

TrackFit: Uncertainty Quantification, Optimal Filtering and Interpolation of Tracks for Time-Resolved Lagrangian Particle Tracking

S. Gesemann^{1*}

¹ DLR (German Aerospace Center), Institute of Aerodynamics and Flow Technology, Department of Experimental Methods, Göttingen, Germany

* sebastian.gesemann@dlr.de

Abstract

Advanced Lagrangian Particle Tracking methods (such as the STB algorithm (Schanz et al. 2016)) are a very useful tool for uncovering properties of flow. As a measurement technique, the results of such methods are perturbed by different sources of errors and noise. This work addresses the problem of optimal filtering of particle tracks as well as estimating uncertainties of derived quantities such as location, velocity and acceleration of observed particles. The behavior and performance of this new filtering method (“TrackFit”), first introduced at Gesemann et al. (2016) is analyzed and compared to the Savitzky–Golay filter (Savitzky and Golay (1964)) which is commonly used for these purposes. The optimal choice of parameters of this filtering method as well as the uncertainty quantification of the reconstructed tracks can be extracted from a spectral analysis of the recorded raw particle tracking data. This is in contrast to a Savitzky–Golay filter where the choice of parameters might often be driven by experience and gut feeling. Estimating the power spectral density (PSD) of the particle trajectory signals for the purpose of optimal filtering parameter selection represents a challenge due to possibly short trajectory signals. In the following work we will present a method for PSD estimation that is applicable in this scenario. In addition, we show that regardless of the choice of Savitzky–Golay filter parameters, the resulting filter will not approximate the ideal noise reduction filter well unlike the “TrackFit” described in this work.

1 Introduction

In the interest of optimal noise reduction filtering and interpolation for derived quantities such as velocity and acceleration and their uncertainty estimation, knowledge about the spectral properties of the position-over-time signals is important. We expect the low frequencies of such signals to have a lot of energy and a high signal-to-noise ratio while for higher frequencies we expect a low signal-to-noise ratio due to mostly white measurement noise. Further, we expect the measurement noise to be additive and statistically uncorrelated to the true positional signals. In such a case the optimal filter (optimal with respect to minimizing the sum of squared errors) would reduce to a simplified Wiener filter that ignores cross correlations between signal and noise. Such a filter’s response can be expressed as follows:

$$H(f) = \frac{SNR(f)}{1 + SNR(f)} \quad (1)$$

In equation 1 $SNR(f)$ refers to the local signal power to noise power ratio close to the frequency f . For a high signal-to-noise ratio the filter gain will approach 1 while for a low signal-to-noise ratio the gain will approach zero to suppress the noise.

Two challenges arise in the context of Lagrangian Particle Tracking (LTP). The particle trajectories may span only a small number of time steps (order of 30). In order to estimate the ideal filter transfer function, an accurate estimate of the power spectral densities of the true signal and the additive noise is necessary. But for such short signals, spectral estimation requires special attention. The possibly short nature of particle trajectories also poses a problem during the application of a filter especially at the borders of the signal

(the beginning and ending of a trajectory) because the filter would require samples outside of the recorded domain.

The remainder of this paper is structured as follows: Section 2 will cover the problem of spectral analysis for this kind of filtering problem. Section 3 will show the results of our spectral estimation method when applied to a known data set from the 1st LPT and DA challenge for testing whether the method results in a reasonable power spectrum estimate but also to verify that our assumptions about what such a spectrum may look like are true. Section 4 will describe the TrackFit filtering method (see also (Gesemann et al. (2016))) with its properties. In section 5 the TrackFit and Savitzky–Golay (Savitzky and Golay (1964)) filter methods are analyzed and compared. Section 7 closes with a summary and conclusion.

2 Spectral estimation

For our case with potentially short tracks which have strong low-frequency components we have developed an appropriate spectral estimation method. The method can be summarized with the following steps:

1. prefilter (FIR) the position-over-time data
2. compute autocorrelation of the result
3. apply Levinson-Durbin recursion for auto-regressive model on autocorrelation data
4. compensate for prefiltering in auto-regressive model

2.1 Prefilter

A prefilter can be used for spectral analysis as a first stage to flatten the signal spectrum. The spectrum of the resulting intermediate signal would be estimated instead to derive the original spectrum by compensating for the prefilter's response. This approach can be preferable to directly estimating the original signal's spectrum. In particular, an FFT-based estimation tends to suffer from spectral leakage due to the use of finite windows. Spectral leakage manifests in the form of a broader main lobe and the presence of side lobes for a single frequency that is present in the signal. These artefacts will be less noticeable if the signal already has a flat shape. For spectral estimation based on the auto-regressive model, prefiltering also has benefits. The effect of different prefilters is shown in section 2.2.

Apart from flattening the spectrum, such a prefilter should also have a short impulse response, e.g. two or three samples. This becomes obvious when the filter has to be applied to potentially short signals. Suppose a short signal with 25 particle locations is available. After applying a prefilter with an impulse response of three samples only 23 samples of the result can be used for estimating autocorrelation coefficients. The longer the filter's impulse response, the shorter the intermediate signals will get.

The prefilters we chose are first and second order digital FIR filters with the following transfer functions H_{p1} and H_{p2} :

$$\begin{aligned} H_{p1}(z) &= 1 - z^{-1} \\ H_{p2}(z) &= 1 - 1.9z^{-1} + 0.9z^{-2} \end{aligned} \tag{2}$$

These filters attenuate the low frequencies and amplify the higher frequencies.

2.2 Spectral estimation via auto-regressive model

The auto-regressive model (AR) basically describes a random process as filtered white noise using an all-pole IIR filter for coloring the noise. With the help of the Levinson-Durbin recursion the filter coefficients can be extracted from autocorrelation coefficients. This allows another form of spectral estimation in that the filter represents the spectral shape (with an average response of 0 dB over a linear frequency axis) and the noise power represents an overall offset for the power spectral density. See listing 1 for a GNU Octave script that shows this process.

```

% let acf be a vector with the autocorrelations
% for lags zero to p (inclusive)

% Auto-regressive model AR(p) computed from ACF using
% Levinson-Durbin recursion
[a,v] = levinson(acf);
% a = order p all-pole filter coefficients
% v = variance of white noise

% Power spectral density estimate by evaluating the
% filter's transfer function and scaling with variance
[h,f] = freqz(1,a); % transfer function
p = abs(h).^2 * v; % power spectral density estimate
loglog(f(2:end), p(2:end));

```

Listing 1: GNU Octave example script for estimating the power spectral density based on autocorrelation coefficients via an auto-regressive model of order p .

2.3 Test and comparison of spectral estimation methods

To test the various spectral estimation methods a ground truth spectrum has been chosen that is believed to be representative of typical particle tracks *including measurement noise*, see figure 1. The lower frequency part is dominated by true particle locations while the higher frequency part is dominated by flat white measurement noise. The ground truth is the sum of two power spectra: signal (PSD_s) and noise (PSD_n). On a normalized frequency axis (ranging from zero to one for the Nyquist frequency) they are defined as follows:

$$PSD_s(f) = \frac{1}{(3.77f)^2 + (641.52f^2)^2 + (17363.515f^3)^2} \quad (3)$$

$$PSD_n(f) = 0.00592 \quad (4)$$

Based on the chosen ground truth spectrum and the transfer function of the prefilter autocorrelation coefficients can be computed directly using a Fourier transform. Only 21 autocorrelation coefficients have been selected (for lags 0 to 20 inclusive) to test how the spectral estimation methods perform.

The FFT-based method extends the 21 coefficients for lags 0 to 20 to 41 coefficients for lags -21 to 21 by mirroring and zero-padding and applies a Hann window to the result. The windowed autocorrelation function is then used as input to a zero-padded FFT. The resulting magnitudes are squared and plotted against their frequency. The AR-based method applies the computations shown in listing 1 of section 2.2 with $p = 20$.

As can be seen in figure 1, FFT-based methods struggle to deal with such a spectrum given only a small symmetric window of 41 autocorrelation coefficients. We can clearly see the side lobes of lower frequencies that dominate the estimation of the power spectrum estimates for higher frequencies. The prefilters help reduce the damage by amplifying the higher frequencies and attenuating them again after the estimation. But the most important feature of the spectrum for the purpose of noise reduction is the frequency region in which the power spectrum flattens again because this is where the signal-to-noise ratio crosses the 0 dB level and the point at which a noise reduction filter should start to attenuate the high frequencies. Unfortunately, the FFT-based methods fail to resolve this reliably in this case even with the prefilters enabled ($k > 0$).

The auto-regressive approach tracks the true spectrum more closely in comparison. Most errors of the estimates are below 1% and only rise with the lower frequencies to about 6%. The estimated curves oscillate around the ground truth which is not surprising given that they are expressed using an order 20 polynomial. Between different prefilters the curves look similar. But the equation systems to compute the filter coefficients are very different in their "numerical difficulty", see table 1. High condition numbers can make the method unreliable in the light of finite precision arithmetic and rounding errors. The use of appropriate prefilters can lower the condition number and thus improve numerical stability of the method. In fact, the condition number can be seen as a measure of how well the prefilter is able to "whiten" the signal.

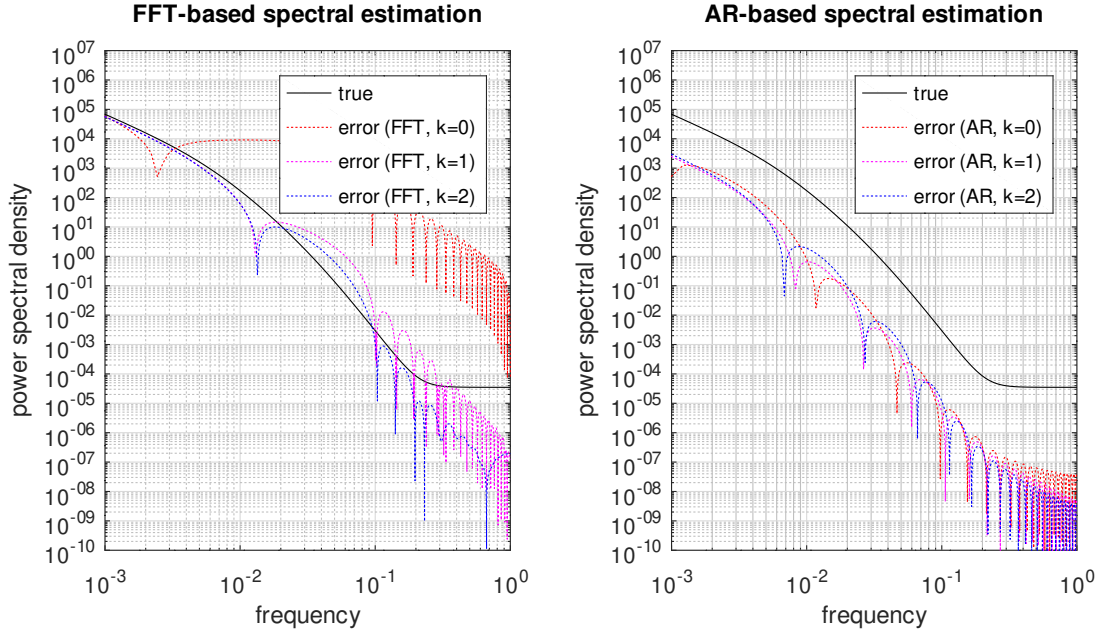


Figure 1: Power spectral density estimation method comparison for a synthetically generated case. The black curve represents the true signal spectrum. For the spectrum estimates only the errors to the ground truth are shown. k refers to the prefilter order with $k = 0$ meaning no prefilter.

	$k = 0$	$k = 1$	$k = 2$
condition number	$2.6 \cdot 10^8$	$3.3 \cdot 10^3$	$8.1 \cdot 10^1$

Table 1: Condition numbers for different prefilter orders

3 Spectral estimation test on fluid simulation data

The question arises whether the black curve in figure 1 for a hypothetical power spectrum is representative of measured particle motion in a fluid. Therefore, we applied our spectral estimation method on a known data set of the first data assimilation challenge (Sciacchitano et al. (2021a)). This data set covers particle locations for 25 time steps based on a fluid simulation and simulated white measurement noise.

Figure 2 shows the power spectral density estimate of the simulated particle data for the particles' z coordinates which closely matches the spectrum assumption from the previous section. In particular, we have the same slope of about -6 in the log-log plot right before the measurement noise becomes dominant and the curve flattens. We have observed this slope to be typical in these sorts of fluid experiments.

4 Filtering with TrackFit

Instead of low-pass filtering by direct convolution, a filtering effect can also be achieved by posing and solving an overdetermined linear equation system in the least squares sense such as the one in equation 5:

$$\begin{pmatrix}
 1w & 4w & 1w & & & & & & & \\
 & 1w & 4w & 1w & & & & & & \\
 & & 1w & 4w & 1w & & & & & \\
 & & & \ddots & \ddots & \ddots & & & & \\
 & & & & 1w & 4w & 1w & & & \\
 -1 & 3 & -3 & 1 & & & & & & \\
 & -1 & 3 & -3 & 1 & & & & & \\
 & & -1 & 3 & -3 & 1 & & & & \\
 & & & \ddots & \ddots & \ddots & \ddots & & & \\
 & & & & -1 & 3 & -3 & 1 & &
 \end{pmatrix} c = \begin{pmatrix} 6w \cdot s \\ 0 \end{pmatrix} \quad (8)$$

Here, the solution c is a set of scale factors which control the contribution of the various shifted B-spline base functions to the trajectory. Given a signal s with n samples, $n + 2$ scaling coefficients will be computed to cover every interval with four cubic base splines.

The choice of regularizing with the third order derivative is motivated by the slope of estimated frequency spectra of position-over-time signals in simulations as well as real STB experiments. A derivative order of k will approximate the response of the ideal Wiener filter for a signal slope of $-2k$ in a log-log plot right before the power spectral density curve starts to flatten.

Using this modification a continuous curve is reconstructed instead of a discrete signal. Thus, the curve can be evaluated at any point in time including its temporal derivatives for velocity and acceleration.

5 Analysis and Comparison

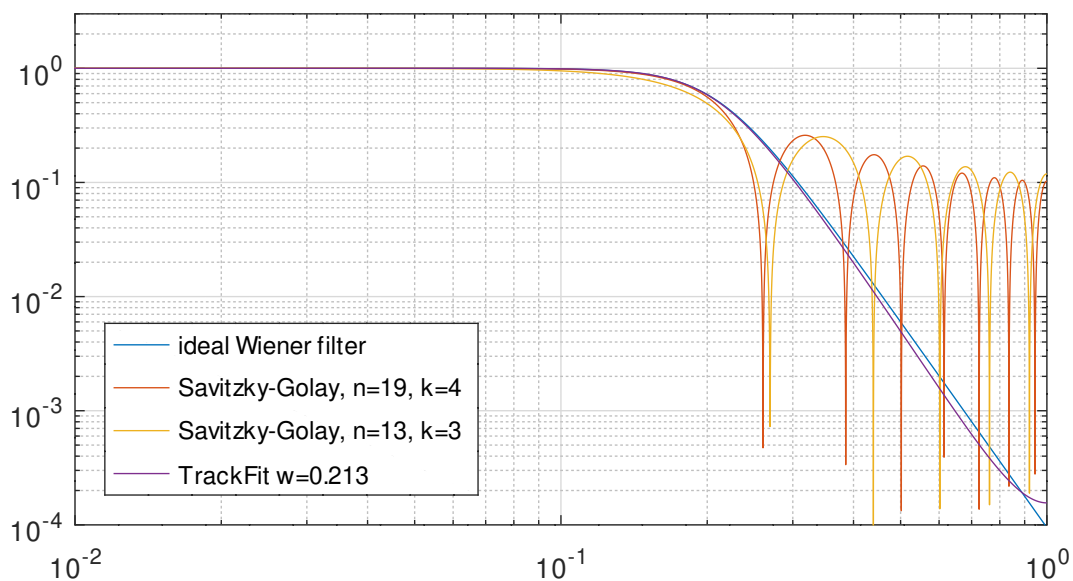


Figure 3: Comparison of ideal filter gain with Savitzky-Golay (window size n and polynomial order k) filters and TrackFit (cutoff frequency parameter w).

Based on the assumed ground truth spectrum that is the sum of the true signal and white noise the ideal filter transfer can be determined according to equation 1 and compared to the response of the real filter realizations of the Savitzky-Golay filter and TrackFit, see figure 3. For the Savitzky-Golay filter the window lengths n were selected for polynomial order $k = 3$ and $k = 4$ so that the filter's response would approximate

the ideal filter response. But these filters tend to have a poor suppression of high frequencies which would retain more of the measurement noise than necessary. The TrackFit approximates the ideal Wiener filter very well for this type of spectrum. It retains the necessary low frequency information and rejects the high frequency measurement noise just as the Wiener filter would.

6 Uncertainty Estimation

Each B-spline coefficient computed with TrackFit for a particle trajectory can be represented as a weighted sum of the raw particle location data with the help of the equation system's pseudo-inverse. In addition, each derived quantity from the B-spline curve (such as location, velocity or acceleration) is also a weighted sum of the B-spline coefficients. The concatenation of two linear mappings is itself linear. Therefore, any of these derived quantities can be written as a weighted sum of the raw unfiltered particle location data, for example, the velocity of a particular particle at time instant t based on the noisy particle location samples s_j :

$$v(t) = \sum_j \alpha_j(t) \cdot s_j \quad (9)$$

Under the assumption of uncorrelated white measurement noise with a power level σ_{le}^2 that can be extracted from the flat high frequency portion of the power spectral density estimates, it is possible to perform Gaussian error propagation. Assuming the variance of the location error in s_j is estimated to be σ_{le}^2 then the variance of the error of velocity σ_{ve}^2 can be written as

$$\sigma_{ve}^2(t) = \sigma_{le}^2 \sum_j \alpha_j^2(t) \quad (10)$$

This gives us uncertainty estimates of any quantity based on the spectral power density estimates of the measurement data and the noise reduction/interpolation approach which effectively controls the α scale factors that weight each individual particle location measurement.

7 Summary and Conclusion

In this work we have presented a method to perform spectral analyses on possibly short particle trajectories which is tailored to strong low frequency content and fast decay using a compensating prefilter. The resulting spectral information allows approaching the ideal noise reduction filter.

Further, the benefits of TrackFit (Gesemann et al. (2016)) have been highlighted. Specifically, it performs a joint noise reduction approximating the ideal Wiener filter with an interpolation that offers consistent temporal derivatives for velocity and acceleration.

In addition, it was shown how Gaussian error propagation can be applied based on the estimated level of the measurement noise floor from the spectra and the filtering coefficients α .

We have applied these methods in various experiments as well as both the first LPT challenge (Sciacchitano et al. (2021b)) and the first DA challenge (Sciacchitano et al. (2021a)) with good results.

Acknowledgements

The project leading to this contribution has received funding in the frame of the project HOMER from the European Union's Horizon 2020 research and innovation program under grant agreement No. 769237.

References

- Gesemann S, Huhn F, Schanz D, and Schröder A (2016) From noisy particle tracks to velocity, acceleration and pressure fields using b-splines and penalties. in *18th international symposium on applications of laser and imaging techniques to fluid mechanics, Lisbon, Portugal*. pages 4–7
- Savitzky A and Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures.. *Analytical Chemistry* 36:1627–1639

Sciacchitano A, Leclaire B, and Schröder A (2021a) Main results of the first data assimilation challenge. in *14th International Symposium on Particle Image Velocimetry*

Sciacchitano A, Leclaire B, and Schröder A (2021b) Main results of the first lagrangian particle tracking challenge. in *14th International Symposium on Particle Image Velocimetry*