac{1}{2\pi f \Delta \tau} \left( e^{-2\pi i f (t_j + \Delta \tau)} - e^{-2\pi i f (t_j - \Delta \tau)} \right) \right] \\
\times \left\{ \sum_{i=0}^{j-1} w_{bw,i} u_i \left[ \frac{1}{2\pi f \Delta \tau} \left( e^{-2\pi i f (t_i + \Delta \tau)} - e^{-2\pi i f (t_i - \Delta \tau)} \right) \right] \right\}^* \quad \text{for } f \neq 0 \\
\sum_{j=1}^{N-1} w_{bw,j} u_j \left( \sum_{i=0}^{j-1} w_{bw,i} u_i \right)^* \quad \text{for } f = 0 
\end{cases}
\]

\( (18a) \)

\[
W_{FB}(f) = \sum_{j=1}^{N-1} w_{bw,j} \left( \frac{1}{\Delta \tau} \int_{t_j - \Delta \tau}^{t_j + \Delta \tau} e^{-2\pi i ft} \, dt \right) \left[ \sum_{i=0}^{j-1} w_{bw,i} \left( \frac{1}{\Delta \tau} \int_{t_i - \Delta \tau}^{t_i + \Delta \tau} e^{-2\pi i ft} \, dt \right) \right]^*
\]

\[
= \begin{cases} 
\sum_{j=1}^{N-1} w_{bw,j} \left[ \frac{1}{2\pi f \Delta \tau} \left( e^{-2\pi i f (t_j + \Delta \tau)} - e^{-2\pi i f (t_j - \Delta \tau)} \right) \right] \\
\times \left\{ \sum_{i=0}^{j-1} w_{bw,i} \left[ \frac{1}{2\pi f \Delta \tau} \left( e^{-2\pi i f (t_i + \Delta \tau)} - e^{-2\pi i f (t_i - \Delta \tau)} \right) \right] \right\}^* \quad \text{for } f \neq 0 \\
\sum_{j=1}^{N-1} w_{bw,j} \left( \sum_{i=0}^{j-1} w_{bw,i} \right)^* \quad \text{for } f = 0 
\end{cases}
\]

\( (19a) \)
Following the derivations in [15] to obtain an estimate of the variance of the mean estimator, if local normalization is used, $U_{FB,1}(f)$ and $U_{FB,2}(f)$ change accordingly to

$$U_{FB,1}(f) = \sum_{j=1}^{N-1} w_{bw,j} \left( \frac{1}{\Delta\tau} \int_{t_j - \Delta\tau}^{t_j + \Delta\tau} e^{-2\pi i f t} dt \right) \left[ \frac{1}{\Delta\tau} \int_{t_j - \Delta\tau}^{t_j + \Delta\tau} e^{-2\pi i f t} dt \right]$$

(20a)

$$= \left\{ \begin{array}{cl}
\sum_{j=1}^{N-1} w_{bw,j} \left[ \frac{i}{2\pi f\Delta\tau} \left( e^{-2\pi i f(t_j + \Delta\tau)} - e^{-2\pi i f(t_j - \Delta\tau)} \right) \right] & \\
\times \left\{ \sum_{i=0}^{j-1} w_{bw,i} u_i^2 \left[ \frac{i}{2\pi f\Delta\tau} \left( e^{-2\pi i f(t_i + \Delta\tau)} - e^{-2\pi i f(t_i - \Delta\tau)} \right) \right] \right\} & \\
\sum_{j=1}^{N-1} w_{bw,j} \left( \sum_{i=0}^{j-1} w_{bw,i} u_i^2 \right)^* & 
\end{array} \right. \quad \text{for } f \neq 0$$

(20b)

$$= \left\{ \begin{array}{cl}
\sum_{i=0}^{j-1} w_{bw,i} u_i^2 \left[ \sum_{j=1}^{N-1} w_{bw,j} e^{-2\pi i f t_j} e^{-2\pi i f t_i} \right] & \\
\left\{ \sum_{j=1}^{N-1} w_{bw,j} \left( \sum_{i=0}^{j-1} w_{bw,i} u_i^2 \right) \right\}^* & 
\end{array} \right. \quad \text{for } f = 0$$

(20c)

and

$$U_{FB,2}(f) = \sum_{j=1}^{N-1} w_{bw,j} u_j^2 \left( \frac{1}{\Delta\tau} \int_{t_j - \Delta\tau}^{t_j + \Delta\tau} e^{-2\pi i f t} dt \right) \left[ \frac{1}{\Delta\tau} \int_{t_j - \Delta\tau}^{t_j + \Delta\tau} e^{-2\pi i f t} dt \right]$$

(21a)

$$= \left\{ \begin{array}{cl}
\sum_{j=1}^{N-1} w_{bw,j} u_j^2 \left[ \frac{i}{2\pi f\Delta\tau} \left( e^{-2\pi i f(t_j + \Delta\tau)} - e^{-2\pi i f(t_j - \Delta\tau)} \right) \right] & \\
\times \left\{ \sum_{i=0}^{j-1} w_{bw,i} \left[ \frac{i}{2\pi f\Delta\tau} \left( e^{-2\pi i f(t_i + \Delta\tau)} - e^{-2\pi i f(t_i - \Delta\tau)} \right) \right] \right\} & \\
\sum_{j=1}^{N-1} w_{bw,j} u_j^2 \left( \sum_{i=0}^{j-1} w_{bw,i} \right)^* & 
\end{array} \right. \quad \text{for } f \neq 0$$

(21b)

$$= \left\{ \begin{array}{cl}
\sum_{i=0}^{j-1} w_{bw,i} u_i^2 \left[ \sum_{j=1}^{N-1} w_{bw,j} e^{-2\pi i f t_j} e^{-2\pi i f t_i} \right] & \\
\left\{ \sum_{j=1}^{N-1} w_{bw,j} u_j^2 \left( \sum_{i=0}^{j-1} w_{bw,i} \right) \right\}^* & 
\end{array} \right. \quad \text{for } f = 0$$

(21c)

### 3.4 Estimation of the variance of the mean estimator

Following the derivations in [15] to obtain an estimate of the variance of the mean estimator from the data set, the previous correlation estimates $R_u(\tau_k)$
and $R_w(\tau_k)$ can be re-used, yielding

$$
\sigma_u^2 = \frac{1}{T_BF} \sum_{k=-\left\lfloor K/2 \right\rfloor}^{\left\lfloor (K-1)/2 \right\rfloor} R_w(\tau_k) + \frac{1}{\pi} \sum_{i=0}^{N-1} w_{bw,i} w_{lw,i} u_i^2 \left[ 1 - \frac{1}{T_BF} \sum_{k=-\left\lfloor K/2 \right\rfloor}^{\left\lfloor (K-1)/2 \right\rfloor} R_w(\tau_k) \right]$$

(22)

with

$$W = W_{FB}(0) + W_{FB}^*(0) \quad (23)$$

Note, that this equation has been modified compared to \cite{15}.

3.5 Remarks

A statistical bias due to the correlation of the velocity and the instantaneous data rate are suppressed due to the implementation of the weighting schemes. Since the self-products have been removed from the sums, noise components in the data will not cause systematic errors in the derived statistical functions. An example program can be found at \cite{1}.

References


