



## **Report**

Course Project: Statistics of Turbulence and the Onset of Chaos

Name: Matteo Calafà  
Date: May 19, 2022  
Course: Turbulence ME-467  
Instructor: Tobias Schneider

# 1 Part I: Statistical Analysis of Turbulence

## 1.1 Introduction

This first part regards the classical Kolmogorov turbulence theory and a comparison with the results from an experimental setup executed in the Warhaft Wind and Turbulence Tunnel at Cornell University [3]. From the downstream velocity measurements, we aim to achieve significant feedbacks with the theory. In particular, air is driven by a fan that later passes inside the tunnel and measured by 6 anemometers, each one distant  $1m$  from the previous.

In this contest, we aim to compare results with the classical turbulence theory and what is predicted from the turbulence decay theory. However, to make comparisons reasonable, it is firstly needed to analyse how much the theory can fit the experimental setup. In other words, if the necessary conditions and hypotheses are fully accomplished or, at least, sufficiently satisfied.

First of all, it is a preliminary observation to notice that the theory restricts to  $Re \rightarrow \infty$ , something that is obviously not satisfied in experiments. However, since the K41 theory is well known to reflect results from several historical experiments, it is clear that this assumption is only approximate and usually it is only required to have Reynolds number big enough to see the evidence of turbulence, something that is well-satisfied for air at an average speed of  $10m/s$ .

Additional theoretical requirements are homogeneity and isotropy at small scales. This is again a condition that cannot be fully satisfied in experiments (see for instance the collisions with the instruments or the not perfect distribution of the inlet) but can be reasonably assumed as true in our setup where the inlet from the fan is constant and precise (and therefore, the forced turbulence is stationary), the direction of the mean flow is approximately horizontal (thanks to the grid between the inlet and the tunnel) and the tunnel is large enough. Moreover, turbulence is fully-developed thanks to the reasonable distance of the anemometers from the inlet.

Other important hypotheses for the K41 theory are the *small scale scaling* and the *finite dissipation rate* for  $\nu \rightarrow 0$  that are usually considered true in experiments, also because they were originally inspired from empirical results.

Moreover, the driving forces (that in this case are represented by the fan action) are evidently applied only on large scales and this will be seen later in the energy spectrum analysis (section 1.2.3).

As anticipated, the second goal of this part is to analyse in the experiment the characteristics of the turbulence decay under the light of the theory for unforced turbulence. This is the main reason of multiple measurements at different positions that, therefore, should prove a decay in the turbulence quantities because of the absence of applied forces except at the inlet. However, the theory requires some further assumptions since K41 hypotheses are not enough. First of all, it is assumed that the velocity associated to large scales has an *infrared asymptotic self-similarity*, i.e.  $v_l \sim Cl^h$  for  $l \rightarrow \infty$  and a certain constant parameter  $h$  (that will be estimated in section 1.3). Then, the *Principle of Permanence of Large Eddies* must hold. On the other hand, these hypotheses are usually considered as sufficiently valid during experiments: also these laws came from experiments and inspired the theory.

We note the fact that these results lead to the temporal law of the main turbulence quantities while the classical theory deals with stationary turbulence. However, this is justified by the fact that the power law decay is still very "slow" with respect to the internal turbulence time scales.

To conclude, we will assume for a while the *Taylor froze hypothesis* even if, in section 1.2.1, it will be shown not to be properly fulfilled.

## 1.2 Data Analysis

### 1.2.1 Velocity Signal in the Spatial Domain

The measurements correspond to streamwise velocities detected at a frequency  $f = 20kHz$ . This implies that each measurement follows the previous one after

$$\Delta t = 1/f \quad (1)$$

We can exploit the Taylor frozen hypothesis to detect how much two measurements are distant in space knowing the time lag. Indeed the distance in space between two following measurements is

$$\Delta x \simeq - \langle u \rangle \Delta t \stackrel{(1)}{=} - \langle u \rangle / f \quad (2)$$

where  $\langle u \rangle$  is the mean velocity at each anemometer. Therefore, we can now compute the "virtual" distance in space between two measurements knowing how many measurements there are between them: the number of measurements is exactly the ratio of the space distance over  $\Delta x$ .

With this preface, we can execute a first analysis showing in Fig. 1 the measurements at each anemometer against the virtual space distance.

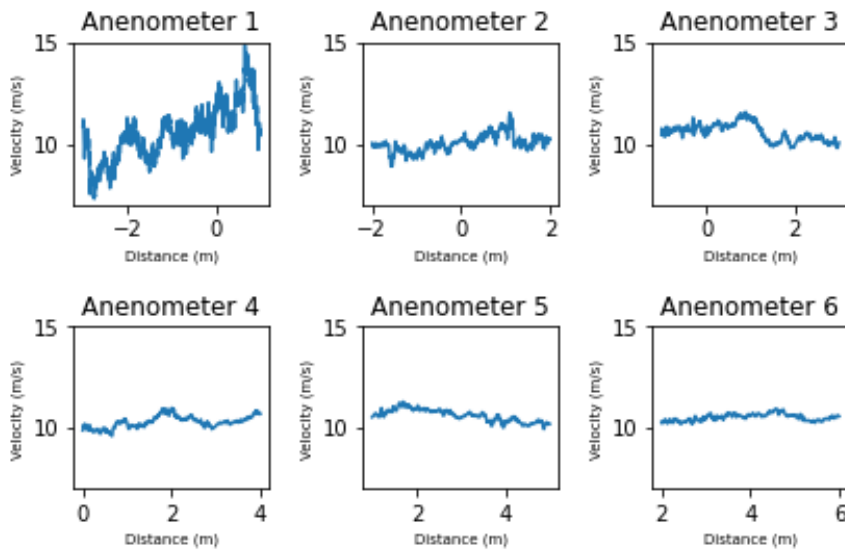


Figure 1: PLOT A - Streamwise velocity measurements for each anemometer. The plots are shown with respect to space thanks to the Taylor frozen hypothesis.

The  $x$  values represent the spatial position of the anemometer and the previous  $4m$  as in our choice. First of all, we notice that the mean velocity is almost constant for each anemometer ( $\approx 10m/s$ ). This is expected because of conservation of mass and constant inlet velocity. I.e., the same quantity of air that goes inside the tunnel should go outside at the same moment if the density is considered constant or almost constant. Because the velocity at the inlet is constant, it implies that the mean velocity should be approximately constant everywhere.

On the other hand, oscillations are more evident in the first anemometer and they later vanish. This is expected and due to two main reasons: the first one is due to the difference between the entrance/unstable region of the tunnel and the fully-developed region where the flow is more asymptotically stable.

The second reason is instead related to the decay of unforced turbulence. Clearly, the only applied force is at the inlet of the tunnel where turbulence is therefore more intense.

In Table 1 we instead show the results of mean velocities and turbulence intensities. To obtain them, it is enough to compute respectively the mean and the standard deviation of the velocities measurements at each anemometer, indeed:

$$U := \langle u \rangle \quad I := \frac{\sqrt{\langle (u - U)^2 \rangle}}{U}$$

These results clearly confirm the previous discussion: the mean velocity is almost constant in each anemometer and turbulence intensities instead decrease. Clearly, turbulence intensities are related to the turbulence decay and the decrease of oscillations in Fig. 1.

Table 1: Mean velocity and turbulence intensity

Param.	Dim.	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$d$	$m$	1.0	2.0	3.0	4.0	5.0	6.0
$U$	$m/s$	10.52223399	10.52201583	10.52117473	10.52232686	10.52234804	10.52185721
$I$	adim.	0.12180673	0.05480351	0.03954964	0.0320026	0.02713776	0.02407247

Certainly, the Taylor frozen hypothesis is just an approximation of the real phenomenon since turbulence has a certain effect on the flow variations. In general, the approximation is valid if the turbulence intensity is very small but this is not our case (especially at the beginning where this values is more than 0.12). Moreover, the turbulence intensity changes a lot in space and this gives incoherent results: see for instance in Fig. 1 the velocity measurements in the second anemometer at the position  $x = -2m$ . This trend is completely different from velocity measured by the first anemometer and associated to the same spatial position. This is because we miss the measurements in the intermediate positions where turbulence intensities have inevitably different values.

However, the error in the estimation of intermediate velocity profiles can be easily bounded. The mean velocity is approximately the same in all the positions while the turbulence intensity can be achieved through an extrapolation. In particular, the turbulence intensity follows a trend similar to the ones presented in section 1.3 (it is indeed equivalent to the square root of the kinetic energy except for a constant coefficient). Therefore, from the results in the table and the assumption of the trend, a regression analysis can be performed to estimate these values in all the intermediate positions.

### 1.2.2 Correlation Length of the Velocity Signal

The autocorrelation function and the integral scale are defined as:

$$C(l) := \frac{\langle u(x+l)u(x) \rangle}{\langle u^2(x) \rangle} \quad L_{int} := \int_0^\infty C(l) dl$$

Moreover, we can approximate  $L_{int}$  with the correlation length  $L_C$  that is the value for  $l$  such that  $C(l)$  drops down to  $1/e$ .

In table 2 we show the measures of the correlation length and the integral scale. To compute the autocorrelation, we exploited the optimized `correlate` function from `Scipy`. To compute the integral scale, we used the same function and after a rectangular quadrature formula for the integral. In Fig. 2, we instead graphically show the trend of the correlation length.

Table 2: Correlation length and integral length

Param.	Dim.	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$L_C$	m	0.36669985	0.63447755	0.77330634	0.90386788	1.00909318	1.08532957
$L_{int}$	m	0.35895344	0.62704922	0.76843644	0.88769361	1.00077964	1.07598959

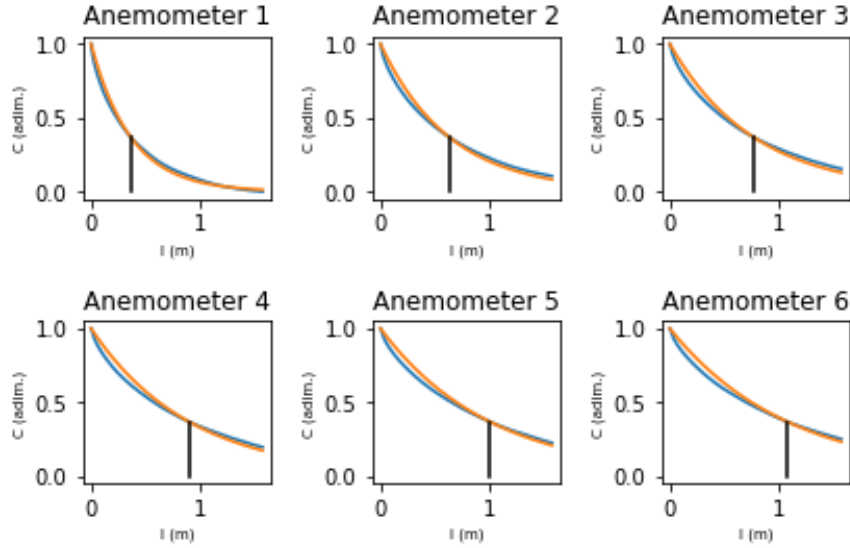


Figure 2: PLOT B - Correlation lengths with respect to the space distance  $l$  (blue curves). Orange lines represent instead the functions  $e^{-\frac{l}{L_C}}$ . Finally, the black lines are the associated  $L_C$  values.

We notice that the two scale lengths are very similar which confirms the fact that the first theoretically approximates the second. Let us clarify this point. Giving a glance at the plots, we recognize soon an exponential decay of the correlation function. Since  $C(0) = 1$  always (by definition and also from the plots) we conclude that it is approximately in the form:  $C(l) = e^{-al}$  for a certain steep coefficient  $a$ . It implies that:

$$L_{int} := \int_0^\infty C(l) dl = \int_0^\infty e^{-al} dl = \frac{1}{a}$$

We now notice that  $1/a$  is also the value such that  $C(1/a) = e^{-1}$ , from here it follows the equivalence of the two definitions of length scale. In addition, these exponential laws are added to Fig. 2 to show the effective overlap of the autocorrelation function with its analytical approximation.

Let us now discuss the trend of the lengths. Intuitively, we can say that the decaying of turbulence implies a stronger correlation between distant points. In other words, when the turbulence is strong, the noisy dynamics implies the velocity vectors to be almost independent despite they might be close. Another reason to convince us of this correct trend is the relation between these length scales and the integral length scale presented in section 1.2.3. We will show later the trend of the latter but we now anticipate that it will increase following the theoretical predictions. Therefore,  $L_C$  and  $L_{int}$  correctly increase as  $l_0$  does.

### 1.2.3 Energy Spectrum of the Flow

We define the spectral energy density as  $E(k) := \tilde{E}(k) + \tilde{E}(-k)$  where:

$$\tilde{E}(k) := \frac{1}{2} \left| \frac{1}{\sqrt{2\pi L}} \int_0^L u(x) e^{-ikx} dx \right|^2$$

where  $k \in \mathbb{R}$  represents the frequencies and  $L$  a maximal measurable distance. In Fig. 3 we report the measured energy spectrum for each anemometer.

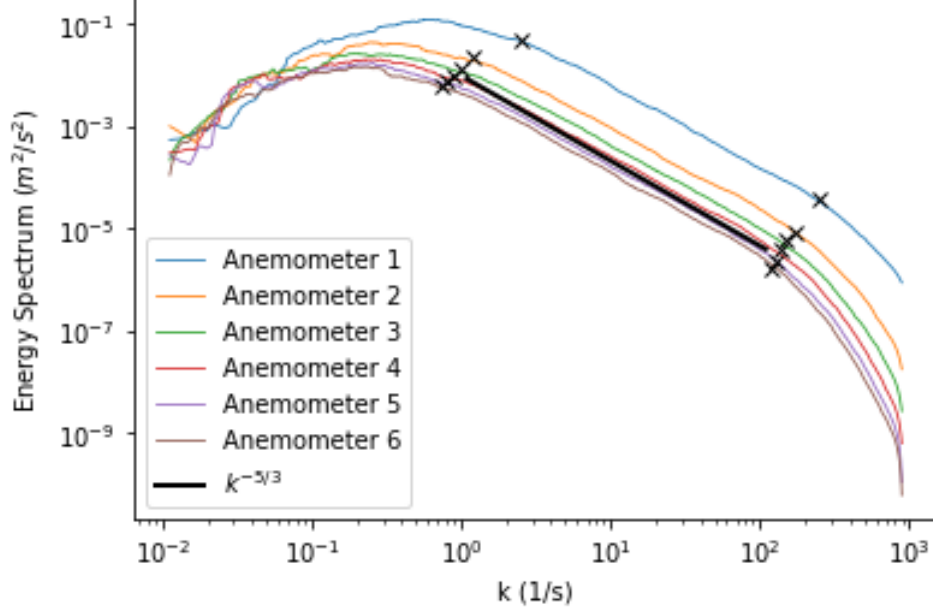


Figure 3: PLOT C - Log-log plot of the energy spectrum of the velocity measurements for each anemometer. Also a black line that indicates the  $-5/3$  slope has been inserted.

The energy spectrum has been calculated exploiting the optimized `fft` algorithm from `Scipy` for the positive frequencies and the inverse `ifft` for the negative frequencies. After that, it is just needed to sum the two contributions paying attention to consider the correct multiplicative coefficient (i.e.  $(\Delta x)^2$ ). As a confirmation, we report here the relative error between

$$\frac{1}{2} \langle u^2 \rangle \text{ and } \int_0^\infty E(k) dk$$

for the first anemometer that turns out to be:  $2.47 \cdot 10^{-7}$ . These quantities should be theoretically equal because of the Parseval theorem and the relative error is very small because the energy spectrum has been calculated very precisely using the full dataset (this is indeed possible in short times, i.e. less than one minute, thanks to the optimized functions cited above).

To conclude the implementation aspects, the original noisy spectrum has been later filtered with a *Savitzky-Golay filter* ([1]). More precisely, a standard application of the filter would overshoot in the left part of the plot because, in a log-log plot with originally equidistant  $k$  values, the density of the number of points increases towards the right part of the plot. Therefore, a regression-based filter as Savitzky-Golay would be unbalanced in the two different regions. Our proposed solution is to interpolate the original spectrum at points that follow an exponential grow so that, in a log-log plot, they appear equi-distant and the filter can be then well-balanced.

We can finally comment the plots in Fig. 3 and see how they respect the predictions of the Kolmogorov theory. We can indeed clearly distinguish the three regions (from left to right: the large scales with driving forces, the inertial range and the dissipation region with small scales). In particular, the inertial range clearly respects the  $5/3$  law since the straight line is perfectly parallel to the trend of all the anemometers. Moreover, the large scales lines seem to respect the order predicted in the turbulence

decay theory (indeed, this will be shown later in section 1.3).

From the energy spectrum, we can detect the starting and ending points of the inertial range that correspond to the integral and Kolmogorov frequencies. These points have been signed in the plot with black crosses. Finally, from these frequencies, we can get the integral and Kolmogorov length scales that are reported in table 3.

Table 3: Integral and Kolmogorov length scales

Param.	Dim.	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$L_{int,E}$	m	2.51327412	5.23598776	6.28318531	6.98131701	7.85398163	8.37758041
$\eta_E$	m	0.02513274	0.03695991	0.0418879	0.0448799	0.04833219	0.05235988

Reminding what has been discussed in section 1.2.2 about the increase of the correlation length, also here the length scales increase with the distance  $d$  because of the same reason as before.

As a last result, we want to check effectively the relationship between the two integral scales. This is illustrated in Fig. 4. Here, we show that the two scales are in fact approximately proportional. This fact should not surprise since turbulence length scales are known to be successfully represented by both autocorrelation and spectral length scales and therefore they must scale together (see for instance [2]). To conclude, we observe that every time  $L_{int,E}$  and  $\eta_E$  will be used to achieve new results, we will always incur in some inaccuracies due to the graphical detection of these values from the plot that, by its nature, is not precise.

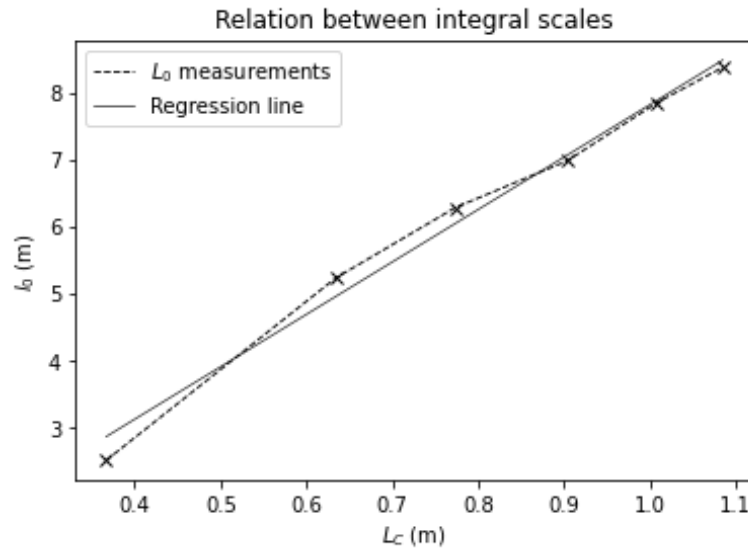


Figure 4: EXTRA - Relation between the integral scale length and integral length of section 1.2.2. Despite some inaccuracies due to the rough  $L_0$  detection on the plot in Fig. 3, the two values are clearly proportional.

### 1.2.4 The Dissipation Rate and Different Reynolds Numbers

We start by defining the energy dissipation rate and the two Reynolds numbers as:

$$\epsilon := \frac{1}{2} \frac{\sqrt{\langle u^2 \rangle^3}}{L_C}, \quad Re_\lambda := \frac{\sqrt{\langle u^2 \rangle} \lambda}{\nu}, \quad Re := \frac{u_0 l_0}{\nu}$$

where  $\lambda$  is the Taylor length scale,  $\nu$  is the viscosity,  $u_0$  is the velocity associated to the large eddies (i.e. the eddies of length  $l_0 = L_{int,E}$ ). In table 4 we report these quantities that have been computed starting from the results of the previous sections.

Table 4: Dissipation rate, Taylor Reynolds and Reynolds numbers

Param.	Dim.	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$\epsilon$	$m^2/s^3$	2.87076074	0.15110392	0.04658421	0.02112303	0.01153716	0.00748596
$Re_\lambda$	Adim.	969.527172	855.4148643	802.221265	780.2180855	759.1436670	741.4848821
$Re$	Adim.	310190.581	290554.460	253269.393	227787.447	217626.8078	204905.321

We can again give both a physical interpretation and a mathematical explanation to the decreasing trend of  $\epsilon$ . First of all, as  $\epsilon$  represents the rate of turbulent kinetic energy dissipated in thermal energy, it is clear that the decaying turbulence implies a lower rate of energy shift and then a lower dissipation rate for higher  $d$ . The second motivation is due to the fact that, by definition of  $\epsilon$  and the previous relation between length scales, we can state that  $\epsilon \sim E^{3/2}/l_0$ . Using the rates from the turbulence decay theory (section 1.3, equations 3, 6), we can soon achieve an order of  $2h/(1-h)$  for  $\epsilon$  with respect to  $d$ , therefore a power decreasing law.

Comparing instead the Reynolds numbers, the decreasing trend is again expected from the theory of turbulence decay and, moreover,  $Re_\lambda$  scales as the square root of  $Re$  (except for a coefficient  $\approx 2$ ) as predicted from K41.

To conclude, these trends are all coherent with the turbulence decay results and therefore also with the trend of  $I$  from section 1.2.1.

## 1.3 Turbulence Decay

In table 5 we report the computations of the kinetic energies per unit mass. This is defined as  $\mathcal{E} = 3/2 \langle u^2 \rangle$  and the 3 factor comes from the number of dimensions. Indeed:

$$\mathcal{E} = \frac{1}{2} \langle \vec{u} \cdot \vec{u} \rangle = \frac{1}{2} \langle u_x^2 + u_y^2 + u_z^2 \rangle = \frac{3}{2} \langle u^2 \rangle$$

Where the last equality comes from the isotropic assumption. Indeed, what we measure is the streamwise velocity  $u_x$  that is assumed to be equal also in the other directions.

Table 5: Turbulence kinetic energy per unit mass

Param.	Dim.	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$
$\mathcal{E}$	$m^2/s^2$	2.46405129	0.49877643	0.25971965	0.1700926	0.12231075	0.09623149

We finally report the laws predicted from the theory of turbulence decay:

$$l_0 \propto (d - d_0)^{1/(1-h)} \quad (3)$$



$$u_0 \propto (d - d_0)^{h/(1-h)} \quad (4)$$

$$Re \propto (d - d_0)^{(1+h)/(1-h)} \quad (5)$$

$$E \propto (d - d_0)^{2h/(1-h)} \quad (6)$$

Where  $d_0$  and  $h$  are constants. We will equivalently analyse the scaling of  $\mathcal{E}$  instead of  $E$  (what changes is the mass/density term).

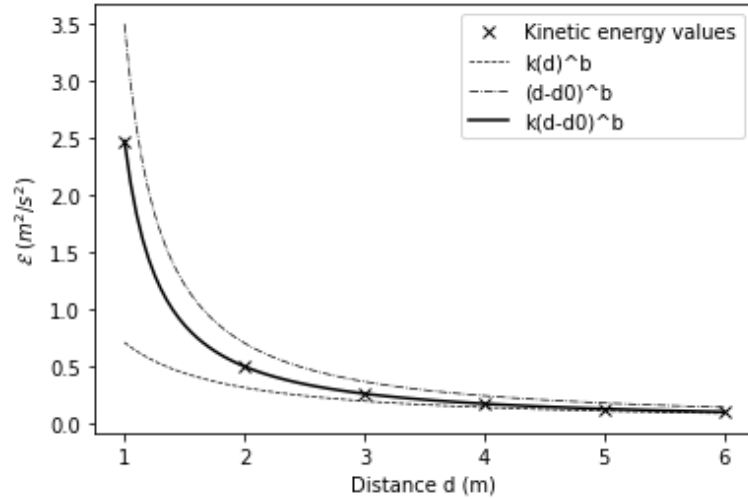


Figure 5: EXTRA - Fitting of the turbulence kinetic energy per unit mass. The other dashed lines are just to show the importance of both  $k$  and  $d_0$  parameters.

**First method:** The first objective is to find  $d_0$  and  $h$  from the measurements of  $\mathcal{E}$  through a fitting method. Certainly linear regression cannot be used because of the non linearity of the exponent, however we can notice that:

$$\exists K, d_0, b \text{ such that } \mathcal{E}(d) \approx K(d - d_0)^b \Leftrightarrow \ln \mathcal{E} \approx \ln K + b \ln (d - d_0) \quad (7)$$

Therefore, applying the logarithm to the data, we only have one non-linear coefficient that is  $d_0$ . The proposed algorithm is then the following:

1. We iterate over a set of hypothetical values for  $d_0$  (numerous enough).
2. For each of these, we compute the optimal  $K$  and  $b$  through linear regression.
3. Given the choice of  $d_0, K, b$ , we compute the resulting MSE error.
4. At the end of the loop, we take the  $d_0$  that gave the lowest MSE error.
5. We can finally get the associated  $K$  and  $b$  again from linear regression.

This is a simple and quick algorithm since iterations are performed over only one parameter. The resulting values are then:  $d_0 \simeq 0.65$ ,  $h \simeq -1.46$ . In addition we show in Fig. 5 the successful fitting method. From now on, for simplicity, we will consider the optimal exponent as  $h = -1.5 = -3/2$ .

**Second method:** Taking again equation 7, we hence have proven that for correct values of  $d_0, b, K$  the log-log plot of  $\mathcal{E}$  against  $d$  is approximately a straight line. Fig. 6 shows a log-log plot of  $\mathcal{E}$  against  $(d - d_0)$  for different values of  $d_0$ . For the same reason as before, only the correct  $d_0$  generates a straight line. Indeed, only the pink and grey lines are close to be straight, therefore  $d_0$  is between 0.6 and 0.7 as confirmed by the fitting method. Taking it in the middle ( $= 0.65$ ), we can estimate  $q$  looking at the slope of the lines. The black line shows the trend of  $\mathcal{E}$  predicted by choosing the previous value for  $d_0$  and the correct  $q$  to have parallel lines. Looking at the numerical values of  $d_0$  and  $q$ , we conclude that the graphical and fitting methods are coherent. Indeed, for both, the correct shifting parameter is approximately  $d_0 = 0.65$  and the exponent  $h = -3/2 \Leftrightarrow q = -6/5$ .

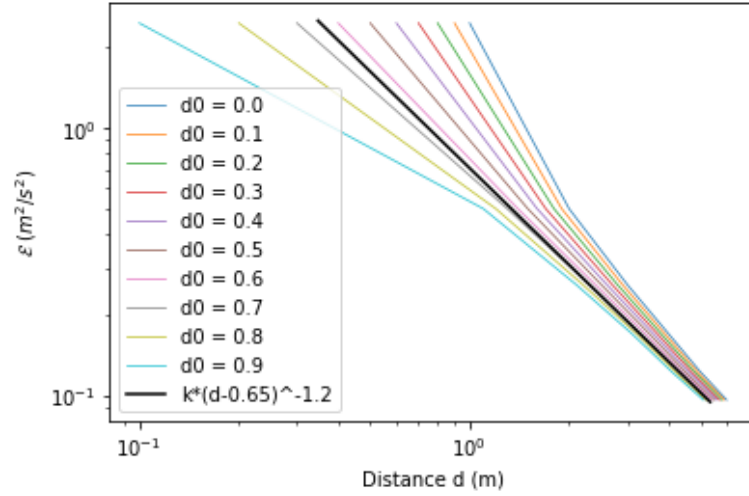


Figure 6: PLOT D - Log-log plot of the measured  $\mathcal{E}$  with respect to  $(d - d_0)$  for different  $d_0$ . As explained in equation 7, only the correct  $d_0$  generates a straight line. Therefore, the correct  $d_0$  is between 0.6 and 0.7. Hence, we have chosen  $d_0 = 0.65$ . After that, the exponent  $q$  has been chosen such that the line has the same slope. Finally, the graphical method prediction is shown through the black line.

**Third method:** Another confirmation comes from the scaling property:

$$\mathcal{E} \propto l_0^{2h} \quad (8)$$

First of all, we inform that we can replace  $l_0$  with  $L_C$  thanks to the reasons given at the end of section 1.2.3. Therefore, we aim now to study this relation and a comparison between the experimental values of  $\mathcal{E}$  against  $L_C$  is shown in Fig. 7.

The power law of  $\mathcal{E}$  is evident from the straightness of the line. Moreover, these measures coincide almost exactly with the ones from the *Saffman's decay* under which  $h = -3/2$ . This is expected since we have already shown from the previous methods that  $h$  is approximately  $-1.5$ .

Seeing the excellent overlapping of the experimental values with the Saffman's decay predictions, it should not surprise that the  $d$  to equation 8 with fixed  $h = -1.5$  returns again a parameter  $d_0$  that is approximately 0.65.

**Comparison of the methods:** To conclude this part about the estimation of  $d_0$  and  $h$ , we can observe that the three methods are indeed equivalent since they return the same values (except for

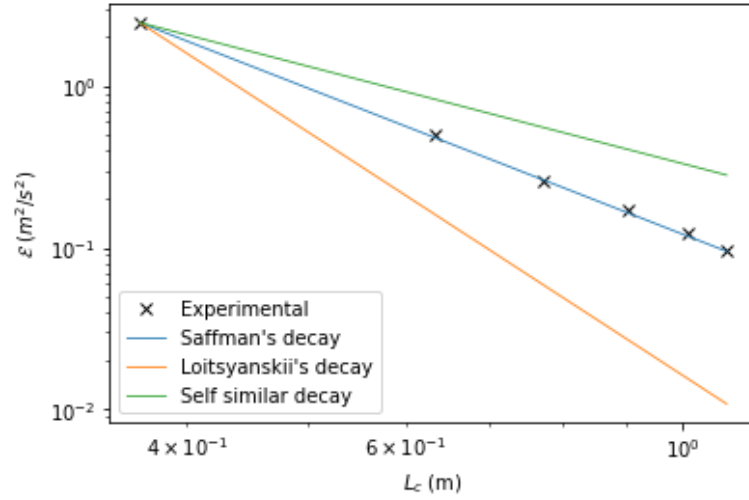


Figure 7: PLOT E - Log-log plot of the measured  $\mathcal{E}$  with respect to  $L_C$ .

approximations). Advantages and disadvantages of each one are listed below:

- The fitting method is very precise. However, it can be a bit slower if compared to the others because of the several calculations (even if it is in general quite fast). In addition, it is harder to interpret results only from the numerical values.
- The graphical method is instead good to interpret results. However, it is not precise at all.
- The relation with  $L_0$  is precise and has theoretical supports. However, it restricts to only 3 possible theories and it is computationally slow as the first method (since it still needs to iterate over  $d_0$ ).

**Turbulence decay and energy spectrum:** In Fig. 8 we show again the energy spectrum. This time we want to study it with respect to the prediction of its trend for large eddies. Despite some noise (we have already discussed the issue of a fitting/filtering method in a logarithmic scale, see section 1.2.3), the lines seem parallel to the predicted line. Therefore, our theoretical results and the choice of  $h = -3/2$  seem to confirm again the experiments.

**Discussion about the virtual origin  $d_0$ :** The virtual origin  $d_0$  is certainly not only a fitting parameter but represents, in a certain way, the starting position for the decaying turbulence predicted trends. In other words, the laws in equations 3, 4, 5, 6 do not have any validity for distances  $d < d_0$ , neither mathematically (exponentials of negative bases) nor physically. Until the distance  $d_0$ , the turbulence is expected not to freely decay but, instead, to be still forced. This is the reason why our  $d_0$  value is between  $0m$  and  $1m$ , i.e. in the inlet region where applied forces have still a dominant role in the turbulence process.

We expect  $d_0$  to be approximately the same for all the above relations (eq. 3, 4, 5, 6) since, the condition of forced/unforced turbulence should hold independently for all the above quantities. Moreover, mathematically, the cited relations are all achieved from one of those and therefore the  $d_0$  is the same for each one. However, these are scaling relations and there are inaccuracies in the values achieved for  $l_0$  and  $Re$ , so we do not expect to achieve experimentally the exact same value of  $d_0$  for all these quantities. We have indeed performed some fitting methods on equations 3 and 5 fixing the exponent  $h = -3/2$  and we achieved respectively  $d_0 = 0.75m$  and  $d_0 = 0.25m$ . Despite these values are not

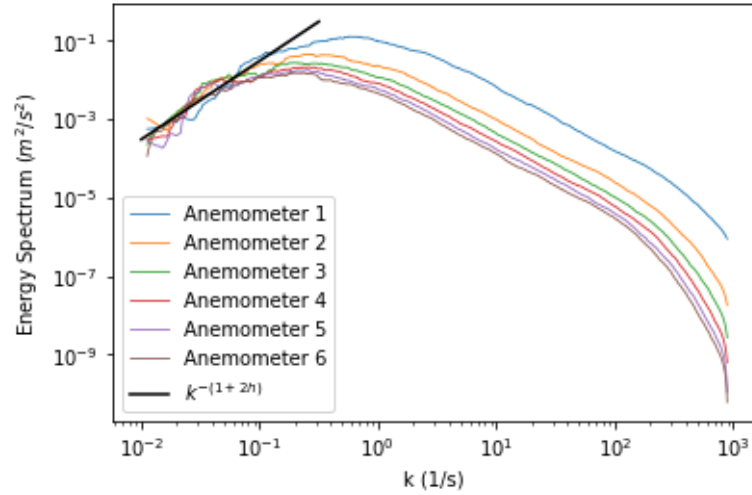
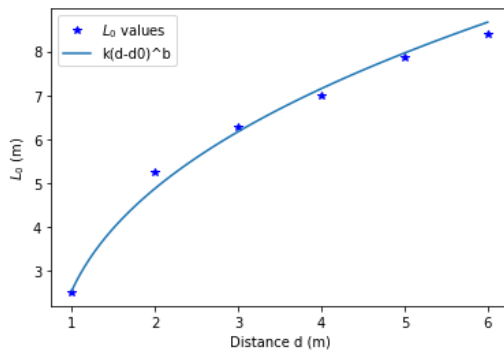
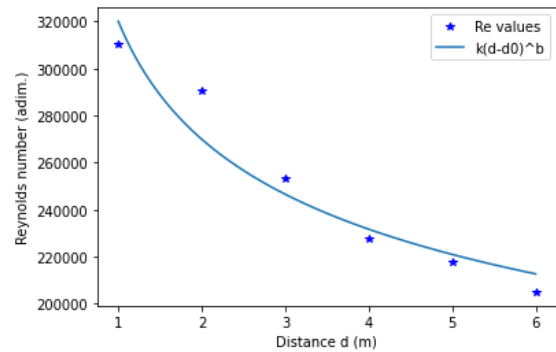


Figure 8: PLOT F - Energy spectrum (already shown in Fig. 3) with the theoretical trend for the large scales. Despite some noise, the spectrum seems to scale as the predicted trend.

similar with the previous  $d_0 = 0.65m$ , we still have values in the same interval  $0m - 1m$  to confirm our theoretical predictions. In addition, we show in Fig. 9 that the choice for the  $h$  value is suitable also for the  $l_0$  and  $Re$  values.



(a) Fitting for  $l_0$  values



(b) Fitting for  $Re$  values

Figure 9: EXTRA - Fitting method applied to equations 3 and 5 fixing  $h = -3/2$ . We clearly see that this value for  $h$  is well suitable also for these quantities.

### 1.3.1 Velocity Increments

In Fig. 10 we show the behaviour of the longitudinal velocity increment  $\delta v_{||}(x, l) := u(x + l) - u(x)$  for the distances  $l = \{1mm, 1cm, 10cm, 10m\}$ .

First of all, considering the turbulent length scales in table 3, we can state that:

- $l = 1mm$  is in the dissipative region ( $l < \eta_E$ ).
- $l = 0.01m$  is between the dissipative and inertial region ( $l \approx \eta_E$ ).

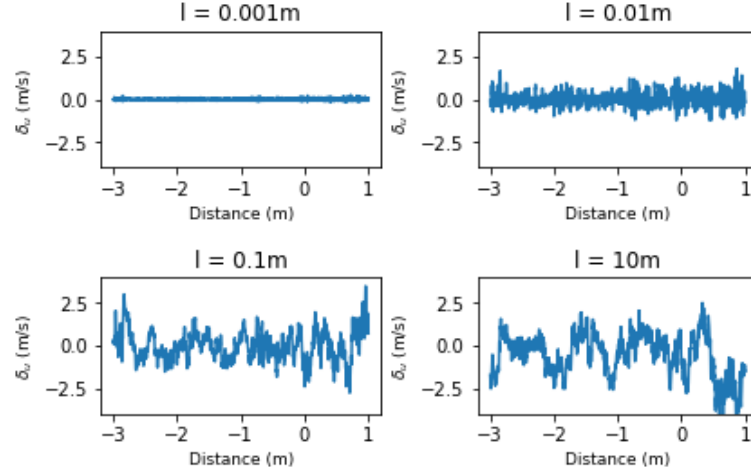


Figure 10: PLOT G - Longitudinal velocity increment for different distances  $l$  and measured at the first anemometer.

- $l = 0.1m$  is in the inertial range ( $\eta_E < l < L_{int,E}$ ).
- $l = 10m$  is in the large scales region ( $l > L_{int,E}$ ) even if the two lengths are quite comparable.

Focusing now on the results from Fig. 10, the behaviour is expected since a small  $l$  implies strong similarity between the two measures (that are indeed near in space) where instead a big  $l$  implies no connection and therefore a completely random difference. This kind of reasoning is very similar to the one made for the autocorrelation functions in section 1.2.2. In both cases, the increase of distance implies a decrease of correlation.

It is important to notice that this increment of randomness has a limit since two measures that are very distant in space are completely uncorrelated and so they show the same pattern whatever is the value of  $l \gg 0$ . As a demonstration, the difference between the case with  $l = 0.1m$  and  $l = 10m$  is not so relevant if compared to the upper plots.

To further investigate the properties of  $\delta u_{||}$  we might wonder whether this has a normal distribution. To do that, we analyse the trend of the flatness that is defined as:

$$f(l) := \frac{\langle \delta u_{||}^4(x, l) \rangle}{\langle \delta u_{||}^2(x, l) \rangle^2}$$

For a centred normal distribution, the fourth moment is known to be equal to  $3\sigma^4$  while the second moment is the variance. Therefore the flatness is  $3\sigma^4/(\sigma^2)^2 = 3$ . The flatness of  $\delta u_{||}$  is instead shown in Fig. 11.

It is reasonable that the flatness tends to the one from the normal distribution because, as we stated above, big distances imply differences of independent measures and, hence, normally distributed.

The reason why the flatness is higher for small distances must be associated to the not normality of  $\delta u_{||}(x, l)$  for small  $l$ . In particular, higher flatnesses are characteristic of the so-called *leptokurtic* distributions, i.e. distributions that are more squeezed at zero and with fatter tails. An equivalent interpretation is the presence of more *outliers*. Therefore, we can reasonably think of  $\delta u_{||}$  for small  $l$  as more concentrated in zero (see indeed Fig. 10) where, however, there are many outliers due to noisy turbulent effects. Physically, the difference between two near measures is usually small because

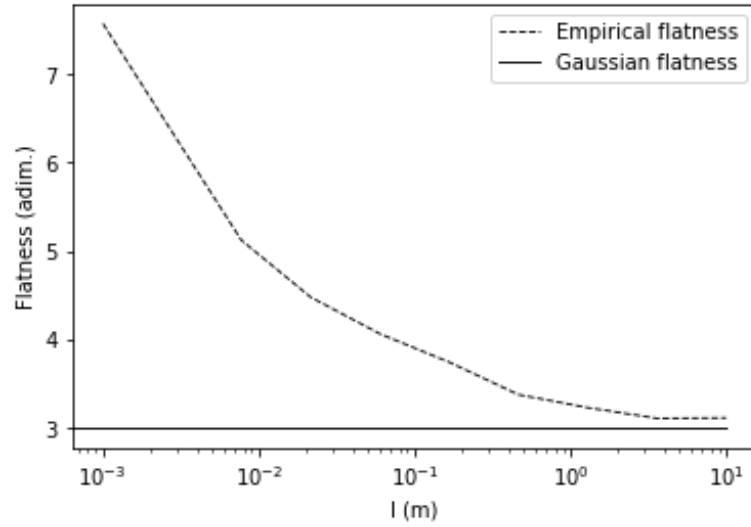


Figure 11: PLOT H - Longitudinal velocity increment flatness for different  $l$ . It is easy to see the convergence towards the normal flatness 3.

of the high correlation, however, some noisy sudden effects can sometimes cause completely different measures.

### 1.3.2 Structure Functions and Energy Dissipation

In Fig. 12 and 13 we show the trend of  $S_2(l)$  and  $S_3(l)$ .

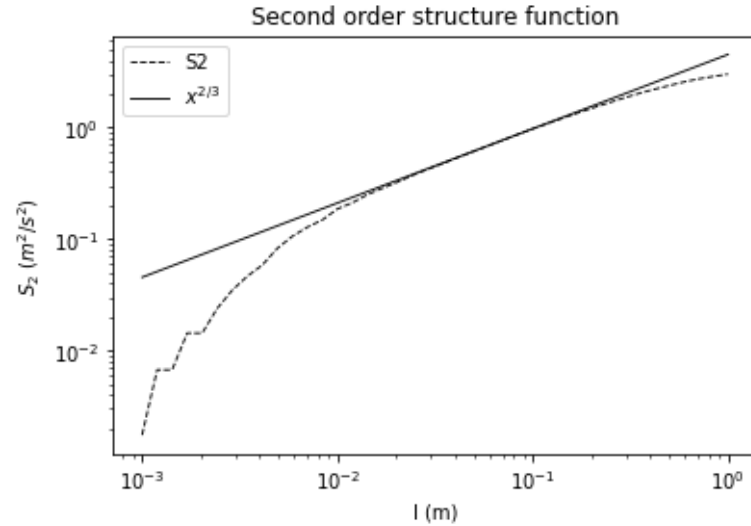


Figure 12: PLOT I - Second order structure function against distance  $l$ . In addition, the theoretical trend from K41.

First of all, we note that  $S_2$  successfully satisfies the  $2/3$  law in a certain range. We know from the K41 that the  $5/3$  law for the energy spectrum is related to the  $2/3$  law of  $S_2$  (more precisely, from the latter we can obtain the former with the Wiener formula). We would like to investigate now whether

the  $2/3$  law has the same range of validity of the  $5/3$  law in Fig. 3. Since these plots are referred to only the first anemometer, we already computed the Kolmogorov and integral length scale for the first anemometer as  $\eta_E \approx 0.025m$  and  $L_{int,E} \approx 2.51m$  (table 3). The clear  $2/3$  slope is instead observed only in a smaller range: the lower bound is bigger than  $0.01m$  and the upper bound is smaller than  $1m$ . On the other hand, even if this range is shorter, it is centred (not shifted) as the one from the energy spectrum and has more or less the same order of magnitude.

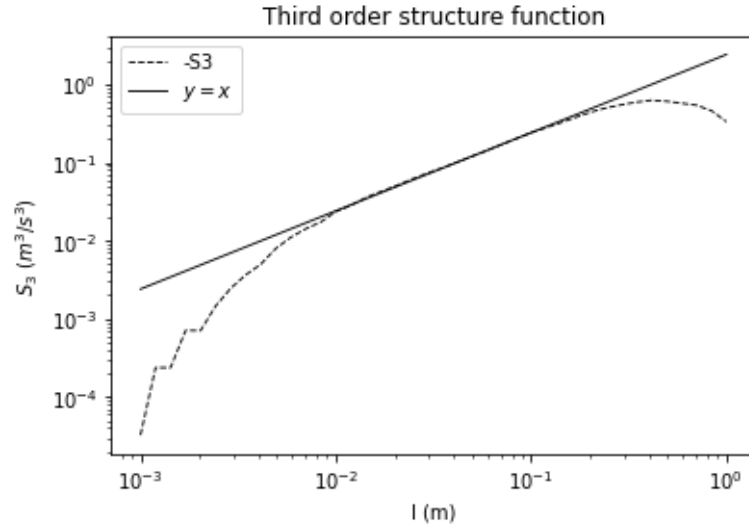


Figure 13: PLOT J - Third order structure function against distance  $l$ . In addition, the theoretical trend from K41.

The same could be noticed in Fig. 13 for  $S_3$  which respects the  $4/5$  law for a certain range that, also in this case, is similar but shorter than the one from the energy spectrum. To conclude, the trends of  $S_2$  and  $S_3$  support the K41 except these intervals do not completely correspond to the one from  $E(k)$ . A last notice is about the fact that from  $S_2$  and  $S_3$  we could potentially achieve a value for the energy dissipation rate  $\epsilon$ . Indeed the  $2/3$  and  $4/5$  laws state that:

$$S_2(l) = C_2 \epsilon^{2/3} l^{2/3} \quad \text{and} \quad S_3(l) = C_3 \epsilon l \quad (9)$$

where  $C_3 = -4/5$  and  $C_2$  can be assumed to be 2.2. Taking a value for  $l$  that is valid for these laws (e.g.  $l = 0.07m$  is in the middle of the validity ranges in Fig. 12 and 13) we can then use the inverses of the equations 9 to get  $\epsilon$ . These values are shown in table 6.

Table 6: Dissipation rate value from  $S_2$  and  $S_3$  relations

Method	$\epsilon$ value ( $m^2/s^3$ )
Standard definition (from table 4)	2.87
From $S_2$ ( $2/3$ law)	2.97
From $S_3$ ( $4/5$ law)	3.01

These satisfying results confirm again the validity of the theory in our experiment.

#### 1.4 Discussion (Limit: 1 page)

## 2 Part II: Nonlinear Dynamics and the Emergence of Chaos

### 2.1 Introduction

This second part is about the analysis of the dynamics of the generalized Baker's Maps defined as:

$$\begin{aligned} x_{n+1} &= \begin{cases} \alpha_1 x_n & \text{if } y_n < \beta \\ (1 - \alpha_2) + \alpha_2 x_n & \text{if } y_n \geq \beta \end{cases} \\ y_{n+1} &= \begin{cases} y_n / \beta & \text{if } y_n < \beta \\ (y_n - \beta) / (1 - \beta) & \text{if } y_n \geq \beta \end{cases} \end{aligned} \quad (10)$$

where the domain is  $[0, 1] \times [0, 1]$ ,  $\alpha_1 + \alpha_2 \leq 1$  and  $0 < \beta < 1$ . The visual representation of one step is shown in Fig. 14.

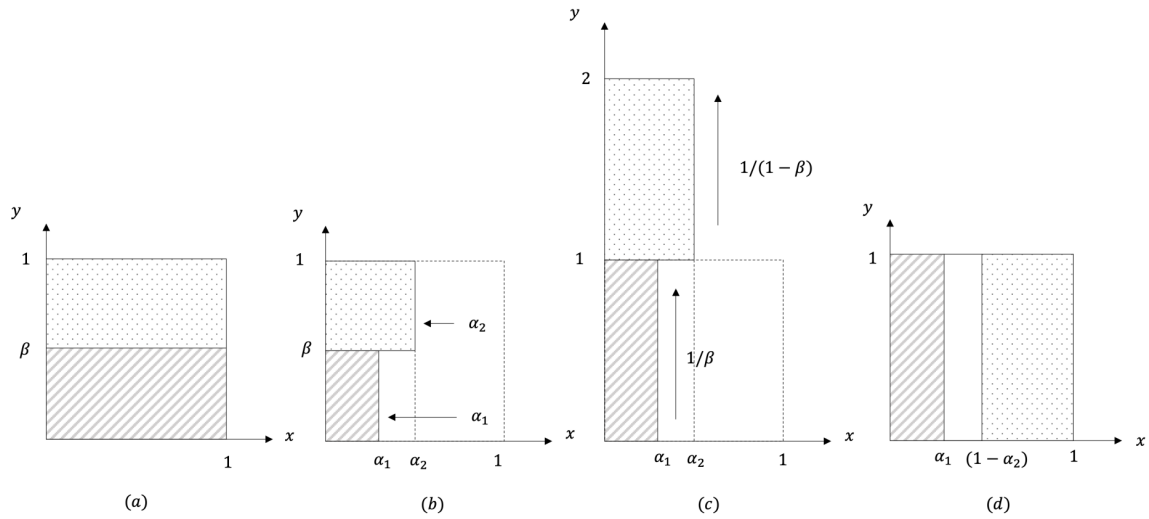


Figure 14: Graphical representation of the generalized Baker's Map step.

### 2.2 Analysis of the Dynamics

#### 2.2.1 Implementation of the Map and (Numerical) Observations

We start by simulating the evolution of this system for some general parameters. A graphical representation is shown in Fig. 15. First of all, thanks also to the colour distinction, we recognize soon the equivalence with the dynamics in Fig. 14. At every step, the map firstly distinguishes two regions based on the vertical coordinate. Later, it shrinks the two regions based on, respectively, the  $\alpha_1$  and  $\alpha_2$  parameters. After that, it vertically stretches the two regions and, in the end, it places the two new regions at the horizontal borders of the square. Further, what we notice is the continuous addition of new bands that get shorter at every step. The two initial colours, in particular, are separated in all the following bands.

Let us now focus on the shape of the attractors for different parameters. These are rectangles that have the full height but a horizontal length that depends on the parameters and the time step. Looking again at Fig. 14 we can define explicitly these horizontal coordinates:



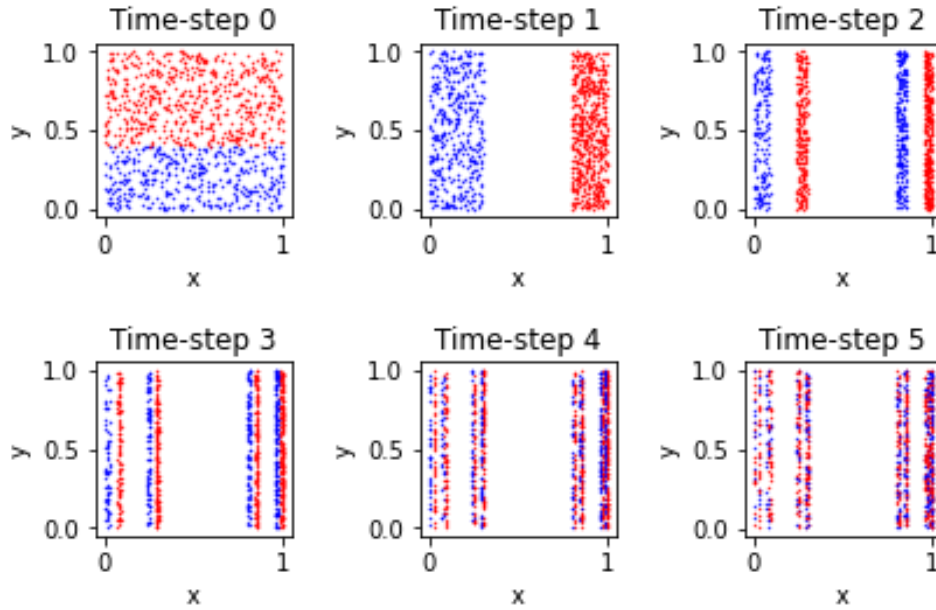


Figure 15: Simulated evolution of the map for  $\alpha_1 = 0.3, \alpha_2 = 0.2, \beta = 0.4$ . 1000 samples are initialized uniformly in the grid  $[0, 1] \times [0, 1]$ . The blue points are the ones with initial  $y_0 < \beta$ , red points the ones with  $y_0 \geq \beta$

- For  $n = 0$ : one band from 0 to 1.
- For  $n = 1$ : two bands from 0 to  $\alpha_1$  and from  $1 - \alpha_2$  to 1.
- For  $n \geq 2$ : the same coordinates as in  $n - 1$  multiplied by  $\alpha_1$  plus the same coordinates as in  $n - 1$  multiplied by  $\alpha_2$  and shifted of  $1 - \alpha_2$ .

As a confirmation, we have generated the map until step 3 and drawn the expected bands in Fig. 16. Here, it is evident the above statements are true since all the samples stay inside these borders.

Now that the general structure of the map has been described, we can analyse the chaotic nature of such system. As a first qualitative description, we take two points that are very close and look at the evolution of the system for such initial points. Two trajectories of this kind are shown in Fig. 17. Here it is clear that, despite the initial points are very close, eventually the two trajectories split in two different paths. We decided to plot explicitly the Euclidean distance between the trajectories for different parameters. This plot is shown in Fig. 18.

To conclude, two near points tend to diverge for every choice of the parameters (except for the special case  $\beta = 0.5$  as we will see later). Moreover, every trajectory does not convergence and continues to move in different part of the square (as seen in Fig. 17). We can conclude from these qualitative results that the generalized Baker's Map seems to be chaotic.

We remind that, in contrast with continuous systems, a 2-dimensional map can potentially be chaotic.

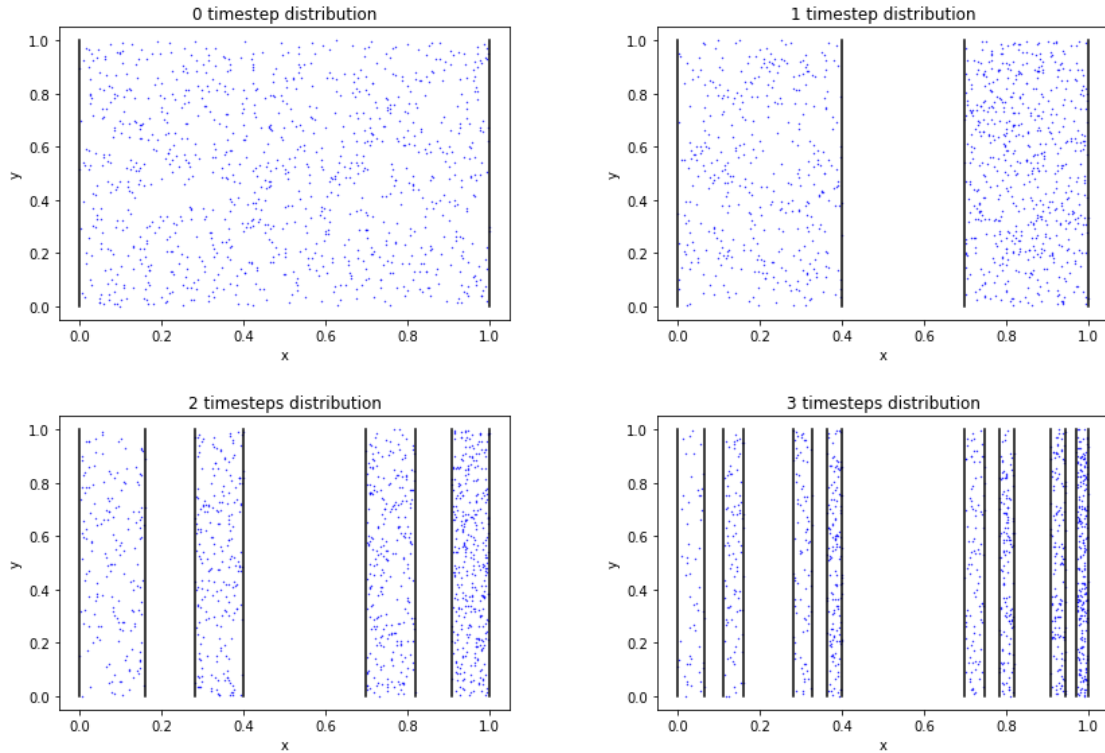


Figure 16: Simulated evolution of the generalized Baker's Map (for  $\alpha_1 = 0.4, \alpha_2 = 0.3, \beta = 0.4$ , 1000 samples) with the predicted coordinates for the bands.

### 2.2.2 Strange Attractor and Fractal Dimensions

In this section, we aim to prove that the map attractors are fractals and we aim to study their characteristics. To do that, we restrict our analysis to the case  $\alpha := \alpha_1 = \alpha_2$ . In this case, the bands described in section 2.2.1 have a simpler form since they are completely symmetric.

We are going now to compute the *box counting dimension* both analytically and numerically. The idea is to cover the square with a grid composed by many square boxes, then the dimension is defined as:

$$D_0 := \lim_{r \rightarrow 0} \frac{\log(1/N(r))}{\log r}$$

where  $r$  is the length of the little boxes and  $N(r)$  is the number of boxes that cover the fractal object. In this case, nothing needs to be done along the  $y$  axis that is always all covered. Therefore, we can compute the dimension as  $D_0 = 1 + D_{0,x}$  where  $D_{0,x}$  is the dimension if we consider only the  $x$  axis. In addition, it is clear that  $\beta$  has no influence in the box counting dimension (bands coordinates depend only on  $\alpha$  as we saw before).

First of all, let us find a suitable box length for each time step. We aim to get the biggest length such that boxes are aligned with the bands. In this contest, suppose  $\alpha \in \mathbb{Q}$ , so it is defined by two natural numbers such that  $\alpha = a/b$ . We can recursively observing that:

- $n = 0$ : the band is in  $[0, 1]$ , therefore one box of length  $r = 1$  contains all the information.
- $n = 1$ : the bands edges are at the coordinates  $\{0, \alpha, 1 - \alpha, 1\} = \{0, a/b, (b - a)/b, 1\}$ . Therefore, boxes of length  $r = 1/b$  can be perfectly placed to align the bands ( $a$  boxes in the first band,  $b - 2a$  boxes in the second,  $a$  boxes in the last one.)

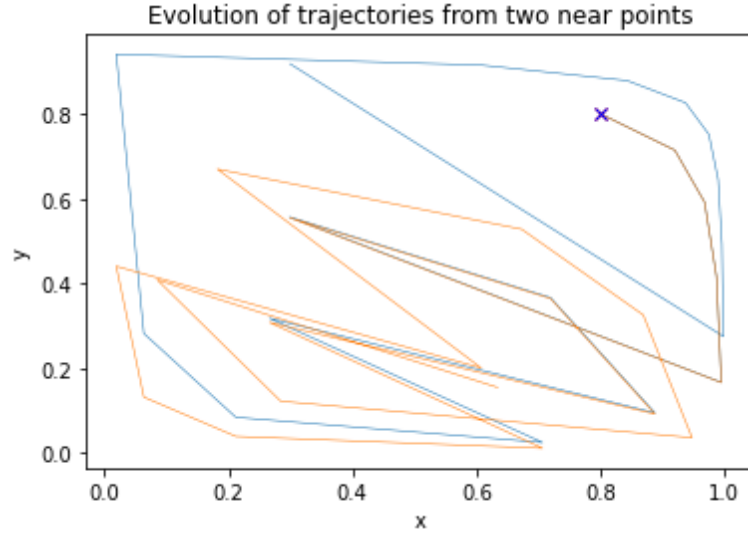


Figure 17: Simulated evolutions taking two very close points, i.e.  $[0.8, 0.8]$  and  $[0.7999, 0.7999]$  ( $\alpha_1 = 0.3, \alpha_2 = 0.4, \beta = 0.3$ ). The two points eventually take two independent paths.

- $n = 2$ : the bands edges are at the coordinates  $\{0, \alpha^2, (1 - \alpha)\alpha, \dots\}$ . The same reasoning can be adopted to reveal that  $r = 1/b^2$  is a divisor of all the coordinates and therefore boxes can be aligned with the bands.
- $n \geq 3$ : by induction, boxes of size  $1/b^n$  are aligned with the bands.

Having detected a good value for the box sizes at each step  $n$ , we know perform the recursive box counting:

- $n = 0$ : one band in  $[0, 1]$  and one box in  $[0, 1]$  therefore, trivially, 1 box contains all the points.
- $n = 1$ : there are 2 bands of length  $\alpha$ . Since boxes and bands of length  $1/b$  are aligned, it means that every band contains  $\alpha/(1/b) = a$  boxes. Therefore,  $2a$  boxes are occupied.
- $n = 2$ : there are 4 bands, each of size  $\alpha^2$ . Therefore  $4 \cdot (\alpha^2/(1/b)^2) = 4a^2$  boxes are occupied.
- $n \geq 3$ : by induction,  $(2a)^n$  boxes are occupied.

We finally have both the box size and the number of occupied boxes for each  $n$ . The resulting box counting dimension is then:

$$D_0 := 1 + \lim_{r \rightarrow 0} \frac{\log(1/N(r))}{\log r} = 1 + \lim_{n \rightarrow \infty} \frac{\log(1/(2a)^n)}{\log(1/b)^n} = 1 + \frac{\log(2a)}{\log(b)}$$

Let us make some checks with two trivial cases:

1. If  $\alpha = 1/2$ , the bands correspond to a full cover of the square at each time step (indeed,  $\alpha$  coincides with  $1 - \alpha$ ). Therefore, the attractive set is the full square at every time step. Not by chance, with  $b = 2a$  we get  $D_0 = 1 + \log(2a)/\log(2a) = 2$ . I.e., the attractive set is not a strange attractor but the normal 2-dimensional square.
2. If  $\alpha = 1/3$ , one can easily see that the limit to the strange attractor along  $x$  is the same as the construction of the Cantor set (indeed, it starts with three equally wide bands). In this case,  $D_{0,x} = \log(2)/\log(3)$  that is the same dimension of the Cantor set.

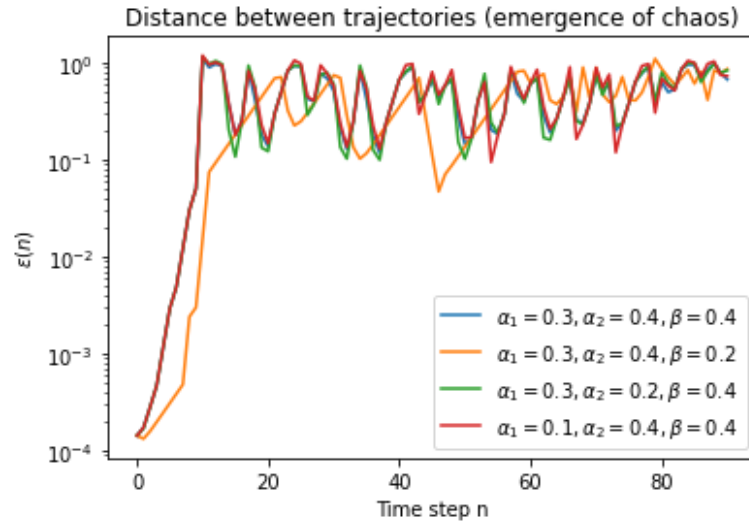


Figure 18: Euclidean distances between trajectories starting at  $[0.8, 0.8]$  and  $[0.7999, 0.7999]$ . Chaos emerges for every choice of parameters. A lower value for  $\beta$  seems to reduce a bit the speed of divergence.

To achieve a numerical result, we simulate the trajectory for  $5 \cdot 10^6$  uniformly distributed initial points and we use boxes of the size defined above. In particular, we perform the box counting operation for  $n = 5$ . There are other possible choices for  $n$  since, from the theory above, for every  $n \geq 1$  we should get the same  $D_0$ . However, we increase the value  $n$  to assure the  $\lim_{n \rightarrow \infty}$  operation. On the other hand, bigger values for  $n$  would imply smaller boxes and therefore huge memory allocations. This is the reason why we considered  $n = 5$  as a reasonable value.

For the implementation details, it is easy to verify which box the point belongs to using the division of the coordinate over the box length without the remainder. Results are shown in table 7.

Table 7: Analytical and numerical box counting dimensions

$\alpha$	Theoretical value	Numerical value
1/5	$1 + \frac{\log(2)}{\log(5)} \approx 1.4307$	1.4307
1/3	$1 + \frac{\log(2)}{\log(3)} \approx 1.631$	1.631
2/5	$1 + \frac{\log(4)}{\log(5)} \approx 1.861$	1.824
1/2	$1 + \frac{\log(2)}{\log(2)} = 2$	2.000

As expected, except for  $\alpha = 2/5$  where some bands are probably not covered by the multiple trajectories, in all the other cases all the bands are correctly covered by points and therefore the numerical solution gives the exact analytical solution. Again, the case  $\alpha = 1/3$  and  $\alpha = 1/2$  give the values that are well-known also without the detailed study of the strange attractors.

### 2.2.3 Chaos and Lyapunov Exponents

Other important characteristics of chaotic systems are well described by the Lyapunov exponents. We aim to compute these values analytically and numerically in the case of  $\alpha = \alpha_1 = \alpha_2$  and  $\beta = 1/2$  and compare with the behaviour of the system. Consider an initial displacement  $\vec{\epsilon}_0$  as we already saw in

Fig. 18. Then, the Lyapunov exponent is defined as:

$$\lambda := \lim_{n \rightarrow \infty} \lim_{\vec{\epsilon}_0 \rightarrow 0} \frac{1}{n} \log \left| \left( \prod_{j=0}^{n-1} J_M(\vec{x}_j) \right) \frac{\vec{\epsilon}_0}{|\vec{\epsilon}_0|} \right|$$

where  $J_M$  represents the Jacobian of the map. We know that for a 2-dimensional map there are 2 Lyapunov exponents and, indeed, the value of  $\lambda$  depends on the choice of the displacement direction. Let us firstly compute the Jacobian of the map. A quick analysis to equation 10 with the conditions  $\alpha = \alpha_1 = \alpha_2$  and  $\beta = 1/2$  reveals that the partial derivatives are the same in both the cases  $y_n < \beta$ ,  $y_n \geq \beta$ . Hence, the Jacobian has the following simple form:

$$J_M(\vec{x}_n) = \begin{bmatrix} \alpha & 0 \\ 0 & 2 \end{bmatrix} \quad \forall n \geq 0$$

To detect the two exponents, we choose the displacements as parallel to the eigenvectors of  $J_M$  that are trivially  $\vec{h}_1 = [1, 0]$  and  $\vec{h}_2 = [0, 1] \forall n \geq 0$ . Therefore:

$$\begin{aligned} \lambda_1 &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \begin{bmatrix} \alpha^n & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right| = \lim_{n \rightarrow \infty} \frac{1}{n} \log(\alpha^n) = \log(\alpha) < 0 \quad \forall \alpha \leq 0.5 \\ \lambda_2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \begin{bmatrix} \alpha^n & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right| = \lim_{n \rightarrow \infty} \frac{1}{n} \log(2^n) = \log(2) \approx 0.693 \end{aligned}$$

As confirmed by the theory, the Lyapunov exponents are represented by the logarithm of the eigenvalues. These values have also been computed numerically for  $n = 50$  with different choices for the initial displacement  $\epsilon_0$  and  $\alpha = 0.4$ . Results are shown in table 8.

Table 8: Analytical and numerical Lyapunov exponents for different initial displacements

$\epsilon_0$	Theoretical value	Numerical value
[0.1, 0]	$\approx -0.916$	-0.916
[0, 0.1]	$\approx 0.693$	0.693
[0.5, 0]	$\approx -0.916$	-0.916
[0, 0.5]	$\approx 0.693$	0.693
[0.2, 0.5]	$\approx 0.693$	0.692
[0.1, 0.05]	$\approx 0.693$	0.677

Results are expected from the theory. Indeed, every initial displacement that is not parallel to the first eigenvalue generates a limit that tends to be bigger eigenvalue and, hence, to the bigger Lyapunov exponent. In formulae, every  $\vec{\epsilon}_0$  that is not proportional to  $[1, 0]$  returns the second Lyapunov exponent, otherwise the first.

Let us see now the relation between these values and the chaotic nature of the system. First of all, as expected, one exponent is positive. Indeed, chaos is generated only if at least one exponent is positive. Not by chance, the positive exponent is the one related to the  $y$  direction. It is indeed in this direction that bands are stretched while they are shrunk along the  $x$  direction. In addition, if we sum the exponents we get a negative value except for  $\alpha = 0.5$  that gives zero sum. This should not surprise since, as already seen in Fig. 15 and 16, the total area tends to decrease and therefore the system is dissipative. On the contrary, the choice of  $\alpha = 0.5$  gives the full cover of the square (fractal dimension = 2 as seen before) and hence the area is constant and the system conservative.

In the specific case of  $\beta = 0.5$ , the system presents a very singular behaviour that can be well explained

by the Lyapunov exponents. First of all, we show in Fig. 19 the trend of the displacement in this scenario.

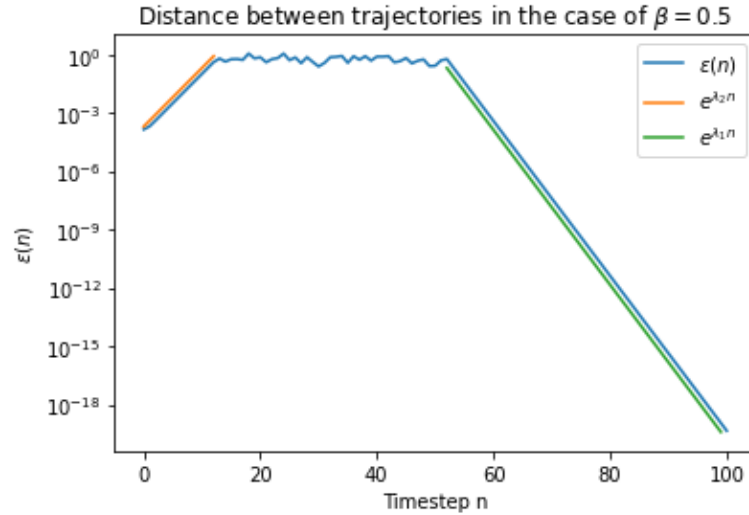


Figure 19: Euclidean distances between trajectories starting at  $[0.8, 0.8]$  and  $[0.7999, 0.7999]$  for  $\beta = 0.5$  and  $\alpha = 0.4$ . Chaos emerges but, after some iterations, the two trajectories come back nearby. In addition, the increase and decrease of distance follow exactly the two Lyapunov exponents.

The fact that the distance initially increases with the same scale of  $e^{\lambda_2 n}$  ( $\lambda_2 = \log(2)$ ) is clearly expected from the theory but what is not expected is the decrease in the second part. This is observed only in the case  $\beta = 0.5$  and therefore, it is necessary to look at the law of  $y$  more in details. Equation 10 applied to this specific case and only for the  $y$  component is the following:

$$y_{n+1} = \begin{cases} 2y_n & \text{if } y_n < 0.5 \\ 2y_n - 1 & \text{if } y_n \geq 0.5 \end{cases}$$

What is observed from the simulations is that  $y$  eventually gets the exact 0 value. This is clearly not possible from the theory because the point  $(x, y) = [0, 0]$  is not a stable equilibrium point. Therefore, this result should be motivated by computational reasons. We know that every number  $y \in [0, 1]$  is represented by an infinite series of digits  $\{I_i\}_{i \in \mathbb{N}}$  such that  $I_i \in \{0, 1\}$  and:

$$y = \sum_{i=1}^{\infty} \frac{I_i}{2^i}$$

In computer memories, however, this series is not infinite but interrupted at a certain  $i = N$ . Let us see now the effect of the map on one number with this representation. Hence, we take  $y_n$  in this form and distinguish the two cases.

1.  $y_n < 0.5$ : this implies  $I_1 = 0$ . Therefore:

$$y_{n+1} = 2y_n = \sum_{i=2}^N \frac{2I_i}{2^i} = \sum_{i=1}^{N-1} \frac{I_{i+1}}{2^i}$$

2.  $y_n \geq 0.5$ : this implies  $I_1 = 1$ . Therefore:

$$y_{n+1} = 2y_n - 1 = 2 \left( \frac{1}{2} + \sum_{i=2}^N \frac{I_i}{2^i} \right) - 1 = \sum_{i=2}^N \frac{2I_i}{2^i} = \sum_{i=1}^{N-1} \frac{I_{i+1}}{2^i}$$

Hence, in both the cases, the following value for  $y$  has the same representation except one term of the sum less. This means that, if  $y_0$  is represented by  $N$  digits (this depends on the way memory is allocated), then  $y_1$  consists of  $N - 1$  digits,  $y_2$  is represented by  $N - 2$  digits until, after  $N$  steps,  $y_N$  is represented by 0 digits, i.e.  $y_N = 0$ .

As expected, this is not a result provided by the theory but due to computational reasons. We can now focus on the role of the first Lyapunov exponent.

We already know that the system is chaotic because of the positive second Lyapunov exponent that is related to the  $y$  direction. However, in this specific case, we know that  $y_n$  eventually gets the exact zero value and, so, it will not change its value after the step  $N$ . This means that our original system transforms to a 1-dimensional system where the  $y$  component is fixed to zero. Explicitly, the system becomes:

$$\begin{cases} x_{n+1} = \alpha x_n \\ y_{n+1} = 0 \end{cases} \quad \text{if } n > N$$

It is now easy to compute the dynamics of the system since we can only focus on  $x$ . This means that  $\lambda_2$  does not affect the system no more and the only Lyapunov exponent left is  $\lambda_1$ , that is negative. Therefore, the system cannot be chaotic any more. Indeed,  $x = 0$  is a stable equilibrium point and we expect trajectories to converge to it. At this point, the speed of convergence is as usual governed by  $e^{\lambda_1 n}$  as confirmed in Fig. 19.

### 2.3 Discussion (Limit: 1/2 page)

## Appendix

### List of Sources

- [1] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36:1627–1639, January 1964.
- [2] Arsenii Trush, Stanislav Pospisil, and Hrvoje Kozmar. Comparison of Turbulence Integral Length Scales Determination Methods. *WIT Transactions on Engineering Sciences*, 128:113–123, 2020.
- [3] Kyunghwan Yoon and Z. Warhaft. The evolution of grid-generated turbulence under conditions of stable thermal stratification. *Journal of Fluid Mechanics*, 215:601–638, 1990. doi: 10.1017/S0022112090002786.

### List of Collaborators

Anna Peruso

### Personal Statement

I hereby certify that I fully respect the stated Honor Code and specifically that:

1. My report is my original work prepared solely by me;
2. All sources used are cited;
3. All people I collaborated with are listed.

---

Signature (Matteo Calafà)

---

Date