

An open source MATLAB[®] package to perform basic statistical analysis of turbulence data and other complex systems along with its application to Fokker-Planck equation and Integral fluctuation theorem

André Fuchs **Swapnil Kharche** **Matthias Wächter** **Joachim Peinke**
Univ. of Oldenburg Univ. Grenoble Alpes Univ. of Oldenburg Univ. of Oldenburg

Abstract

We present a user-friendly open-source MATLAB[®] package developed by the research group Turbulence, Wind energy and Stochastics (TWiSt) at the Carl von Ossietzky University of Oldenburg. This package enables to perform a standard analysis of given turbulent data and extracts the stochastic equations describing the scale-dependent cascade process in turbulent flows through Fokker-Planck equations. As a precondition, Markovian properties of the process in scale are tested. Such a stochastic scale-dependent cascade process allows a comprehensive statistical description in terms of the complexity of the data. Cascade trajectories can be defined as single events, for each of which a total entropy production can be determined. For such entropy fluctuations a rigorous law of non-equilibrium stochastic thermodynamics, namely the integral fluctuation theorem, will be verified. As the analysis of the scale-dependent cascade process through a hierarchy of spatial and temporal scales in turbulent flows is an integral part of turbulence theory, this interdisciplinary treatment of the turbulent cascade process has the potential for a new way to link the statistical description of turbulence (via common two-point increment statistics), non-equilibrium stochastic thermodynamics and local turbulent flow structures. The presented package can be used also for the analysis of other data with turbulent like complexity.

Keywords: MATLAB[®], Fokker-Planck Equation, Fluctuation theorems, Turbulence, Stochastic Processes, Markov Processes.

1. Introduction

This open-source package is created for the analysis of turbulent velocity time series (one-dimensional). Although, it can be used for any other type of data containing a time series. Within this document, a detailed procedure to use this package will be discussed. The user is expected to have all the theoretical background associated with the turbulent cascade process, Fokker-Planck Equation (FPE), Integral fluctuation theorem (IFT) and its interpretation. Several research articles deal with this method for analyzing turbulent data. However, with the implementation in this open source package, for the first time the application in practice becomes easy and fast for everyone. For a quick and simple overview, we recommend the following annual review article [Peinke et al. \(2019\)](#). With this open source package one should be able to perform all analysis steps of this paper.

This package has been used continuously within our lab since 2018. It also has been successfully used by a large number of students (practical exercises which were part of the fluid dynamics lecture at the University of Oldenburg) to ensure stability across different machines and operating systems. This package has proved its value in a number of publications like:

- Reinke, Nico, et al. "On universal features of the turbulent cascade in terms of non-equilibrium thermodynamics." *Journal of Fluid Mechanics* 848 (2018): 117-153.
- Ali, Naseem, et al. "Multi-scale/fractal processes in the wake of a wind turbine array boundary layer." *Journal of Turbulence* 20.2 (2019): 93-120.
- Peinke, Joachim, MR Rahimi Tabar, and Matthias Wächter. "The Fokker–Planck approach to complex spatiotemporal disordered systems." *Annual Review of Condensed Matter Physics* 10 (2019): 107-132.
- Fuchs, André, et al. "A Rigorous Entropy Law for the Turbulent Cascade." *Turbulent Cascades II*. Springer, Cham, 2019. 17-25.
- Fuchs, André, et al. "Small scale structures of turbulence in terms of entropy and fluctuation theorems." *Physical Review Fluids* 5.3 (2020): 034602.
- Fuchs, André, et al. "The entropy and fluctuation theorems of inertial particles in turbulence." arXiv preprint arXiv:2104.03136 (2021).
- Kharache, Swapnil, et al. "Energy dissipation and total entropy production in SHREK experiment." *Progress in Turbulence VIII*. 2021. (in review process)

The package itself or parts of it, as well as results obtained by using the package, have also been presented at several international conferences, such as:

- Fuchs, André, et al. "A Rigorous Entropy Law for the Turbulent Cascade." 71rd Annual Meeting of the APS Division of Fluid Dynamics. 2018.
- Fuchs, André, et al. "Fine structure of turbulence determined by entropy variation." Workshop on different states of turbulence and transitions from one state to the other: small and large-scale aspects and their interrelations. Grenoble. 2019

- Fuchs, André, et al. "Fine structure of turbulence determined by particle image velocimetry & entropy variation." Workshop on turbulence in the context of Eulerian and Lagrangian views, with a particular focus on cascade features connected to stochastic processes, large deviation theory and instantons. Grenoble. 2019
- Fuchs, André, et al. "An open source Matlab package for solving Fokker-Planck Equation and validation of Integral Fluctuation Theorem." 73rd Annual Meeting of the APS Division of Fluid Dynamics. 2020.
- Kharche, Swapnil, et al. "Energy dissipation and total entropy production in SHREK experiment." Interdisciplinary Turbulence initiative Conference in Turbulence. 2021
- Fuchs, André, et al. "An open source Matlab package for solving Fokker-Planck Equation and validation of Integral Fluctuation Theorem." Interdisciplinary Turbulence initiative Conference in Turbulence. 2021
- Fuchs, André, et al. An open source Matlab package for stochastic data analysis in the context of Fokker-Planck Equation and Integral Fluctuation Theorem. No. EGU21-9608. European Geosciences Union General Assembly. 2021.

Note, this list is not complete, only a selection is listed here.

2. Implementation and architecture

The software is implemented in MATLAB® (2020a). The results presented in this paper were obtained using the same version. Before using this script the following toolboxes should be included in your MATLAB license.

- **Curve Fitting Toolbox**
- **Optimization Toolbox**
- **Parallel Computing Toolbox**
- **Signal Processing Toolbox**
- **Statistics and Machine Learning Toolbox**

As these MATLAB toolboxes are essential for the application of the package, the compatibility to Octave can not be provided. But to enhance the accessibility, standalone applications (64-bit) for Windows, macOS and Linux are also created to run the MATLAB code on target machines that do not have a MATLAB license.

To apply our method, two important features must be met. First, the set of data from which one extracts the sequence of velocity increments in scale must be a **stationary process** (performing multipoint analysis this condition is not always required). Second, the process in scale must be **Markovian**. These preconditions are tested during the execution of the program. Overall, the analysis consists of three main steps: pre-processing, estimation of Kramers-Moyal coefficients and estimation of the total entropy variation. The central part of this paper is the description of the main program, subroutines and selection of the relevant parameters. The following section will give an overview of all subroutines that are accessible

F_s	L	λ	<i>inc_bin</i>	Δ_{EM}
8000 Hz	0.067 m	0.0066 m	93	22 samples

Table 1: Information that the user must enter during the analysis to reconstruct the results shown in this paper. With the sampling frequency F_s , integral length scale L , Taylor length scale λ , number of bins to be used to divide the velocity increment series *inc_bin* and Einstein-Markov length Δ_{EM} .

in this package. Note that all the abbreviations used in this document are listed at the end in the list of nomenclature/abbreviations.

`main` is the primary script in which all the different functions to carry out various analyses are included. All subroutines, which will be briefly introduced in the next chapter, are called in a logical order. These subroutines are also accessible from the command line and can be included by the user in other applications.

3. Part I: Pre-processing

3.1. Exemplary dataset and system requirements

To demonstrate the application of this program an exemplary dataset obtained in a turbulent air jet experiment [Renner et al. \(2001\)](#) is used within this document. The local velocity was sampled by a time-resolved hot-wire measurement. The data acquisition comprises 1.25×10^7 samples at a sampling frequency of 8 kHz. The package additionally includes this exemplary dataset. We also provide the generated plots/results allowing the user to verify if the software is operating correctly. Table 1 lists all parameters that the user must enter during the analysis to reconstruct the results shown in this paper.

The system requirements (memory, disk space, processor, total run time) demanded by the script depend very much on the size of the data set to be examined, the resolution in relation to the number of bins as well as the available number of CPUs (required memory will be allocated to each worker during parallel computations).

3.2. Loading data and variables

`uiimport` is the first command in the program which asks the user to select interactively the data file which will be used for the analysis. The name of the variable/time series to be analyzed must be `data`. At this point, it is possible to specify the percentage of the total data that will be used to perform the analysis (for example, the first 20 % of the data). Note that this parameter has a significant effect on the overall performance of the script.

`save_path` opens a dialog box to navigate through the local directories in order to select a folder for saving figures and files.

`Fs` generates a pop-up dialog box to enter the sampling frequency of the data in Hz.

`kin_vis` generates a pop-up dialog box to enter the value of kinematic viscosity ν in m^2/s of the fluid for which the experimental data has been acquired.

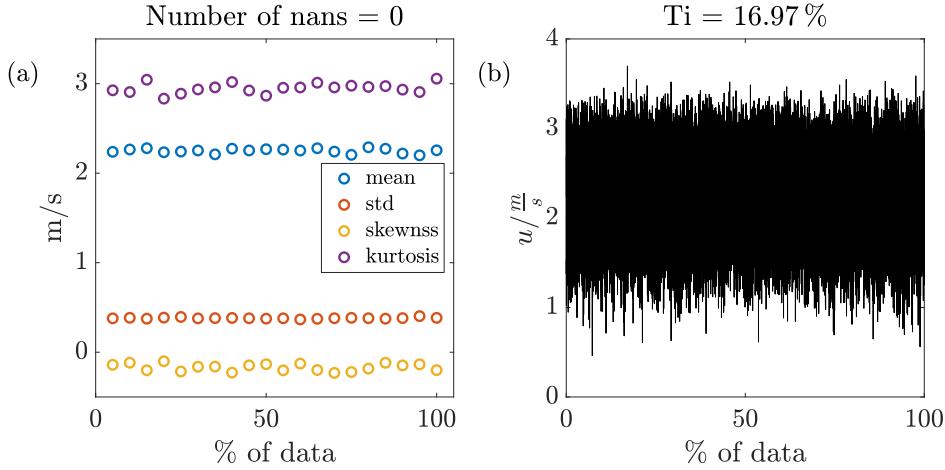


Figure 1: (a) plot of the mean, standard deviation, skewness and kurtosis of each section of a length of 5% of the data (b) plot of the data itself as a function % of the data.

`increment_bin` generates a pop-up dialog box to specify the number of bins to be used to divide the velocity increment series. A first estimation of the number of bins is made using

$$inc_bin = 10 \frac{\max(data) - \min(data)}{u'} \quad (1)$$

Note that this parameter has a significant effect on the overall performance of the script.

3.3. Test of stationarity and filtering of the time series

`plot_stationarity` This function plots the mean, standard deviation, skewness and kurtosis of each section of a length of 5% of the data and the data itself to check the stationarity condition in Figure 1. In the title of the figure, the number of nans, which will be removed from the time series and the turbulence intensity is printed.

`plot_pdf` This function plots the probability density function (PDF) of the data with the specified number of bins (specified in the function `increment_bin`) in Figure 2. It also plots the Gaussian distribution which has the same standard deviation and mean value as of the data. In the title of the figure the range of the data (the difference between the maximum and minimum values of sample data), the skewness and flatness of the data are printed.

`spectrum` This function calculates the energy spectral density (ESD) of the time series using `fft` function of MATLAB. Also, the ESD with and without averaging (moving average with equally spaced frequency interval in log-space) as a function of frequency will be plotted in Figure 3.

`low_freq` generates a pop-up dialog box to select whether the data should be filtered (low-pass filter) or not. If a filter is to be applied then in the next step the frequency in Hz at which the data will be filtered by a low-pass filter have to be specified (for example 1800 Hz). If the pop-up dialog box is denied, it is set to the value `low_freq=Fs/2`.

`frequency_filter` This function returns the filtered data and the filtered energy spectral density in the frequency domain (see Figure 4) using the low-pass filter at the previously set frequency using `butter` and `filtfilt` function of MATLAB.

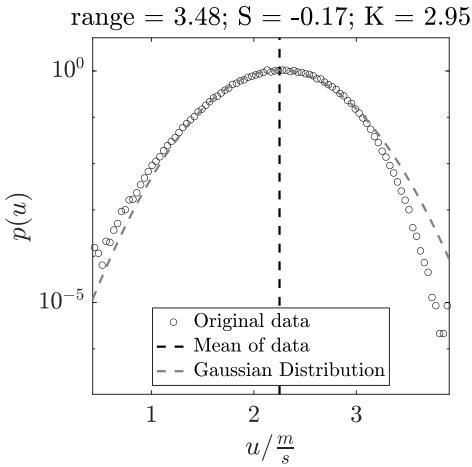


Figure 2: Probability density function (PDF) of the data. The grey dashed line corresponds to a Gaussian distribution with the same standard deviation and mean value (vertical black dashed line) as of the data.

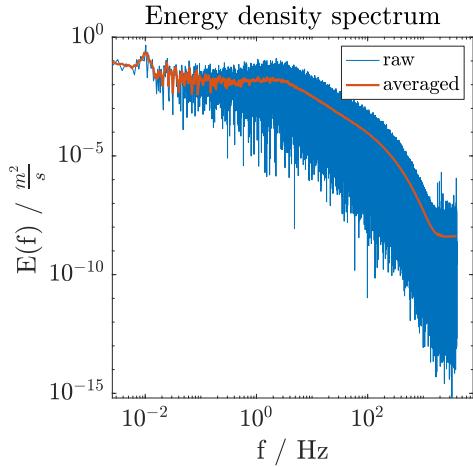


Figure 3: Energy spectral density (ESD) in the frequency domain. The red solid line corresponds to the averaged ESD as a function of frequency.

The filtered data named as **data_filter** will be used for all the further data post-processing. If the filtering was negated in the previous step, **data_filter** and **data** are equal. In addition, different representations/normalization of the energy respectively dissipation spectrum density with respect to frequency f , scale r , wave number k will also be plotted in Figure 5. The energy spectral density is normalized (Parseval's theorem) so that

$$u'^2 = \int_0^\infty E(k) dk \quad (2)$$

$$k = \frac{2\pi f}{\langle u \rangle} \quad (3)$$

$$E(k) = \frac{E(f) \langle u \rangle}{2\pi}. \quad (4)$$

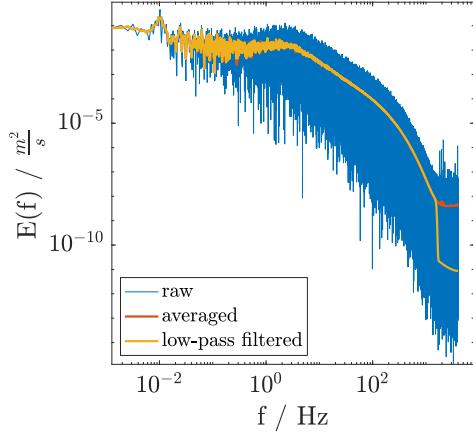


Figure 4: Energy spectral density (ESD) in the frequency domain. The yellow solid line corresponds to the averaged and filtered energy spectral density in the frequency domain using the low-pass filter.

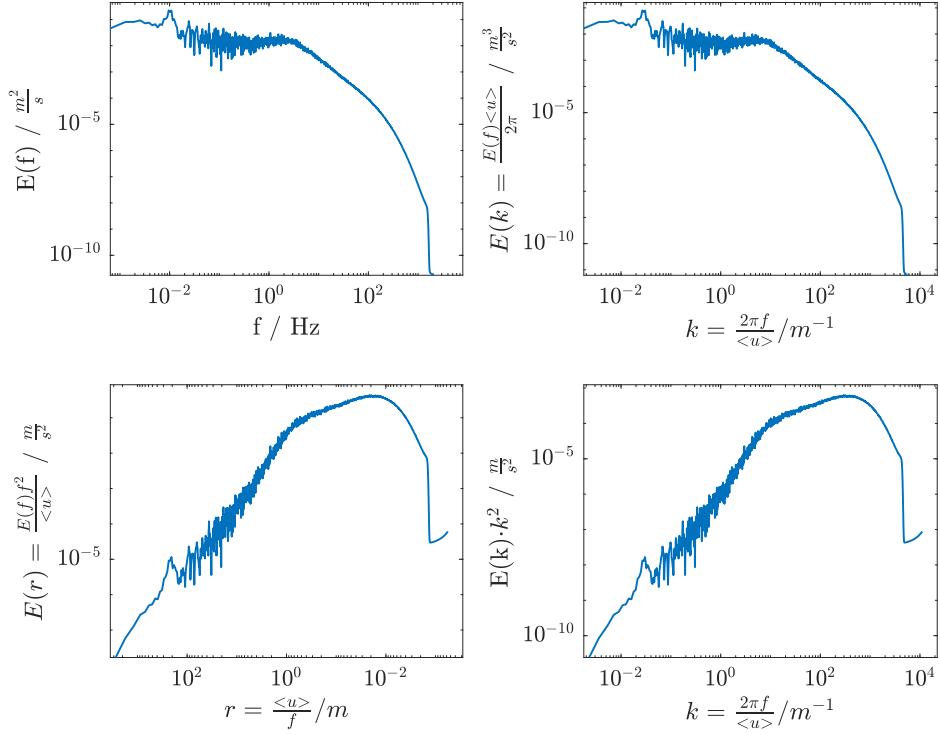


Figure 5: Different representation/normalization of the energy spectral density with respect to frequency f and wave number k and dissipation spectral density with respect to scale r and wave number k .

3.4. Estimation of fundamental turbulence length scales

Each scale can be represented either in terms of the number of samples or in meters. Although for physical interpretation, the units of scales can be taken as meters but for the post-processing of data, the number of samples is an easier choice. From data series of the longitudinal velocities $u(t)$ (the component in direction of the mean flow) we construct the data sets of their longitudinal increments

$$u_\tau(t) = u(t) - u(t + \tau) \quad (5)$$

labeled by the time-separation τ which, assuming the Taylor hypothesis of frozen turbulence [Taylor \(1938\)](#) (necessary condition: turbulence intensity is less than 20 %), the temporal information corresponds to spatial velocity increments, $u_r(t) = -u_\tau(t)$ with $r = -\tau\langle u \rangle$. For example consider a hot wire signal with a sampling frequency of 50 kHz which is characterized by a mean velocity of 10 m/sec and an integral length scale of 0.1 m. According to the basic equation of velocity = distance/time, the integral length scale of 0.1 m is equal to the turnover time of 0.01 sec or 500 samples.

length_scales In this function, the integral length scale L , Taylor length scale λ and Kolmogorov length scale η are estimated using different methods of calculation. Within this function, pop-up dialog boxes will be generated to enter the values of the integral, Taylor, Kolmogorov length scale in m and energy dissipation rate in m^2/s^3 based on which the further processing of data (solving the FPE and extracting cascade trajectories) should be referred. The entered length scales will be rounded towards the nearest integer sample. The proposed value in the pop-up dialog box is the median length scale for all methods.

The **integral length scale L** is estimated by using:

1. the **energy spectrum density** which requires the range of frequency that will be used to linearly extrapolate the value of ESD at a frequency of 0 Hz [Hinze \(1975\); Roach \(1987\)](#) (see Figure 6). Therefore the user will be asked to enter the **f_start** and **f_end** in Hz (**f_start < f_end**). (For example: $f_{\text{start}} = 0.03$ Hz and $f_{\text{end}} = 2.7$ Hz.)

$$L = \lim_{f \rightarrow 0} \left[\frac{E(f) \langle u \rangle}{4u'^2} \right] \quad (6)$$

2. from the **autocorrelation coefficient R_{uu}** with respect to scales [Frisch \(1995\); Pope \(2001\); Bourgoin et al. \(2017\)](#) plotted in Figure 7

$$L = \int_0^\infty R_{uu}(r) dr, \quad (7)$$

where r is the scale in meter. The cumulative integral gives the asymptotic value at a specific scale r which is characterized by the integral length scale L . This method, however, leads to large errors if the correlation does not decay exponentially. In the case of a non-monotonic decrease of R_{uu} , the autocorrelation function is

3. integrated up to the first zero-crossing of the autocorrelation function [O'Neill et al. \(2004\)](#)

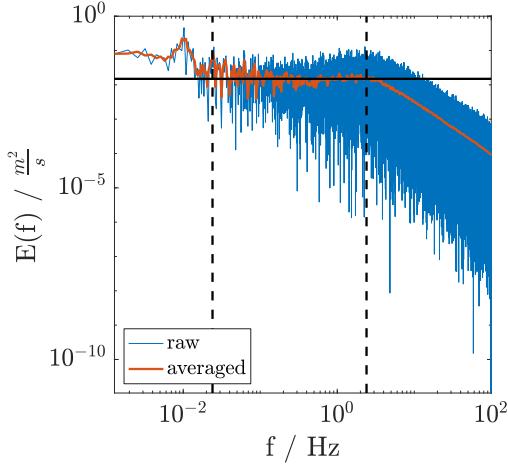


Figure 6: Representation of the linear extrapolation of ESD at a frequency of 0 Hz for estimating the integral length scale with the method according to Hinze (1975); Roach (1987). The two vertical dashed lines correspond to the range of frequency that will be used to linearly extrapolate (solid black line).

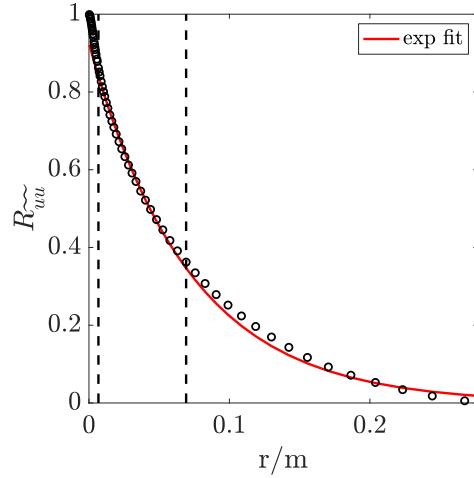


Figure 7: Representation of the autocorrelation coefficient $R_{uu}^{-\tilde{}}$ as a function of scale in meter. The two vertical dashed lines correspond to the range of scale that will be used for the extrapolation (red solid line) by an exponential function for estimating the integral length scale with the method according to Hinze (1975); Tritton (2012).

4. integrated up to the first 1/e crossing of the autocorrelation function Tritton (2012)
5. extrapolated by an exponential function Hinze (1975); Tritton (2012) (fit region: $r=[0 : r_e]$, see Figure 7), with $R_{uu}^{-\tilde{}}(r_e) = 1/e$. The integral can be solved analytically

$$L = \int_0^\infty ae^{-br} dr = \frac{a}{b}. \quad (8)$$

6. calculated from the second order structure function $S^2(r)$ which holds the relationship with autocorrelation function such as

$$R_{\tilde{u}\tilde{u}}(r) = 1 - \left[\frac{S^2(r)}{2u'^2} \right], \quad (9)$$

where u' can be calculated directly from the second order structure function. At a sufficiently large length scale compared to the energy injection scale of the experiment the second order structure function truncates to the asymptotic value of $2u'^2$ [Mordant et al. \(2001\)](#).

7. calculated using the `xcorr` inbuilt function of MATLAB using the ‘unbiased’ normalization option followed by the estimation of cumulative integral. The peaks in the cumulative integration of $R_{\tilde{u}\tilde{u}}$ with respect to r show the fact that autocorrelation function is fluctuating around a constant value. After this peak, the signal is considered to be uncorrelated and hence the scale at which this peak occurs can be taken as a measure of the integral length scale.
8. The integral length scale can be estimated via the zero crossings proposed by [Mora and Obligado \(2020\)](#). To verify that zero crossing are well resolved the signal has to be filtered with a low-pass filter with a characteristic length η_c . This method consists of estimating the Voronoi tessellation of the 1D zero-crossings positions data set. It is compared with a Random Poisson Process, which has no correlations between the scales. The method proposes that the integral length scale is equal to the characteristic length η_c for which

$$\frac{\sigma_{vor0}}{\sigma_{RPP}} = 1 \quad (10)$$

σ_{vor0} is the standard deviation of the Voronoi cells normalized by their mean value and σ_{RPP} is the equivalent value for an Random Poisson Process, that is equal to $\sqrt{(1/2)}$. In Figure 8 the standard deviation (normalized) of the Voronoi cells as a function of the characteristic length η_c is plotted. Finally, the integral length scale is defined as the value of η_c that correspond to $\sigma_{vor0}/\sigma_{RPP} = 1$. If we observe that $\sigma_{vor0}/\sigma_{RPP} > 1$ for all values of η_c , the method cannot provide the value of the integral length scale and longer signals are needed (nevertheless the extrapolation of the value remains possible).

The **Taylor length scale** λ is estimated by using:

1. A parabolic fit $R_{\tilde{u}\tilde{u}}(r) = 1 - \frac{r^2}{\lambda^2}$ to the autocorrelation function is used to estimate λ at the origin $r = 0$ (see Figure 9). The range of the positive curvature is therefore used for estimation (close to $r = 0$ the auto-correlation function has an inflection point). Since this method needs a well-resolved auto-correlation function it strongly depends on the resolution of the sensor.

Assuming isotropy at small scales and Taylor’s frozen turbulence hypothesis the **Taylor length scale** λ is estimated using the relation

$$\lambda^2 = \frac{u'^2}{\langle (\partial \tilde{u} / \partial x)^2 \rangle}. \quad (11)$$

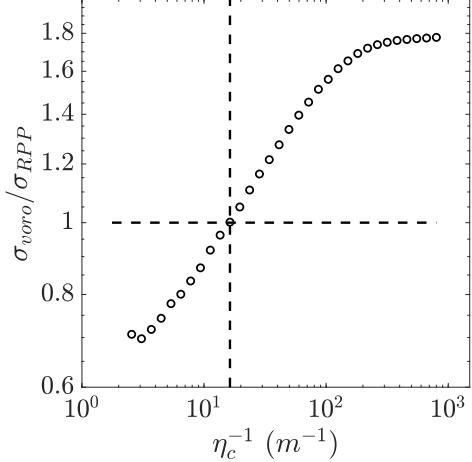


Figure 8: Standard deviation (normalized) of the Voronoi cells as a function of the characteristic length η_c of a low-pass filter for estimating the integral length scale with the method according to [Mora and Obligado \(2020\)](#).

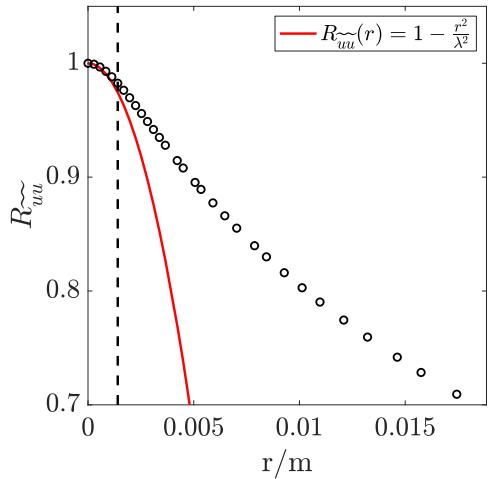


Figure 9: Autocorrelation function with parabolic fit as a function of the scale r for estimating the Taylor length scale. The vertical dashed line correspond to the range of scale that will be used for the extrapolation (red solid line) by a parabolic fit.

The numerical differentiation in the denominator is approximated by:

2. the simple difference quotient. Due to the finite spatial resolution of the measuring sensor, this method will yield an incorrect result. In order to correctly compute the velocity derivatives, the spatial resolution must be of the order of the Kolmogorov microscale (see [Hussein and George \(1990\)](#)).
3. the procedure proposed by [Aronson and Loefdahl \(1993\)](#).

$$\lambda = \lim_{r \rightarrow 0} \sqrt{\frac{u'^2 r^2}{S^2(r)}}. \quad (12)$$

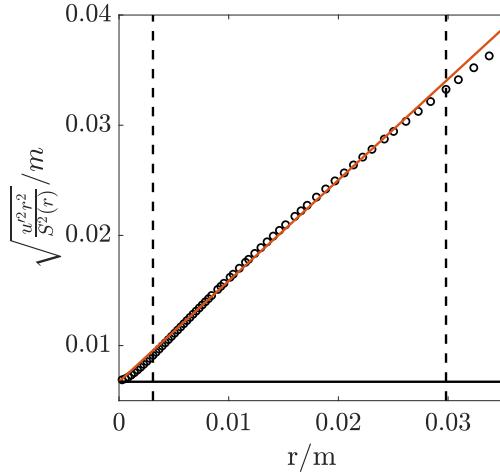


Figure 10: Development of Eq. 12 with linear extrapolation (red solid line) for estimating the Taylor length scale with the method according to Aronson and Loefdahl (1993). The two vertical dashed lines correspond to the range of scale that will be used for the extrapolation.

In Figure 10 the development of Eq. 12 as a function scale r is plotted. For the extrapolation we use a linear fit. The fit-region is used for spatial lags that are larger than 4 times the scale that corresponds to the low-pass filter frequency `low_freq`. The larger limit of fit-region must be set by the user.

4. using the dissipation spectrum. This procedure has been proposed by Hinze (1975). The upper limit of the integration is set to the cutoff frequency or the full (5. method) dissipation spectrum.

$$\left\langle \left(\frac{\partial \tilde{u}}{\partial x} \right)^2 \right\rangle = \int_0^\infty k^2 E(k) dk. \quad (13)$$

6. & 7. method: λ can be estimated via the zero crossings of the fluctuating velocity,

$$\lambda = \frac{l}{C\pi}, \quad (14)$$

with l the average distance between zero-crossings. C is a constant in the order of unity that quantifies the level of non-Gaussianity of the derivative $\partial \tilde{u} / \partial x$. This method was introduced by Mazellier and Vassilicos (2010); Sreenivasan *et al.* (1983); Mora *et al.* (2019). In Figure 11(a) the density of zero-crossings times the average distance between zero crossings as a function of the characteristic length η_c is plotted.

For values of η_c within the inertial range, a power law $2/3$ is expected Mazellier and Vassilicos (2010), and eventually for smaller filter sizes (or large $1/\eta_c$) a plateau is reached. The presence of this plateau, related to the dissipative range, implies that the density of zero-crossings n_s are well resolved, and therefore l can be deduced using the trivial relation $n_s \cdot l = 1$. If the plateau is not reached, small scales are not resolved and the method can not estimate the Taylor length scale. On the other hand, if after the plateau the value of $n_s \cdot l$ increases again, it means that the cut-off frequency is too

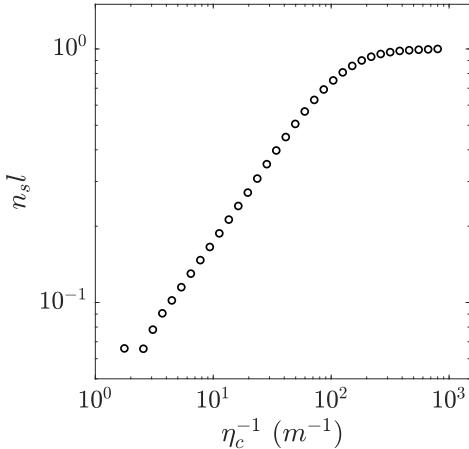


Figure 11: Density of zero-crossings times the average distance between zero crossings as a function of the characteristic length η_c of a low-pass filter for estimating the Taylor length scale with the method according to [Mazellier and Vassilicos \(2010\)](#); [Sreenivasan *et al.* \(1983\)](#); [Mora *et al.* \(2019\)](#).

high and the analysis is affected by small scale noise. Initially, the constant $C = 1$ (6. method) gives a good approximation of λ using Eq. 14. A better estimation can be obtained if $\partial \tilde{u} / \partial x$ is resolved. In that case, C is defined as

$$C = \sqrt{\frac{2}{\pi}} \frac{\sigma_{\partial_x \tilde{u}}}{\langle |\partial_x \tilde{u}| \rangle} \quad (15)$$

with $\langle |\partial_x \tilde{u}| \rangle$ is the mean and $\sigma_{\partial_x \tilde{u}}$ is the standard deviation of $\partial \tilde{u} / \partial x$, where \tilde{u} is filtered with the largest frequency within the plateau of $n_s \cdot l$.

Furthermore, this function estimates and returns the **mean energy dissipation rate** $\langle \epsilon \rangle$, which is estimated by its one-dimensional surrogate using

1. & 2. method: either 2nd or 3rd order structure function. The estimation of dissipation using this method relies on the transfer of energy within the inertial range. This method is particularly useful when the higher frequency content present in the flow is not fully resolved by the measurement device. This is generally the case where for example the length of the hot wire is larger than the Kolmogorov length scale η of the flow under consideration. This function generates a pop-up dialog box to enter the value of Kolmogorov constant C_2 (typically within 2.0 - 2.2) used in the relation between second order structure function $S^2(r)$ and $\epsilon(r)$ based on the assumption of homogeneous isotropic turbulence (HIT) [Pope \(2001\)](#); [Taylor \(1938\)](#). The mean energy dissipation rate $\langle \epsilon \rangle$ is calculated by finding the mean amongst 5 points closest to the peak value of $\epsilon(r)$. In Figure 12 the development of $\epsilon(r)$ using either 2nd or 3rd order structure function is plotted.

$$\epsilon(r) = \frac{1}{r} \left[\frac{S^2(r)}{C_2} \right]^{3/2} \quad (16)$$

$$\epsilon(r) = -\frac{5}{4} \left[\frac{S^3(r)}{r} \right] \quad (17)$$

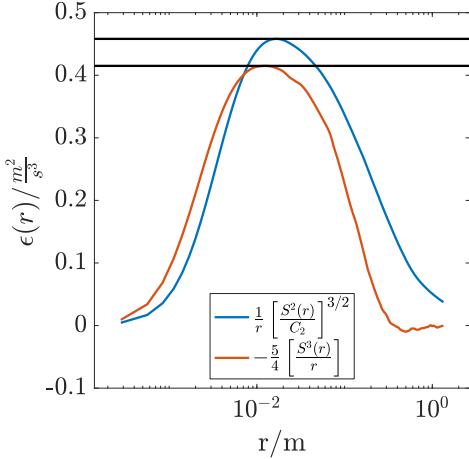


Figure 12: Development of $\epsilon(r)$ using either 2nd or 3rd order structure function (see Eq. 16 and 17). The black solid line marks the peak value of $\epsilon(r)$.

3. by using the chosen Taylor length scale and the following relation

$$\langle \epsilon \rangle = 15\nu \frac{u'^2}{\lambda^2}. \quad (18)$$

4. & 5. method: As a consequence of the K41 phenomenology, the second order structure function implies an energy spectrum of the form

$$E(k) = C_k \langle \epsilon \rangle^{2/3} k^{-5/3}. \quad (19)$$

C_k is the so-called Kolmogorov constant that remains undetermined in Kolmogorov's theory (typically $C_k \approx 0.53 \pm 0.01$ Sreenivasan (1995); Oboukhov (1962)). Following Kolmogorov's $k^{-5/3}$ prediction the two following fits in the inertial range ($\lambda < r < L$) are used to estimate the mean energy dissipation rate $\langle \epsilon \rangle$ (see Figure 13)

$$E(k) = C_k \langle \epsilon \rangle^{2/3} (k - k_0)^{-5/3} \quad (20)$$

$$E(k) = C_k \langle \epsilon \rangle^{2/3} k^{-5/3}. \quad (21)$$

6. via the dissipation spectrum

$$\langle \epsilon \rangle = \int_0^\infty 15\nu k^2 E(k) dk. \quad (22)$$

As proposed in Mora *et al.* (2019) the dissipation spectrum is modeled for large wave number k using a second order polynomial (see Figure 14).

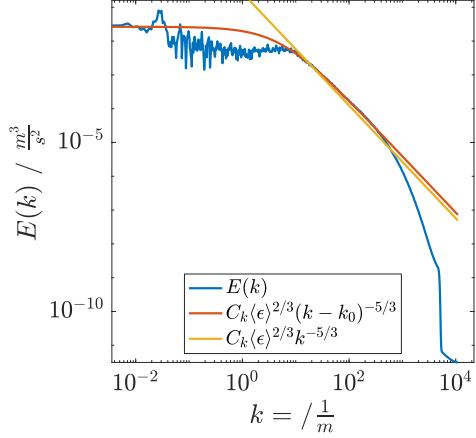


Figure 13: Energy spectral density (ESD) in the frequency domain. The red and yellow solid line corresponds to the fit according to equation 20 and 21 for estimating the mean energy dissipation rate $\langle \epsilon \rangle$.

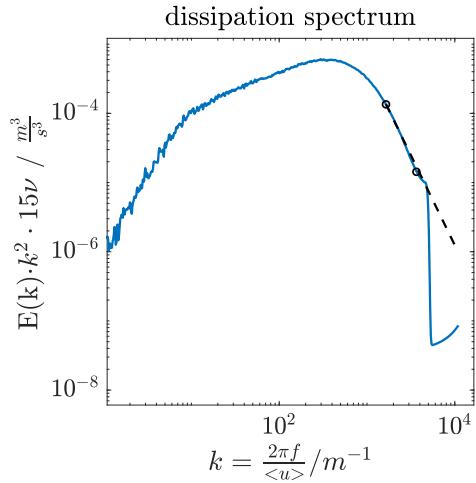


Figure 14: Dissipation spectral density in the wave number domain. The black dashed line corresponds to the fit (second order polynomial) to model the dissipation spectrum for large wave number k .

The **Kolmogorov length scale** η (the smallest size of the eddy in a given turbulent flow) is estimated by using the classical relation given by [Frisch \(1995\)](#)

$$\eta = \left(\frac{\nu^3}{\langle \epsilon \rangle} \right)^{1/4}. \quad (23)$$

In addition, the normalized energy dissipation rate will be returned [Batchelor and Townsend \(1948\); Tennekes and Lumley \(1972\)](#)

$$C_{\epsilon_1} = \frac{\langle \epsilon \rangle L}{u'^3} \quad (24)$$

$$C_{\epsilon_2} = 15 \frac{L}{\lambda} \frac{1}{Re_\lambda}, \quad (25)$$

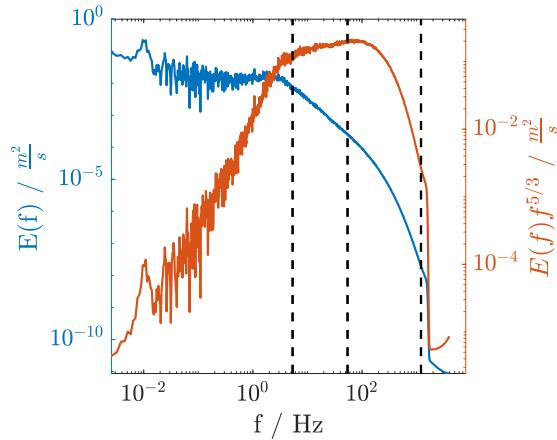


Figure 15: ESD and compensated ESD in the frequency domain. Black dashed vertical lines correspond to the respective frequency of the fundamental length scales.

with the local **Taylor-Reynolds number**

$$Re_\lambda = \frac{u' \lambda}{\nu}. \quad (26)$$

Re_λ allows for reasonable comparisons of experiments with different boundary conditions or energy injection mechanisms (Re_λ is independent of the integral length scale L).

At the end of this function, in Figure 15 a vertical dashed line at the integral L , Taylor λ and Kolmogorov length scale η will be added to different representations of the spectrum.

3.5. Normalization of the data

struc_flip_test This function tests whether the data have to be flipped or not. The decision of flipping of data depends on a simple relation of 3rd order structure function $S^3(r)$ with the dissipation based on the assumption of homogeneous isotropic turbulence (HIT). The thumb rule is that the quantity $S^3(r)$ must be negative. In the literature, the keyword that goes with this picture is vortex stretching. To verify this, $S^3(r)$ as a function of the scale r is plotted, from which it is possible to decide whether it is essential to flip the data or not.

normalization This function is mainly to perform the normalization of the data. Before doing so, it generates the pop-up dialog box which asks the user whether to flip the data or not (based on the previous investigation). After that, this function normalizes the entire data with the quantity of $\sigma_\infty = \sqrt{2}\sigma$, where σ is the standard deviation of the **data_filter** (this method is proposed by [Renner et al. \(2001\)](#)). This function returns the filtered and normalized data as **data_filter**, **siginf** = σ_∞ and **m_data** = mean of the data before normalization. In addition the scale r is given in units of Taylor length scale λ . We use this normalization to compare the results of different data sets.

$$u_r = \frac{u_{r_i}}{\sigma_\infty} \quad (27)$$

$$r = \frac{r_i}{\lambda} \quad (28)$$

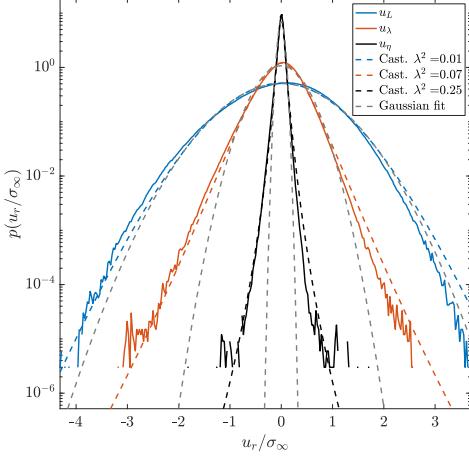


Figure 16: PDF of the velocity increments u_r at the scale $r = L, \lambda, \eta$. The colored dashed line correspond to Castaing fits (form factor λ^2 [Castaing *et al.* \(1990\)](#)) and grey dashed line to Gaussian fits.

This normalization affects also the Kramers-Moyal coefficients (KMCs), which will be introduced in Chapter: Estimation of Kramers-Moyal coefficients (KMCs). The indices i describes the initial parameter of the Kramers-Moyal coefficients without normalization:

$$D^{(1)} = \frac{D_i^{(1)} \lambda}{\sigma_\infty} \quad (29)$$

$$D^{(2)} = \frac{D_i^{(2)} \lambda}{\sigma_\infty^2} \quad (30)$$

$$d_{11} = d_{11_i} \lambda \quad (31)$$

$$d_{20} = \frac{d_{20_i} \lambda}{\sigma_\infty^2} \quad (32)$$

$$d_{21} = \frac{d_{21_i} \lambda}{\sigma_\infty} \quad (33)$$

$$d_{22} = d_{22_i} \lambda. \quad (34)$$

plot_increment_pdf This function plots in Figure 16 the probability density function of the velocity increments at the scale $r = L$, $r = \lambda$ and $r = \eta$.

plot_struct_function This function plots in Figure 17 the k -th order structure function

$$S^k(r) = \langle u_r^k \rangle \quad (35)$$

$$T^k(r) = \langle |u_r|^k \rangle \quad (36)$$

with $k = 2 - 7$ for scales $\lambda \leq r \leq L$. In addition this function plots in Figure 18 the scaling exponent ζ_k

$$S^k(r) \propto r^{\zeta_k} \quad (37)$$

estimated using the extended-self similarity (ESS) method. The scaling of a selected set of known intermittency models is also included for comparison. For this, the user is asked to

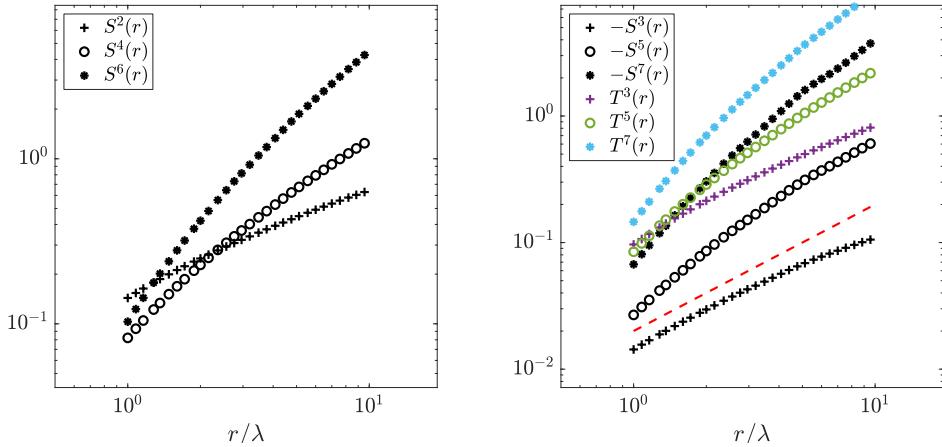


Figure 17: Course of the k -th order structure function S^k and T^k with $k = 2 - 7$ as a function of scale. Dashed red line represents $-4/5$ law.

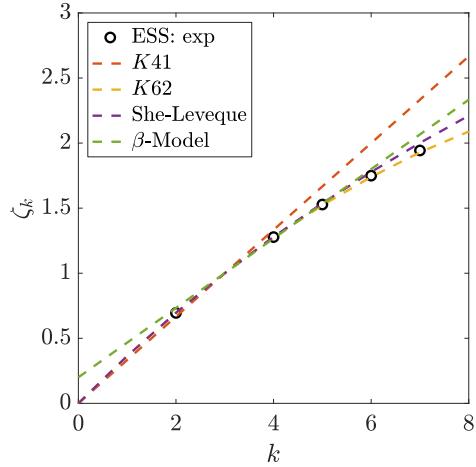


Figure 18: Course of scaling exponent ζ_k as a function of order k of structure function. Dashed lines represent the scaling of a selected set of known intermittency models.

specify the intermittency coefficient μ (experiments suggest a value of $\mu \approx 0.227$ Frisch (1995) and $\mu \approx 0.26$ Arneodo *et al.* (1996)) and the coefficient D of β -Model Anselmet *et al.* (1984).

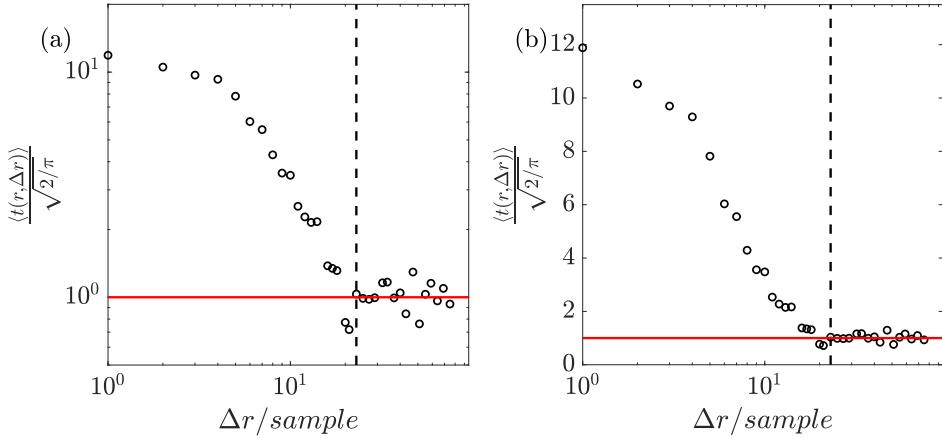


Figure 19: Development of $\langle t(r, \Delta r) \rangle / \sqrt{2/\pi}$ (see Lueck *et al.* (2006) for details) as function Δr in terms of samples (a) on log-log scale and (b) lin-log scale. The vertical dashed line corresponds to the Taylor length scale λ in terms of samples.

4. Part II: Estimation of Kramers-Moyal coefficients

4.1. Examination of the Markov Property/Determination of the Markov-Einstein Length

wilcoxon_test This function determines the Einstein-Markov length Δ_{EM} Renner *et al.* (2001). Above this length scale, the Markov properties hold and below this length scale, the Markov properties cease to hold. The Wilcoxon test (see Figure 19) is a parameter-free procedure to compare two empirically determined probability distributions (two data sets of velocity increments) (see Lueck *et al.* (2006) for details). It is a quantitative test that determines the Δ_{EM} . A sufficient resolution in measurement below Taylor's length scale is expected to perform this test.

markov generates a pop-up dialog box to enter the number of samples which corresponds to Einstein-Markov length. Experimental evidence shows that the Markov property can be assumed to hold for the cascade process coarse-grained by the Einstein-Markov length Lueck *et al.* (2006); Renner *et al.* (2001) (proposed value in the pop-up dialog box)

$$\Delta_{EM} \approx 0.9\lambda. \quad (38)$$

Note, if the resolution of the used sensor ceases at Δ_{EM} , it is possible to enter the number of samples which correspond to a larger scale than Δ_{EM} at which the data might be resolved in scale (for example samples corresponding to λ or 5λ). In addition, a red vertical dashed line at the Einstein-Markov length will be added to the spectrum in the frequency domain.

min_events generates a pop-up dialog box to enter the minimum number of events/counts to occur in a single bin, which will make that specific bin valid for further processing. If the minimum number of events is equal to 400 all the information in those specific bins in which the number of events/counts is less than 400 will be excluded for further post-processing of data. The provision of the minimum number of events/counts is for avoiding the appearance of noise and hence for better fitting of parameters. Based on the experience, we have fixed

the minimum value of the `min_events` to 400. Based on the length of the data and the statistics, it is possible to increase/decrease this number.

`conditional_PDF_markov` This function performs a qualitative/visual check for the validation of Markov property based on the alignment or misalignment of the single conditioned and double conditioned PDFs of velocity increments for three different scales $r_0 > r_1 > r_2$ each of which is separated by Δ_{EM} . To do this, a pop-up dialog box is generated to enter the conditioned value for large scale increment u_{r_0} , for example $u_{r_0} = \pm 1$. Note, the condition $u_{r_0} = 0$ corresponds to the maximum number of statistics. This function also plots various representations of the single and double conditioned PDFs (shown in Figure 20 are only two). If there is not a good agreement between the single conditioned and double conditioned PDF of velocity increments, it is possible to modify the Einstein-Markov length and/or the minimum number of events and repeat this qualitative/visual check for the validation of Markov property.

Note, the results presented hereafter are related to a modified number of bins (changed from 93 to 201). The number of bins has been adjusted to get a more detailed view of the following figures (conditional PDFs and the Kramers-Moyal coefficients). These detailed illustrations will be described in this readme file in an exemplary manner. A smaller number of bins leads to slightly different results, but the general trend remains the same. Furthermore, the results presented hereafter relating to the Fokker-Planck analysis using the multiscaling approach (not multipoint analysis).

4.2. Estimation of conditional moments

`scale_steps` This pop-up dialog box calculates the possible number of steps between the integral length scale and Taylor length scale which are separated by the Markov length. For these steps, the Kramers-Moyal coefficients (KMCs) will be estimated.

`multi_point` In this pop-up dialog box, it must be selected whether to perform the multipoint analysis or not (see annual review article [Peinke et al. \(2019\)](#)). To do the Fokker-Planck analysis using the multiscaling approach, the question in this pop-up dialog box must be denied. If multipoint analysis should be performed, an additional condition on the increment must be specified in the next pop-up dialog box.

`conditional_moment` This function estimates the k -th order conditional moment

$$M^{(k)}(u_r, r, \Delta r) = \int_{-\infty}^{\infty} (u_{r'} - u_r)^k p(u_{r'}|u_r) du_{r'}, \quad (39)$$

$k = 1 - 4$ for all scales $2\Delta_{EM} < r \leq L$ and for each bin (specified in the function `scale_steps` and `increment_bin`) for all values of longitudinal velocity increments u_r . For a fixed scale r the conditional moments are calculated for 5 different scales separations (colored circles in Figure 21) $\Delta r = r - r'$ within the range of $\Delta_{EM} \leq \Delta r \leq 2\Delta_{EM}$. The condition $r' < r$ is fulfilled.

`plot_conditional_moment` This function plots in Figure 21 the first and second conditional moments $M^{(1,2)}(u_r, r, \Delta r)$ as a function of the scale separation Δr . For this purpose, a scale r and the number of a bin (value of the velocity increment u_r) condition must be specified. The proposed value in the pop-up dialog box is $r = L$ and $u_r \approx 0$. A possible deviation from a linear law for small values of Δr is due to the Einstein-Markov length, as the Markov properties cease to hold for very small scale separations.

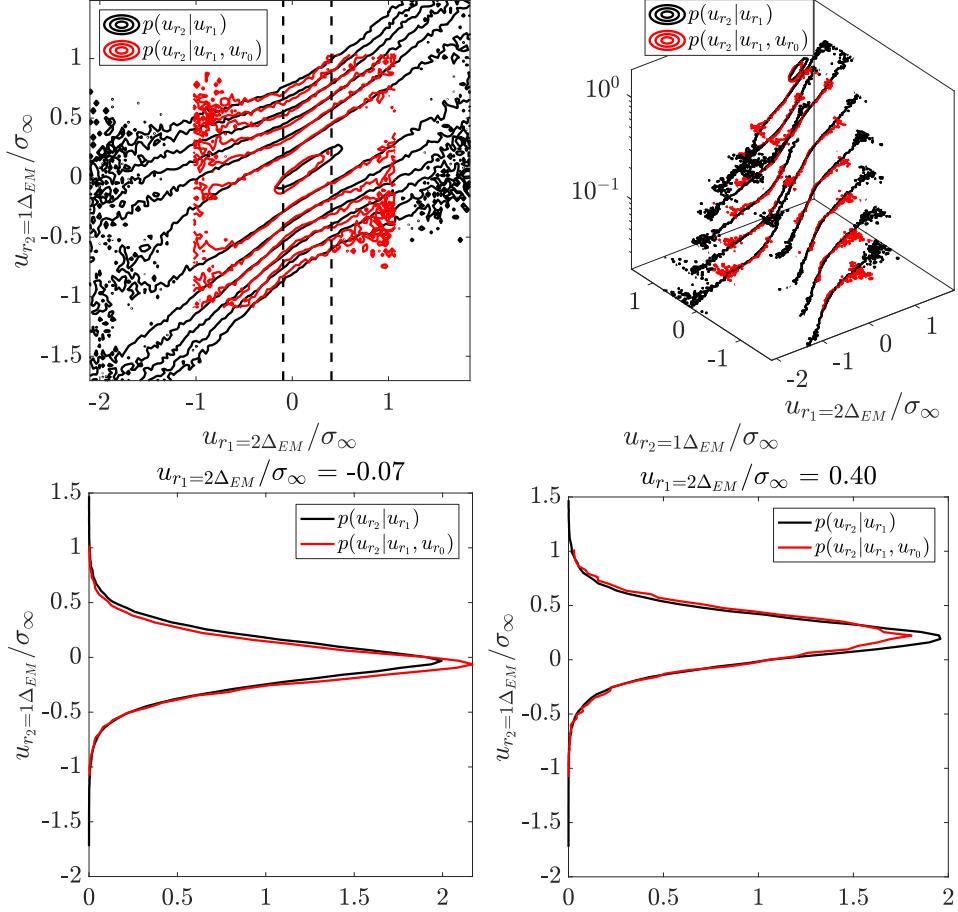


Figure 20: Visualization of Markov properties. Top: Contour plots showing single (black solid lines) and double conditioned PDFs (red solid lines, $u_{r_0} = 0$) of velocity increments for three different scales $r_0 > r_1 > r_2$ each of which is separated by Δ_{EM} . The right figure is a three-dimensional view. The dashed black lines in the top figure correspond to cut through the single and double conditioned PDFs at the marked points of u_{r_2} . Bottom: Cut through the single (black) and double (red) conditioned PDFs at the marked points of u_{r_2} .

4.3. Estimation of Kramers-Moyal coefficients (KMCs)

KM_Calculation This function calculates the Kramers-Moyal coefficients $D^{(k)}(u_r, r)$ with $k = 1 - 4$ for all scales (specified in `scale_steps`) and for each bin (specified in the function `increment_bin`) for all values of velocity increments by a linear extrapolation in Δr of the k -th order conditional moments $M^k(u_r, r)$ (see Figure 21) and the function `KM_plot_raw` plots them accordingly. With $r' < r$:

$$D^{(k)}(u_r, r) = \lim_{r' \rightarrow r} \frac{M^{(k)}(u_r, r, \Delta r)}{k! (r' - r)}. \quad (40)$$

This limit approximation leads to uncertainties in the absolute values of the Kramers–Moyal coefficients, whereas the functional forms of $D^{(k)}(u_r, r)$ are commonly well estimated. In order to overcome this problem, the optimization algorithm described below is performed.

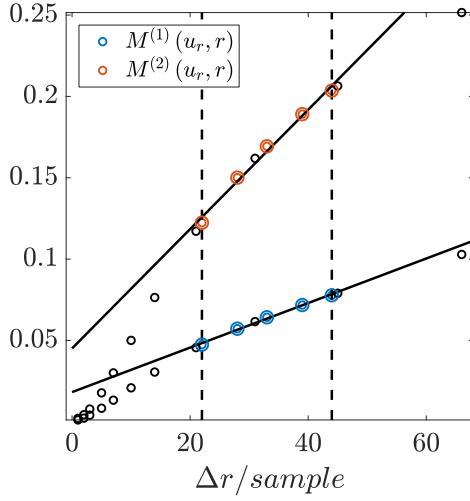


Figure 21: First and second conditional moments $M^{(1,2)}(u_r, r, \Delta r)$ as a function of the scale separation Δr . In addition, a linear extrapolation in Δr (solid black line) of the first and second order conditional moments is plotted (see Chapter: Estimation of Kramers-Moyal coefficients). The vertical dashed lines and the colored circles limit the range used for the linear fit ($\Delta_{EM} \leq \Delta r \leq 2\Delta_{EM}$).

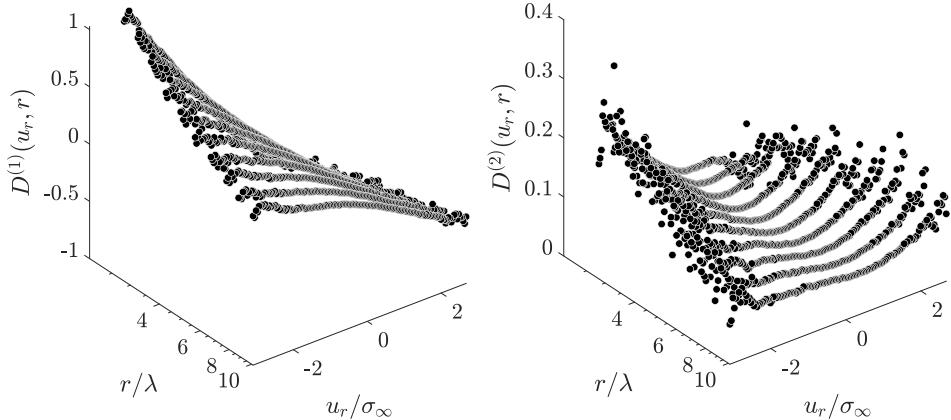


Figure 22: Non-optimized Kramers-Moyal coefficients $D^{(1,2)}(u_r, r)$ with respect to scale r (with $L \approx 10r/\lambda$) and velocity increment obtained by the linear extrapolation method.

4.4. Pointwise optimization of Kramers-Moyal coefficients: conditional PDF

KM_STP_optimization This function performs the pointwise optimization of Kramers-Moyal coefficients $D^{(1,2)}(u_r, r)$ at each scale and value of velocity increment to minimize possible uncertainties in the absolute values of the Kramers-Moyal coefficients. The purpose of this optimization is to find the best Fokker-Planck equation to reproduce the conditional PDFs as these are the essential part of the Markov process. This optimization procedure is proposed in [Kleinhans et al. \(2005\)](#); [Nawroth et al. \(2007\)](#); [Reinke et al. \(2018\)](#) and it includes the reconstruction of the conditional probability density functions $p(u_r'|u_r)$ via the short time

propagator Risken (1996) with $r' < r$

$$p_{stp}(u_{r'}|u_r) = \frac{1}{\sqrt{4\pi D^{(2)}(u_r, r)\Delta r}} \exp\left(-\frac{(u_{r'} - u_r - D^{(1)}(u_r, r)\Delta r)^2}{4D^{(2)}(u_r, r)\Delta r}\right). \quad (41)$$

The scale step size $\Delta r = \Delta_{EM}$ leads to consistent results. Smaller steps than Δ_{EM} do not significantly improve the results. The aim of this optimization is to minimize a weighted mean square error function in logarithmic space Feller (1968) (analogous to Kullback–Leibler entropy)

$$\xi = \frac{\sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} (p_{exp} + p_{stp}) (\ln(p_{exp}) - \ln(p_{stp}))^2}{\sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} (p_{exp} + p_{stp}) (\ln^2(p_{exp}) + \ln^2(p_{stp}))}. \quad (42)$$

This error function is a logarithmic measure of the difference between the experimental p_{exp} and reconstructed p_{stp} conditional probability density function. The optimization procedure systematically changes $D^{(1,2)}(u_r, r)$ until the error function is minimized. This optimization use the function `fmincon` implemented in MATLAB. The constraints were set in a physically and mathematically meaningful way: $d_{11} \leq 0$, $d_{20} \geq 0$ and $d_{22} \geq 0$ (see for more details `FIT_KM`).

In addition, this function generates a pop-up dialog box whether an example optimization is required. If the pop-up dialog box is denied, then this function straightaway performs the optimization for all scales and all velocity increments without plotting anything. If this is confirmed, then the conditional PDFs will be plotted using different representations (shown here are only two see Figure 23) to see the differences between optimized and non-optimized and experimental conditional PDFs. Note, if the variable `scale_steps` is equal to 9 then it is possible to enter any scale number from 1 up to scale number 9 (smallest respectively the largest scale).

`FIT_KM` This function performs the surface fit with a linear function for $D^{(1)}(u_r, r)$ and a parabolic function for $D^{(2)}(u_r, r)$ to the optimized and non-optimized KMCs interpreted in the Itô convention Gardiner (2009)

$$D^{(1)}(u_r, r) = d_{11}(r)u_r, \quad (43)$$

$$D^{(2)}(u_r, r) = d_{22}(r)u_r^2 + d_{21}(r)u_r + d_{20}(r). \quad (44)$$

Coefficients $d_{ij}(r)$ in the fits are functions of scale r of the form $\alpha(r/\lambda)^\beta + \gamma$. After fitting, this function plots the parameters d_{11} , d_{20} , d_{21} and d_{22} as a function of $\frac{r}{\lambda}$ for optimized and non-optimized $D^{(1,2)}(u_r, r)$.

5. Part III: Estimation of total entropy variation

In the remaining part of the script, the calculation leading towards the integral fluctuation theorem will be done. In the spirit of non-equilibrium stochastic thermodynamics Seifert (2012) it is possible to associate with every individual cascade trajectory $[u(\cdot)]$ a total entropy variation ΔS_{tot} . In this investigation it is assumed that a single cascade trajectory represents one realization of the turbulent cascade process and a large number of these trajectories reflect the statistics caused by the process.

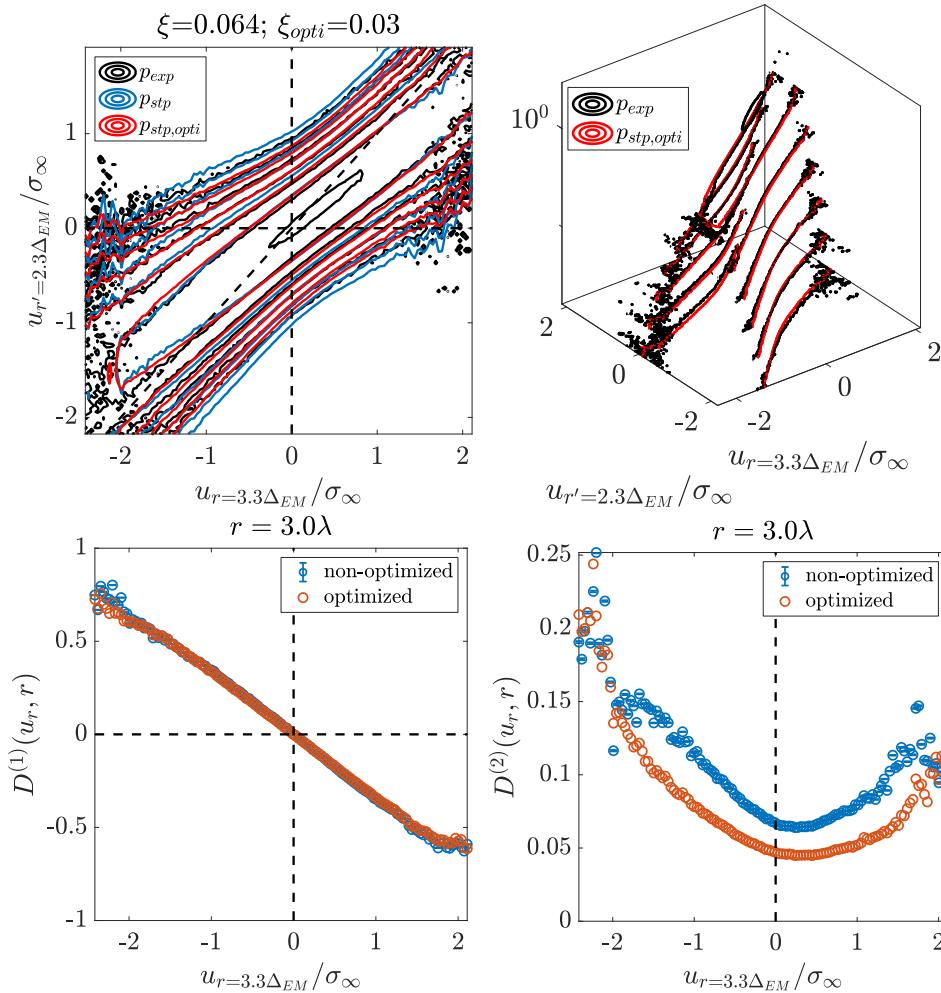


Figure 23: Top: Contour plots showing experimental p_{exp} , non-optimized p_{stp} and optimized conditional PDFs $p_{stp,opti}$ using the short time propagator of velocity increments for a pair of two scales with $r' < r$ each of which is separated by Δ_{EM} . The right figure is a three-dimensional view (only p_{exp} and $p_{stp,opti}$ is shown). Bottom: Non-optimized and optimized Kramers-Moyal coefficients $D^{(1,2)}(u_r, r)$ with respect to velocity increment u_r for a fixed scale $r = 3\lambda = 3.3\Delta_{EM}$ obtained by the optimization algorithm described above.

5.1. Validation of Integral Fluctuation Theorem

trajec Based on velocity increments u_r the cascade trajectories $[u(\cdot)] = \{u_L, \dots, u_\lambda\}$ for different scales from the integral length L to the Taylor length λ can be extracted from the data series of velocities $v(t)$. A pop-up dialog box is generated to select if the start and end of the cascade trajectory should be adjusted. If the pop-up dialog box is denied, then $[u(\cdot)]$ start at the integral length L and end at the Taylor length λ . If this is confirmed, at the beginning of function `checkFT` a pop-up dialog box is generated to specify whether the start and/or the end of the cascade trajectory should be adjusted in multiples of integral respectively Taylor length scale.

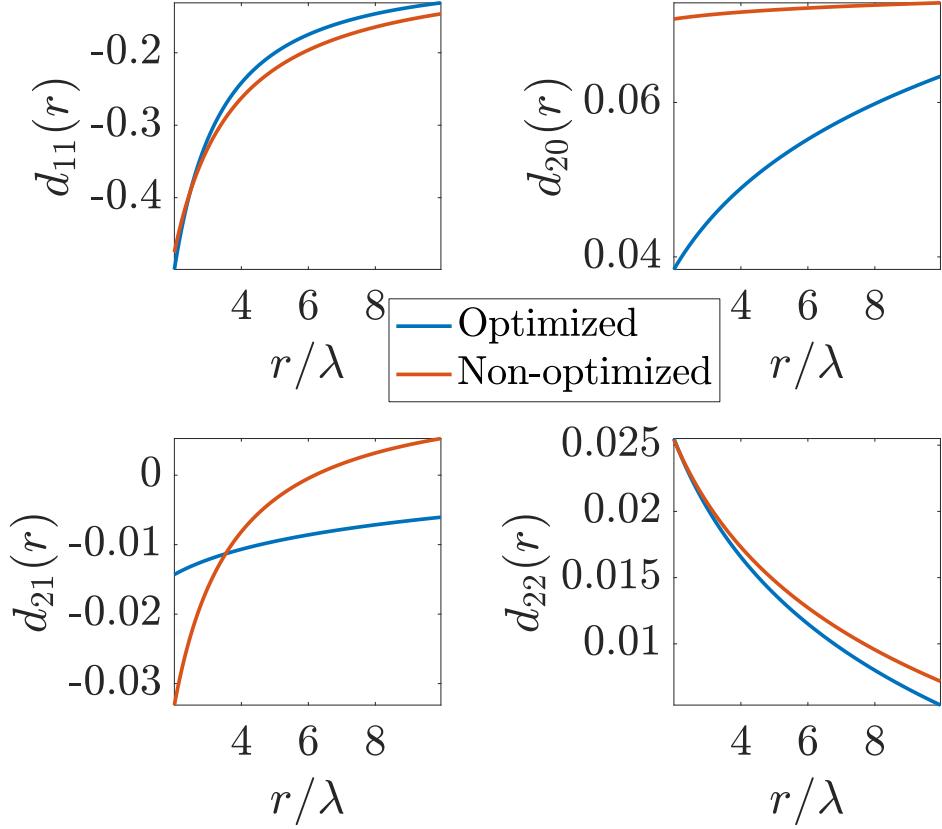


Figure 24: Coefficients $d_{ij}(r)$ of the optimized Kramers-Moyal coefficients using the surface fits with a linear function for $D^{(1)}(u_r, r)$ and a parabolic function for $D^{(2)}(u_r, r)$ (see Eq. 43 - 44) with respect to scale.

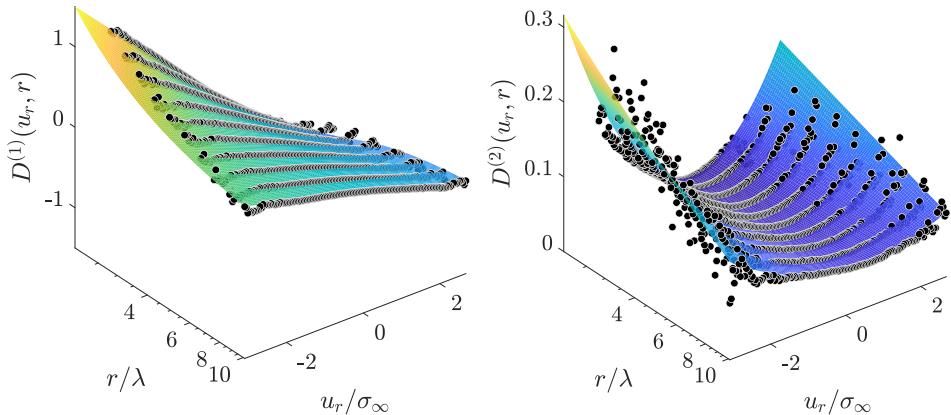


Figure 25: Optimized Kramers-Moyal coefficients $D^{(1,2)}(u_r, r)$ and the surface fits with a linear function for $D^{(1)}(u_r, r)$ and a parabolic function for $D^{(2)}(u_r, r)$ (see Eq. 43 - 44) with respect to scale r and velocity increment u_r .

In addition a pop-up dialog box is generated to select whether the total entropy variation should be calculated for overlapping (**z=1**) or independent cascade trajectories (**z=3**).

dr_ind A pop-up dialog box is generated to define the separation of scales/step increment (in samples) referred to the sequence from large to small scales in the cascade trajectory. The proposed value in the pop-up dialog box is equal to the Einstein-Markov length Δ_{EM} .

data_length A pop-up dialog box is generated to select the percentage of the data length that should be used to perform the calculation of the total entropy variation (for example, the first 20 % of the data).

checkFT The set of measured cascade trajectories results in a set of total entropy variation values ΔS_{tot} (the same number of entropy values as the number of trajectories) given by Seifert (2005, 2012); Sekimoto (2010); Nickelsen and Engel (2013); Reinke *et al.* (2018). This function calculates the system entropy

$$\Delta S_{sys} [u(\cdot)] = -\ln \left(\frac{p(u_\lambda, \lambda)}{p(u_L, L)} \right), \quad (45)$$

medium entropy

$$\begin{aligned} \Delta S_{med} [u(\cdot)] &= - \int_L^\lambda \partial_r u_r \partial_{ur} \varphi(u_r) dr \\ &= + \int_L^\lambda \partial_r u_r \frac{D^{(1)}(u_r, r) - \partial_{ur} D^{(2)}(u_r, r)/2}{D^{(2)}(u_r, r)} dr \end{aligned} \quad (46)$$

and the total entropy variation

$$\Delta S_{tot} [u(\cdot)] = \Delta S_{sys} + \Delta S_{med} \quad (47)$$

for all the independent cascade trajectories. The numerical differentiation is approximated by the central difference quotient:

$$\partial_r u_r = \lim_{r' \rightarrow r} \frac{u_{r'} - u_r}{r' - r} \quad (48)$$

This numerical differentiation is performed for every individual extracted cascade trajectory $[u(\cdot)]$ in a sequence from large to small scales. The integration in scale is approximated by using rectangles and a mid-point rule discretization of the scale intervals, therefore the integral takes the average of beginning and end of the discretization interval. The probabilities of starting and ending of the cascade trajectories, u_L and u_λ , can be estimated from the given data. The results depend slightly on the discretization rules and convention. However, the overall statements do not depend on it.

plot_entropy This function plots in Figure 26 the empirical average $\langle e^{-\Delta S_{tot}} \rangle_N$ of ΔS_{tot} as a function of the number, N (sample size), of cascade trajectories $[u(\cdot)]$ with error bars. In addition, the probability density function of the system, medium and total entropy will be plotted while displaying the value of $\langle \Delta S_{tot} \rangle$ which should be larger than 0. The integral fluctuation theorem (IFT) expresses the integral balance between the entropy-consuming ($\Delta S_{tot} < 0$) and the entropy-producing ($\Delta S_{tot} > 0$) cascade trajectories and states

$$\langle e^{-\Delta S_{tot}} \rangle_{[u(\cdot)]} = \int e^{-\Delta S_{tot}} p(\Delta S_{tot}) d\Delta S_{tot} = 1. \quad (49)$$

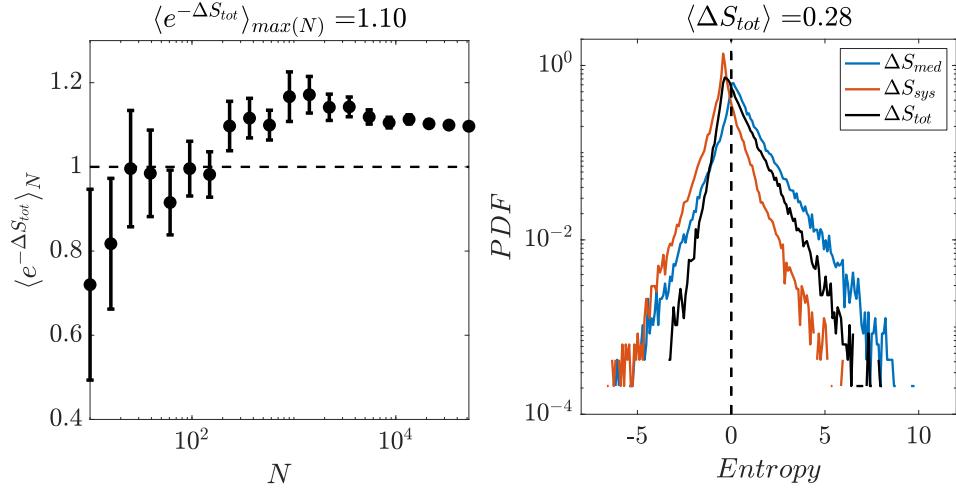


Figure 26: Empirical average $\langle e^{-\Delta S_{tot}} \rangle_N$ of ΔS_{tot} as a function of the number N (sample size) of cascade trajectories $[u(\cdot)]$ with error bars. According to the integral fluctuation theorem (IFT), the empirical average has to converge to the horizontal dashed line. Probability density function of the system S_{sys} , medium S_{med} and total entropy variation S_{tot} .

5.2. Pointwise optimization of Kramers-Moyal coefficients: IFT

In the remaining part of the script, the pointwise optimization of Kramers-Moyal coefficients towards the integral fluctuation theorem will be done. Thereby the separation of scales/step increment (in samples) referred to the sequence from large to small scales in the cascade trajectory is set to a minimum of 1 sample and for the optimization, independent cascade trajectories ($z=3$) are used. Note, we use here a separation that is less than or equal to the Einstein-Markov length of Δ_{EM} .

iter this pop-up dialog box is generated to enter the maximum number of iteration which will be performed for the optimization.

tol_D1, tol_D2 this pop-up dialog box is generated to specify the constraints/tolerance in percent of the coefficients $d_{ij}(r)$ which will be used to perform the optimization.

OPTI_IFT_dij This function performs the optimization of $D^{(1,2)}(u_r, r)$ at each scale and at each value of velocity increment in order to satisfy the integral fluctuation theorem with minimum possible error and plots the optimized d_{ij} as a function of scale (see Figure 27). The optimization procedure systematically changes $D^{(1,2)}(u_r, r)$ until the error function

$$\xi = |1 - \langle e^{-\Delta S_{tot}} \rangle_{max(N)}| \quad (50)$$

is minimized. Within the optimization process, the user is asked which $d_{ij}(r)$ should be optimized. This optimization use the function **fmincon** implemented in MATLAB.

Using the function **checkFT** and **plot_entropy** with **dr_ind=1** and overlapping cascade trajectories (**z=1**) and the optimized Kramers-Moyal coefficients the results presented in Figure 28 are obtained for the calculation of ΔS_{tot} .

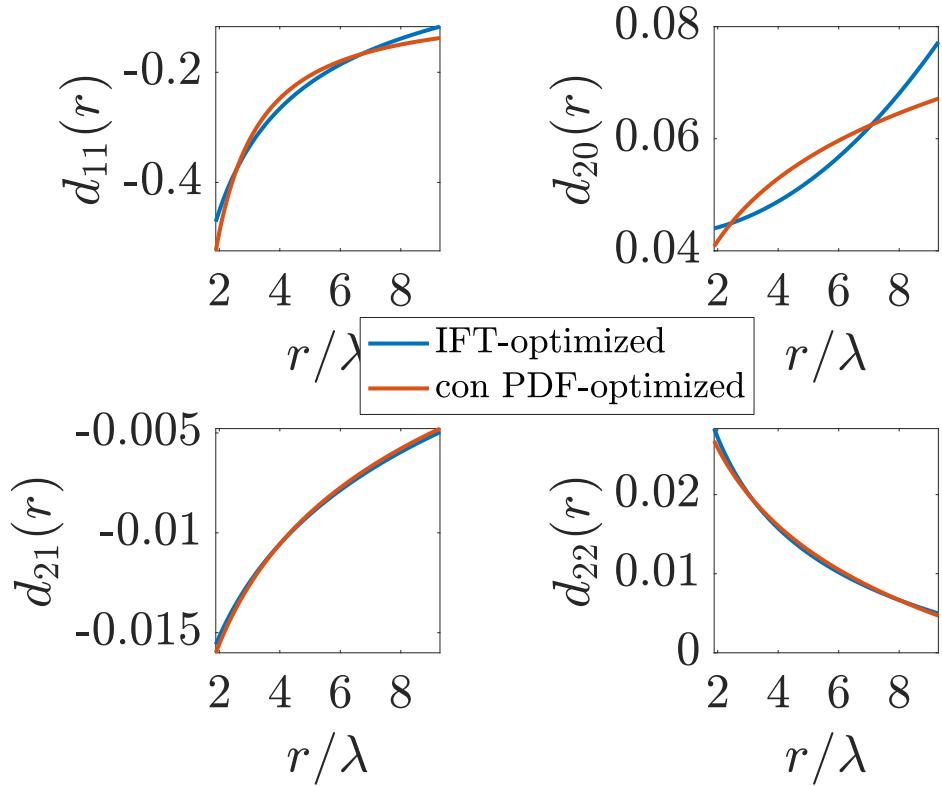


Figure 27: Coefficients $d_{ij}(r)$ of the optimized Kramers-Moyal coefficients using the surface fits with a linear function for $D^{(1)}(u_r, r)$ and a parabolic function for $D^{(2)}(u_r, r)$ (see Eq. 43 - 44) with respect to scale.

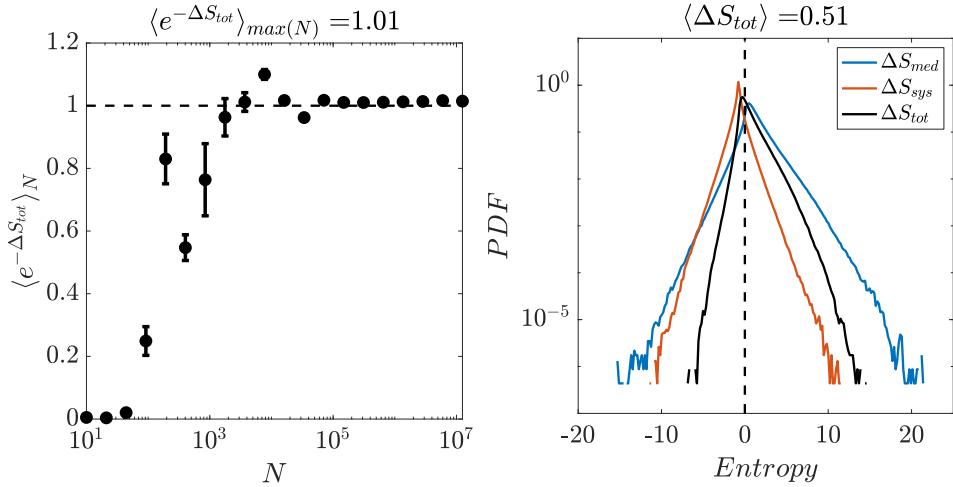


Figure 28: Empirical average $\langle e^{-\Delta S_{tot}} \rangle_N$ of ΔS_{tot} as a function of the number N (sample size) of cascade trajectories $[u(\cdot)]$ with error bars. According to the integral fluctuation theorem (IFT), the empirical average has to converge to the horizontal dashed line. Probability density function of the system S_{sys} , medium S_{med} and total entropy variation S_{tot} .

As it can be seen in Figure 27, this optimization is a fine-tuning of the coefficients, but its impact on the IFT is clearly evident. In this comparison, it must be taken into account that the separation of scales from large to small scales in the cascade trajectory is different and overlapping cascade trajectories ($z=1$) are investigated here.

6. Summary and discussion

6.1. Code Repository

This open source MATLAB package is available as free software, under the GNU General Public License (GPL) version 3, and can be downloaded with all the supplementary material (data, source code and standalone applications (64-bit) for Windows, macOS and Linux) to replicate all the results presented in this paper from the repository on GitHub or Matlab File Exchange Server.

Name: OPEN_FPEIFT

Persistent identifier: GitHub

https://github.com/andre-fuchs-uni-oldenburg/OPEN_FPEIFT

Persistent identifier: Matlab File Exchange Server

https://www.mathworks.com/matlabcentral/fileexchange/80551-open_fpe_ift

Publisher: André Fuchs

Version published: 4.0

Date published: 15/05/21

Operating system: Windows, macOS and Linux

Programming language: MATLAB

6.2. Reuse potential

The development of this user-friendly package greatly enhances practicability and availability of this new method, which allows a comprehensive statistical description in terms of the complexity of turbulent velocity time series (one-dimensional). It can also be used by researchers outside of the field of turbulence for the analysis of data with turbulent like complexity.

Support is available at https://github.com/andre-fuchs-uni-oldenburg/OPEN_FPEIFT, where questions can be posted and generally receive quick responses from the authors.

Acknowledgments

The software resulted from funded research. We acknowledge financial support by VolkswagenStiftung. We acknowledge the following people for helpful discussions and testing pre-version of the package A. Abdulrazek, J. Ehrich, A. Engel, J. Friedrich, A. Girard, G. Gürker, P. G. Lind, D. Nickelsen, N. Reinke, T. Wester, M. Obligado.

References

- Anselmet F, Gagne Y, Hopfinger EJ, Antonia RA (1984). “High-order velocity structure functions in turbulent shear flows.” *Journal of Fluid Mechanics*, **140**, 63–89. doi:10.1017/s0022112084000513.
- Arneodo A, Baudet C, Belin F, Benzi R, Castaing B, Chabaud B, Chavarria R, Ciliberto S, Camussi R, Chillà F, Dubrulle B, Gagne Y, Hebral B, Herweijer J, Marchand M, Maurer J, Muzy JF, Naert A, Noullez A, Peinke J, Roux F, Tabeling P, van de Water W, Willaime H (1996). “Structure functions in turbulence, in various flow configurations, at Reynolds number between 30 and 5000, using extended self-similarity.” *Europhysics Letters (EPL)*, **34**(6), 411–416. doi:10.1209/epl/i1996-00472-2.
- Aronson D, Loefdahl L (1993). “The plane wake of a cylinder: Measurements and inferences on turbulence modeling.” *Physics of Fluids A: Fluid Dynamics*, **5**(6), 1433–1437. doi:10.1063/1.858579.
- Batchelor GK, Townsend AA (1948). “Decay of isotropic turbulence in the initial period.” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, **193**(1035), 539–558.
- Bourgoin M, Baudet C, Kharche S, Mordant N, Vandenbergh T, Sumbekova S, Stelzenmuller N, Aliseda A, Gibert M, Roche PE, Volk R, Barois T, Caballero ML, Chevillard L, Pinton JF, Fiabane L, Delville J, Fournent C, Bouha A, Danaila L, Bodenschatz E, Bewley G, Sinhuber M, Segalini A, Örlü R, Torrano I, Mantik J, Guariglia D, Uruba V, Skala V, Puczyłowski J, Peinke J (2017). “Investigation of the small-scale statistics of turbulence in the Modane S1MA wind tunnel.” *CEAS Aeronautical Journal*, **9**(2), 269–281. doi:10.1007/s13272-017-0254-3. URL <https://doi.org/10.1007/s13272-017-0254-3>.
- Castaing B, Gagne Y, Hopfinger E (1990). “Velocity probability density functions of high Reynolds number turbulence.” *Physica D: Nonlinear Phenomena*, **46**(2), 177–200. doi:10.1016/0167-2789(90)90035-n.
- Feller W (1968). *Probability theory and its applications, vol. 1* New York. John Wiley and Sons, Inc.,.
- Frisch U (1995). *Turbulence*. Cambridge University Press. doi:10.1017/cbo9781139170666.
- Gardiner CW (2009). *Handbook of Stochastic Methods for physics, chemistry, and the natural sciences*. 4th edition edition. Springer, Berlin.
- Hinze JO (1975). *Turbulence*. McGraw-Hill classic textbook reissue series. McGraw-Hill, New York. ISBN 9780070290372. URL <https://books.google.de/books?id=xfRQAAAAMAAJ>.
- Hussein H, George W (1990). “Influence of wire spacing on derivative measurement with parallel hot-wire probes.” In *Forum on Turbulent Flows-1990*, pp. 121–124.
- Kleinhans D, Friedrich R, Nawroth A, Peinke J (2005). “An iterative procedure for the estimation of drift and diffusion coefficients of Langevin processes.” *Physics Letters A*, **346**(1-3), 42–46. doi:10.1016/j.physleta.2005.07.077.

- Lueck S, Renner C, Peinke J, Friedrich R (2006). “The Markov–Einstein coherence length—a new meaning for the Taylor length in turbulence.” *Physics Letters A*, **359**(5), 335–338. ISSN 0375-9601. [doi:10.1016/j.physleta.2006.06.053](https://doi.org/10.1016/j.physleta.2006.06.053).
- Mazellier N, Vassilicos JC (2010). “Turbulence without Richardson–Kolmogorov cascade.” *Physics of Fluids*, **22**(7), 075101. [doi:10.1063/1.3453708](https://doi.org/10.1063/1.3453708).
- Mora DO, Obligado M (2020). “Estimating the integral length scale on turbulent flows from the zero crossings of the longitudinal velocity fluctuation.” [2005.06055](https://arxiv.org/abs/2005.06055).
- Mora DO, Pladellorens EM, Turró PR, Lagauzere M, Obligado M (2019). “Energy cascades in active-grid-generated turbulent flows.” *Physical Review Fluids*, **4**(10), 104601. [doi:10.1103/physrevfluids.4.104601](https://doi.org/10.1103/physrevfluids.4.104601).
- Mordant N, Metz P, Michel O, Pinton JF (2001). “Measurement of Lagrangian Velocity in Fully Developed Turbulence.” [doi:10.1103/PhysRevLett.87.214501](https://doi.org/10.1103/PhysRevLett.87.214501). [physics/0103084](https://arxiv.org/abs/physics/0103084).
- Nawroth AP, Peinke J, Kleinhans D, Friedrich R (2007). “Improved estimation of Fokker–Planck equations through optimization.” *Physical Review E*, **76**(5), 056102. [doi:10.1103/physreve.76.056102](https://doi.org/10.1103/physreve.76.056102).
- Nickelsen D, Engel A (2013). “Probing Small-Scale Intermittency with a Fluctuation Theorem.” *Physical Review Letters*, **110**(21), 214501. [doi:10.1103/physrevlett.110.214501](https://doi.org/10.1103/physrevlett.110.214501).
- Oboukhov AM (1962). “Some specific features of atmospheric turbulence.” *Journal of Fluid Mechanics*, **13**(1), 77–81. [doi:10.1017/s0022112062000506](https://doi.org/10.1017/s0022112062000506).
- O’Neill PL, Nicolaides D, Honnery D, Soria J, et al. (2004). “Autocorrelation functions and the determination of integral length with reference to experimental and numerical data.” In *15th Australasian fluid mechanics conference*, volume 1, pp. 1–4. Univ. of Sydney Sydney, NSW, Australia.
- Peinke J, Tabar M, Waechter M (2019). “The Fokker–Planck Approach to Complex Spatiotemporal Disordered Systems.” *Annual Review of Condensed Matter Physics*, **10**(1), 107–132. [doi:10.1146/annurev-conmatphys-033117-054252](https://doi.org/10.1146/annurev-conmatphys-033117-054252).
- Pope SB (2001). “Turbulent Flows.” *Measurement Science and Technology*, **12**(11), 2020–2021. [doi:10.1088/0957-0233/12/11/705](https://doi.org/10.1088/0957-0233/12/11/705).
- Reinke N, Fuchs A, Nickelsen D, Peinke J (2018). “On universal features of the turbulent cascade in terms of non-equilibrium thermodynamics.” *Journal of Fluid Mechanics*, **848**, 117–153. [doi:10.1017/jfm.2018.360](https://doi.org/10.1017/jfm.2018.360).
- Renner C, Peinke J, Friedrich R (2001). “Experimental indications for Markov properties of small-scale turbulence.” *Journal of Fluid Mechanics*, **433**, 383–409. [doi:10.1017/s0022112001003597](https://doi.org/10.1017/s0022112001003597).
- Risken H (1996). “Fokker-Planck Equation.” In *The Fokker-Planck Equation*, pp. 63–95. Springer Berlin Heidelberg. [doi:10.1007/978-3-642-61544-3_4](https://doi.org/10.1007/978-3-642-61544-3_4).
- Roach P (1987). “The generation of nearly isotropic turbulence by means of grids.” *International Journal of Heat and Fluid Flow*, **8**(2), 82–92. [doi:10.1016/0142-727x\(87\)90001-4](https://doi.org/10.1016/0142-727x(87)90001-4).

- Seifert U (2005). “Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem.” *Physical Review Letters*, **95**(4), 040602. [doi:10.1103/physrevlett.95.040602](https://doi.org/10.1103/physrevlett.95.040602).
- Seifert U (2012). “Stochastic thermodynamics, fluctuation theorems and molecular machines.” *Reports on Progress in Physics*, **75**(12), 126001. [doi:10.1088/0034-4885/75/12/126001](https://doi.org/10.1088/0034-4885/75/12/126001).
- Sekimoto K (2010). *Stochastic Energetics*, volume 799. Springer Berlin Heidelberg. [doi:10.1007/978-3-642-05411-2](https://doi.org/10.1007/978-3-642-05411-2).
- Sreenivasan KR (1995). “On the universality of the Kolmogorov constant.” *Physics of Fluids*, **7**(11), 2778–2784. [doi:10.1063/1.868656](https://doi.org/10.1063/1.868656).
- Sreenivasan KR, Prabhu A, Narasimha R (1983). “Zero-crossings in turbulent signals.” *Journal of Fluid Mechanics*, **137**, 251–272. [doi:10.1017/s0022112083002396](https://doi.org/10.1017/s0022112083002396).
- Taylor GI (1938). “The Spectrum of Turbulence.” *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, **164**(919), 476–490. [doi:10.1098/rspa.1938.0032](https://doi.org/10.1098/rspa.1938.0032).
- Tennekes H, Lumley JL (1972). *A First Course in Turbulence*. The MIT Press. [doi:10.7551/mitpress/3014.001.0001](https://doi.org/10.7551/mitpress/3014.001.0001).
- Tritton DJ (2012). *Physical fluid dynamics*. Springer Science & Business Media.

A. List of Nomenclature/Abbreviations

Latin symbols

C_ϵ	normalized turbulent kinetic energy dissipation rate
$D^{(k)}(u_r, r)$	k -th order Kramers-Moyal coefficients
$E(f)$	energy spectral density in frequency domain
$E(k)$	energy spectral density in wave number domain
E_{kin}	total kinetic energy
F	flatness
F_s	sampling frequency
L	integral length scale
$M^{(k)}(u_r, r, \Delta r)$	k -th order conditional moment
$p(u_r)$	probability density function of u_r
$p(u_{r'} u_r)$	conditioned PDF of velocity increments for a pair of two scales with $r' < r$
$r = -\tau \langle u \rangle$	spatial length scale
Re	Reynolds number
Re_λ	Taylor Reynolds number
S	skewness
$S^k(r)$	k -th order structure function
Ti	turbulence intensity
u	stream-wise velocity
$\langle u \rangle$	mean stream-wise velocity
\tilde{u}	stream-wise velocity fluctuations
u'	root mean square velocity
$u_\tau(t)$	temporal velocity increment at time-scale τ
$u_r(t) = -u_\tau(t)$	spatial velocity increment at length scale r
$[u(\cdot)]$	cascade trajectory

Greek symbols

$\Delta S_{sys} [u(\cdot)]$	system entropy
$\Delta S_{med} [u(\cdot)]$	medium entropy
$\Delta S_{tot} [u(\cdot)]$	total entropy variation
$\langle \epsilon \rangle$	mean energy dissipation rate
ζ_k	scaling exponent of structure functions
η	Kolmogorov length scale
λ	Taylor length scale
μ	intermittency coefficient
ν	kinematic viscosity of the fluid
ρ	fluid density
σ	standard deviation of u

Affiliation:

André Fuchs, Matthias Wächter, Joachim Peinke
 Institute of Physics and ForWind
 University of Oldenburg
 Küpkersweg 70
 26129 Oldenburg, Germany

Swapnil Kharche
 IRIG-DSBT, UMR CEA-Grenoble University
 CEA Grenoble
 17 rue des Martyrs
 38054 Grenoble, France