POLITECNICO DI TORINO

Master's degree course in Aerospace Engineering

Master Thesis

A Lagrangian view of turbulent mixing in channel flow through complex networks

Supervisors
prof.ssa Stefania Scarsoglio
prof. Luca Ridolfi

Candidate
Davide Perrone

Academic Year 2018-2019

Abstract

Turbulent mixing and diffusion are investigated in the Lagrangian framework using tools proper to network theory. Starting from a database of trajectories of non-inertial particles obtained through direct numerical simulation of a channel flow, two innovative analysis tools are developed and discussed.

The first tool aims to characterize the tendency of tracers to group in coherent structures, or *clusters*, despite being driven further and further away by the diffusive action of turbulence. The second tool deals with the diffusion and transport of particles in the only inhomogeneous direction of the channel flow, which is the wall normal one.

In both tools particles (or sets of them) are represented as graph nodes, while different definition of links are given. Network theory offers a powerful framework to analyse complex dynamical behaviour in large datasets.

Contents

List of Figures VI

1 Introduction 1

2 Turbulence and mixing 3
  2.1 Turbulent flows 3
    2.1.1 Governing equation 4
    2.1.2 Reynolds number 4
    2.1.3 Kolmogorov theory 5
    2.1.4 Wall flows 6
    2.1.5 Near wall region 7
  2.2 Lagrangian approach and mixing 10
    2.2.1 Trajectories 10
    2.2.2 Acceleration and intermittency 11
    2.2.3 Particle dispersion 12

3 Turbulence and networks 15
  3.1 Network theory 15
    3.1.1 Degree and strength 15
    3.1.2 Other matrices and graph spectrum 16
  3.2 Applications to turbulent flows 18

4 Channel flow dataset 21
  4.1 Channel flow DNS 21
  4.2 Trajectories integration 22
  4.3 Distance statistics 22
  4.4 Velocity statistics and Lagrangian timescale 26
    4.4.1 Probability distributions 26
    4.4.2 Velocity correlations and timescales 27
    4.4.3 Taylor microscale 28
  4.5 Acceleration statistics 31
    4.5.1 Acceleration timescale 31

5 Clustering network 33
5.1 Definitions 33
5.2 Single time clustering 35
5.3 Clustering along the x and y direction 40
5.3.1 x-clustering 40
5.3.2 y-clustering 41
5.4 Threshold value and degree evolution 43
5.4.1 Variable threshold 43
5.4.2 Fixed threshold 44
5.5 Temporal and spatial evolution 48
5.5.1 Evolution of marginal probability density functions 48
5.5.2 Onset and persistence of clustering 55
5.5.3 Spatial inhomogeneity 57

6 Flow map network 63
6.1 Transport mapping 63
6.2 Network building and an example 64
6.2.1 Strength and degree for the transport network 67
6.3 Network evolution 69
6.3.1 Strength and degree evolution 69
6.3.2 Predominant links 76
6.3.3 Network partitioning 79

7 Conclusions 83

List of Figures

2.1 Flow structures inside channel flow 3
2.2 Channel schematics 6
2.3 Viscous and Reynolds stresses, channel flow at $Re_\tau = 590$. DNS data from Moser, Kim, and Mansour [14] 8
2.4 Velocity profile in a channel flow at $Re_\tau = 950$. DNS data from Kuerten and Brouwers [13] 9
2.5 Reynolds stresses in a channel flow at $Re_\tau = 590$. DNS data from Moser, Kim, and Mansour [14] 10
2.6 Trajectory of a single tracer in a turbulent channel. DNS data from Kuerten and Brouwers [13] 11
2.7 Trajectory of a single tracer with its acceleration superimposed. DNS data from Kuerten and Brouwers [13] 12
3.1 Example graph G 17
3.2 First non trivial eigenvector of G 18
4.1 Mean distance of particles 23
4.2 Mean distance of particles divided by its standard deviation 24
4.3 Skewness of the distance of particles 25
4.4 Kurtosis of the distance of particles 25
4.5 Probability distribution of u^+, $t^+ = 100$ 26
4.6 Mean axial velocity for different y_0^+ 27
4.7 Autocorrelation ρ_{11} for two different y_0^+ 28
4.8 Velocity timescales T_{ii}^+ in wall time units 29
4.9 Graphical depiction of the temporal Taylor microscale 30
4.10 Taylor scale λ_i for different particle release coordinates 30
4.11 Probability distribution function of a_y^+/σ 31
4.12 Acceleration timescale in wall units of a_x, a_y and a_z for different y_0^+ 32
4.13 Ratio T_{ii}^v/T_{ii}^a between velocity and acceleration timescales 32
5.1 Network matrix $\widetilde{\mathbf{\Delta D}}$, $t^+ = 475$, $d_t^+ = 100$ 34
5.2 Detail of the structure of $\widetilde{\mathbf{\Delta D}}$ 35
5.3 Particle snapshot and clustering links at $t^+ = 475$, $d_t^+ = 100$ 36
5.4 Joint probability distribution of degree and wall normal coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$ 37
5.5 Joint probability distribution of distance decrease and centroid wall normal coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$ 38

5.6 Joint probability distribution of distance decrease normalized by mutual distance and mutual distance with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$ . . . 38
5.7 Joint probability distribution of distance decrease normalized by mutual distance and centroid coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$ . . . 39
5.8 Joint probability distribution of y_c^+ and r^+, with its trendline, $t^+ = 475$, $d_t^+ = 100$ . . . 39
5.9 Joint probability distribution of degree and wall normal coordinate, x-clustering, $t^+ = 475$, $d_t^+ = 100$ . . . 40
5.10 Joint probability distribution of distance decrease and centroid wall normal coordinate, x-clustering, $t^+ = 475$, $d_t^+ = 100$ . . . 41
5.11 Joint probability distribution of distance decrease and mutual distance, x-clustering, $t^+ = 475$, $d_t^+ = 100$ . . . 42
5.12 Joint probability distribution of distance decrease and mutual distance, y-clustering, $t^+ = 475$, $d_t^+ = 100$ . . . 42
5.13 Mean degree, variable threshold . . . 43
5.14 Degree standard deviation, variable threshold . . . 44
5.15 Mean degree, fixed threshold . . . 45
5.16 Degree standard deviation, fixed threshold . . . 45
5.17 Mean degree normalized by its standard deviation, fixed threshold . . . 46
5.18 Kurtosis of k, fixed threshold . . . 46
5.19 Evolution of $\langle k^* \rangle$ for different threshold values . . . 47
5.20 Degree of nodes . . . 48
5.21 Non-zero elements of $\widetilde{\boldsymbol{\Delta}\mathbf{D}}$, weight superimposed (constant threshold $d_t^+ = 350$) 49
5.22 Probability distribution function of k for different times . . . 50
5.23 Exponential distribution coefficient λ . . . 51
5.24 Probability distribution function of y_c^+ for different times . . . 51
5.25 Probability distribution function of r^+ for different times . . . 52
5.26 Probability distribution function of ΔD^+ for different times . . . 52
5.27 Mean centroid wall normal coordinate . . . 53
5.28 Mean mutual distance normalized by threshold value . . . 54
5.29 Mean link strength . . . 54
5.30 Mean link strength normalized by mutual distance . . . 54
5.31 Mean duration of links $\langle T_l^+ \rangle$ . . . 55
5.32 Mean distance reduction for full links $\langle \Delta D_l^+ \rangle$ . . . 56
5.33 Mutual distance of a pair of particles. Red circles: start and end of the clustering event; blue circle: $t^+ = 9500$ . . . 56
5.34 Mean duration of links $\langle T_l^+ \rangle$, threshold included . . . 57
5.35 Mean distance reduction for full links $\langle \Delta D_l^+ \rangle$, threshold included . . . 57
5.36 Distribution of tracers in the yz plane, $t^+ = 950$, $d_t^+ = 350$ . . . 58
5.37 Distribution of links in the yz plane, $t^+ = 950$, $d_t^+ = 350$ . . . 59
5.38 Koch's snowflake, $\mathcal{D} \approx 1.26$ . . . 60
5.39 $N(\varepsilon)$ versus ε and fractal dimension, $t^+ = 950$, $d_t^+ = 350$ . . . 60

5.40 Fractal dimension $\mathcal{D}$ versus time, $d_t^+ = 350$ 61
6.1 Distribution of particles in the channel; the starting level is color coded . . 63
6.2 Evolution of $\mathbf{P}$ 65
6.3 Balance between N^k and N^0, initial transient 66
6.4 Graph of the example flow network at $t = t_1$ 66
6.5 Graph of the example flow network at $t = t_2$ 67
6.6 Probability distribution of the ingoing strength 70
6.7 Probability distribution of the outgoing strength 70
6.8 Standard deviation of s, both ingoing and outgoing 71
6.9 Evolution of the ingoing strength 71
6.10 Probability distribution of k^{IN} 72
6.11 Probability distribution of k^{OUT} 72
6.12 Evolution of the ingoing degree 73
6.13 Evolution of the outgoing degree 73
6.14 Mean degree $\langle k \rangle$ 74
6.15 Degree standard deviation 74
6.16 Mean link length 75
6.17 Mean outgoing weighted link length 76
6.18 Graph of the example flow network, cumulative representation 77
6.19 Strength distribution of the cumulative network, $t_0^+ = 0$, $t_f^+ = 15\,200$ 77
6.20 Cumulative transport matrix $\mathbf{P}^c$, $t_0^+ = 0$, $t_f^+ = 15\,200$. 78
6.21 Eigenvector plot, non cumulative flow network at $t^+ = 712$. The colouring identifies the origin zones; red: $0 < y^+ < \delta/2$, blue: $\delta/2 < y^+ < \delta$, yellow: $\delta < y^+ < 3\delta/2$, green: $3\delta/2 < y^+ < 2\delta$ 80
6.22 Eigenvector plot, non cumulative flow network at $t^+ = 9000$. Colour codes as in 6.21 81
6.23 Eigenvector plot, non cumulative flow network at $t^+ = 15\,200$. Colour codes as in 6.21 81
6.24 Eigenvector plot, cumulative flow network from $t^+ = 0$ to $t^+ = 15\,200$. Colour codes as in 6.21 82

Chapter 1

Introduction

Transport phenomena arising in turbulent flows permeate several aspects of our lives, spanning from the diffusion of pollutants in the atmosphere to the combustion of fuel in an aircraft's jet engine. While diffusion of chemical species, momentum and heat takes place even in the absence of turbulence, all of these processes are vastly amplified by its multiscale and chaotic nature.

Despite the great importance of turbulent transport in many of today's engineering and physics problems, several questions remain open. Practical applications often resort on models which are by their own nature incomplete and bounded to a restricted number of cases. On the other hand, the complete resolution of the structure of a flow, *i.e.* the solution of the Navier-Stokes equations for any boundary condition, is not possible in an analytical way nor computationally feasible with a numerical approach. Nonetheless, the numerical solution of the Navier-Stokes equations for selected, simplified cases (such as homogeneous isotropic turbulence or the flow in a channel) provides useful insight on more complex real-life scenarios and is of prime importance in the definition of simplified models [15]. Along with numerical techniques, a great number of data is collected through experiments; relatively new methods, such as Particle Image Velocimetry, offer the possibility to investigate various laboratory flows with high spatial and temporal resolution [26].

The aforementioned approaches generate huge amounts of data, which need to be adequately processed in order to extract useful information. Besides the dimension of the datasets involved, a key factor pushing forward the research of novel post-processing methods is the need to effectively highlight the prominent features of turbulent flows.

One of the most recent advancements amongst these methods is offered by complex networks and their underlying mathematical basis, that is graph theory [1]. It has been shown that complex networks are particularly suitable to describe the mutual interactions of systems composed by a great number of dynamical elements; similar approaches have been already applied to a number of different cases, ranging from biology to the study of social groups. Its application to the study of turbulence is rather recent but promising, since the dynamical and interconnected behaviour of turbulent flows seems suitable for a network representation [17] [9] [7] [18] [20].

The purpose of this work is to develop and test two innovative methods of analysis for wall bounded turbulence, starting from high resolution Lagrangian data and employing complex networks. The results provided by these methods will be accompanied by classical results; also, additional insights on channel turbulence will be searched, in order to exploit the peculiarities of network theory to highlight hidden features of such flows. In general, the capability of graph theory to thoroughly describe turbulent flows will be assessed.

Chapter 2 introduces turbulent flows, their main characteristic and in particular wall bounded turbulence, which is the main subject of this thesis.

Chapter 3 treats the theory of networks, introduces usual metrics needed to describe them and reviews some relevant applications to fluid dynamics.

Chapter 4 analyses the dataset employed in this work, providing some classical results about trajectories, their velocities and accelerations in the Lagrangian viewpoint. The Lagrangian data has been obtained through the direct numerical simulation code of Kuerten and Brouwers [13].

Chapter 5 and 6 constitute the main subject of this thesis; two network-based methods to characterize wall bounded turbulence are introduced. The first method, treated in chapter 5, defines a network starting from particles that get closer one to the other during their motion. The second, addressed in chapter 6, aims to follow the motion of tracers (grouped according to their release location) in the inhomogeneous direction of the channel, which is y.

Chapter 7 provides some closing remarks and suggests some future developments of this subject.

Chapter 2

Turbulence and mixing

2.1 Turbulent flows

The main feature of turbulent flows is their unsteady, irregular and seemingly unpredictable motion; still, they are regulated by a deterministic set of equations and are known to exhibit some common features and repetitive structures. Figure 2.1 shows both the pseudochaotic behaviour and the presence of the previously mentioned structures, which usually appear to be coherent both in time and space and are thus defined *coherent structures*.

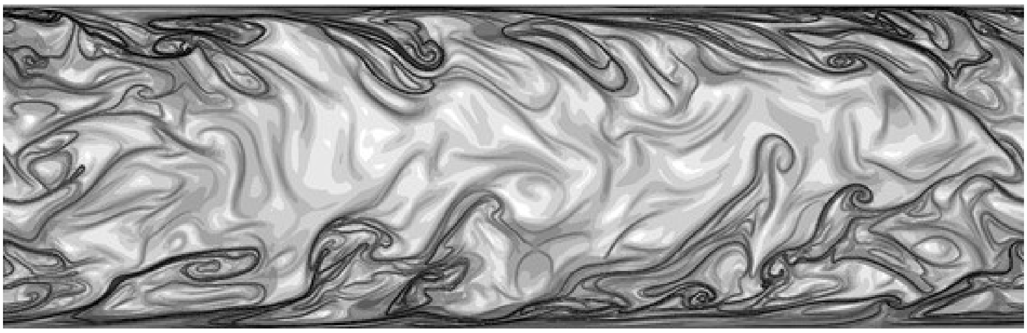


Figure 2.1: Flow structures inside channel flow

While the random behaviour makes the analysis of flow features difficult and usually forces to rely on statistical tools, the existence of repeated patterns and similarities between different cases makes it useful to search for such features, in order to provide a simpler and more immediate characterization of the flow.

Similarities in homogeneous turbulence and in bounded flows are perhaps the most common amongst these phenomena and thus the most studied. Their existence stems directly from several features embedded in the equations of motion of viscous fluids, which are the Navier-Stokes equations.

2.1.1 Governing equation

The Navier-Stokes equations for incompressible, Newtonian fluids are, in cartesian coordinates,

$$\nabla \cdot \mathbf{v} = 0 \tag{2.1}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\rho} \nabla p + \nu \Delta \mathbf{v}, \tag{2.2}$$

where ν is the kinematic viscosity of the fluid [15]. While the right-hand side of the momentum equation (2.2) accounts for the action of normal and tangential stresses, the left-hand side, in particular the non linear term $\mathbf{v} \cdot \nabla \mathbf{v}$, is responsible for the transfer of energy between different scales. This transfer is the foundation for the onset of turbulence, because it is the mechanism by which larger structures embedded in a flow break up and generate other, smaller structures and so on.

The generation of smaller and smaller scales of motion is accountable for the enhanced mixing properties of turbulent flows. The unsteady fluctuations of the velocity field, together with the presence of vortices of every size, radically change the way passive scalar quantities are transported and diffused. It should be noted that such process is way more effective than diffusion caused by thermal agitation.

2.1.2 Reynolds number

The Navier-Stokes equations can be made non dimensional introducing appropriate scales for length L, velocity U and pressure ρU^2, so that $\mathbf{x} = L\mathbf{x}^\star$, $\mathbf{v} = U\mathbf{v}^\star$ and $p = \rho U^2 p^\star$, where the star denotes the non dimensional variable and will be omitted in the following. While the continuity equation is left formally unchanged by the substitution, the momentum equation (2.2) becomes

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \frac{\nu}{UL} \Delta \mathbf{v} \tag{2.3}$$

where UL/ν is defined as the Reynolds number Re. This parameter is proportional to the ratio between inertial and viscous forces, the latter having a stabilizing effect on the dynamics of the flow.

Many cases, such as boundary layers and the flow in a pipe or around a smooth sphere, exhibit a sharp shift in certain properties (most notably, the forces exchanged between fluid and body) for definite values of the Reynolds number. This transition is due to prevailing inertial forces which overcome viscous effects and lead the flow to instability. With growing Re the regularization of the solution induced by the Laplacian of the velocity $\Delta \mathbf{v}$ is diminished, while the growing importance of non linear terms starts the breakup of larger structures and the generation of smaller scales.

It should be noted that, in the limit of very large Reynolds number, the dissipation does not become negligible, since the minimum scales of the flow become smaller and the associated velocity gradients larger.

2.1.3 Kolmogorov theory

All the previously treated concepts receive a formal framework in the theory developed by Kolmogorov in 1941 [11]. The process by which large vortices generate smaller ones and transfer energy to them is called *energy cascade* and was already recognized by Richardson [16]. This transfer is not endless but stops when flow structures have reached a size and a Reynolds number so small that viscosity is effective in dissipating their kinetic energy into heat.

Kolmogorov theory is stated through three main hypotesis, which can be summarised as follow. The first statement is the

Hypothesis of local isotropy *At large Reynolds number, the smaller scales of motion are statistically isotropic*

While the larger scales of size $\mathcal{L}$ are strictly dependent on the boundary conditions, going lower and lower in the cascade the smaller structures lose memory of the external geometry and retain no directional information. This is only true for high *Re* since otherwise there would not be enough separation between scales for this to happen.

It is useful to introduce the dissipation rate ε, that is the ratio at which energy is transferred between vortices of different size. Since dissipation happens only at the bottom of the cascade, ε is assumed to be constant through the cascade and equal to that injected into the system by gradients in the mean flow. The loss of information from the mean flow makes the lesser scales of every turbulent flow *similar* if rescaled with appropriate parameters. The most important features at this state are the injection of energy and its dissipation; this leads to the

First similarity hypothesis *At the locally isotropic scale, all the statistics of the flow are determined uniquely by ν and ε*

Starting from ν and ε, unique length and velocity scales can be derived. In particular the characteristic length scale at which dissipation happens can be defined as $\eta = \sqrt[4]{\nu^3\varepsilon}$ and is the *Kolmogorov scale.*

At very high Reynolds number, it is reasonable to assume the existence of a scale of characteristic length $l \ll \mathcal{L}$, so already statistically isotropic, yet enough large ($l \gg \eta$) to not experience significant dissipative phenomena. This is the

Second similarity hypothesis *In turbulent flows at very high Reynolds number, the statistics of the scale l with $\eta \ll l \ll \mathcal{L}$ are universal and dependent only on ε, not on ν*

This range is called the *inertial range*, since dissipative effects are negligible and only inertial ones are considered. Following from these considerations and from dimensional arguments, a simple power law for the energy spectrum can be obtained:

$$E(k) = C_0 \epsilon^{2/3} k^{-5/2}, \tag{2.4}$$

where C_0 is a universal (Kolmogorov) constant.

2.1.4 Wall flows

While of great theoretical importance, Kolmogorov's predictions often fall short when employed in most naturally occurring flow cases. Most notably, the need for turbulence to be statistically homogeneous both in time and space excludes cases where gradients are present. The presence of a wall introduces a boundary condition on velocity

$$\mathbf{v} = 0 \text{ at } y = 0, \tag{2.5}$$

where y is the wall normal coordinate; equation (2.5) is the *no-slip* condition. The presence of a non zero velocity away from the wall generates what is called a boundary layer, *i.e.* a region in the vicinity of a solid boundary where, even at high Reynolds number, viscosity effects are not negligible because of the great intensity of the velocity gradients caused by the boundary.

One of the simplest cases of flow in presence of a wall is the fully developed channel, that is the flow in a rectangular channel with one dimension much smaller than the other two; entrance effects are also neglected, so a portion of the channel where motion is statistically steady is considered. This configuration will be the main subject of this thesis. As can be seen in figure 2.2, the channel has height 2δ, while dimensions b and L are much larger ($b, L \gg \delta$). The mean flow is non zero only in its streamwise component (that is x), while variations in the mean flow happen only in the wall normal y direction. Since the spanwise dimension z is much larger than the wall normal one, far from the $\pm b/2$ walls the mean flow is independent of z. The mean flow has components U, V and W with only $U \neq 0$ and $\mathrm{d}U/\mathrm{d}y \neq 0$; the channel flow is also statistically symmetrical along the center plane located at $y = \delta$. In numerically simulated channels the $\pm b/2$ walls are usually neglected and substituted by periodic boundaries.

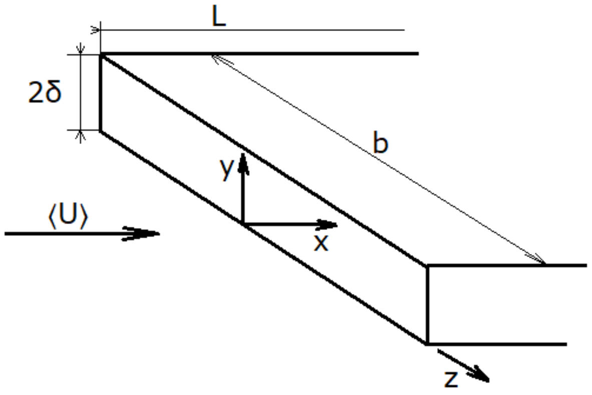


Figure 2.2: Channel schematics

Because of the mean continuity equation and of the boundary condition $\langle V\rangle_{y=0} = 0$, integration leads to $\langle V\rangle = 0$. Taking the mean of the momentum equation (2.2) and projecting it onto y,

$$0 = -\frac{1}{\rho}\frac{\partial\langle p\rangle}{\partial y} - \frac{\mathrm{d}\langle v^2\rangle}{\mathrm{d}y}, \tag{2.6}$$

Since also the fluctuating velocity is zero on the wall ($\langle v^2 \rangle_{y=0} = 0$), equation (2.6) can be integrated and reduces to

$$\langle v^2 \rangle + \frac{\langle p \rangle}{\rho} = \frac{p_w(x)}{\rho}.$$

The axial projection of the mean momentum equation is

$$0 = \nu \frac{\mathrm{d}^2 \langle U \rangle}{\mathrm{d}y^2} - \frac{\mathrm{d}\langle uv \rangle}{\mathrm{d}y} - \frac{1}{\rho} \frac{\partial \langle p \rangle}{\partial x}; \tag{2.7}$$

defining the total shear stress τ as

$$\tau = \rho \nu \frac{\mathrm{d}\langle U \rangle}{\mathrm{d}y} - \rho \langle uv \rangle, \tag{2.8}$$

equation (2.7) can be rewritten as

$$\frac{\mathrm{d}\tau}{\mathrm{d}y} = \frac{\mathrm{d}p_w}{\mathrm{d}x}$$

The flow is statistically steady and this results in equation (2.7) being a simple balance of forces between the mean axial pressure gradient and the cross-stream shear stresses. Shear stresses meanwhile are a function only of y, while the wall pressure is a function only of x, so integration of equation (2.7) is straightforward and results in the solution

$$-\frac{\mathrm{d}p_w}{\mathrm{d}x} = \frac{\tau_w}{\rho} \tag{2.9}$$

where $\tau_w = \tau(y = 0)$ is the wall shear stress; shear stresses are zero at the channel centreline. The pressure drop is hence related to the shear stresses at the channel boundary.

2.1.5 Near wall region

As already stated in equation (2.8), the total shear stress τ is the sum of contributions from viscous and Reynolds stresses (the latter being related to turbulent velocity fluctuations). As can be seen in the data shown in figure 2.3 for a channel flow at $Re_\tau = 590$, at the wall ($y = 0$) the viscous stress is the highest (because of the velocity gradient) while the boundary condition $\mathbf{v} = 0$ imposes that Reynolds stresses are zero. For growing y the relative importance of the viscous term lessens in favour of turbulent fluctuations. Near the wall viscosity (and, subsequently, Reynolds number) plays a major role in determining the characteristic of the flow. The other important parameter is wall shear stress. Starting from these quantities (and adding the density ρ) unique velocity and length scales can be defined, namely the *friction velocity*

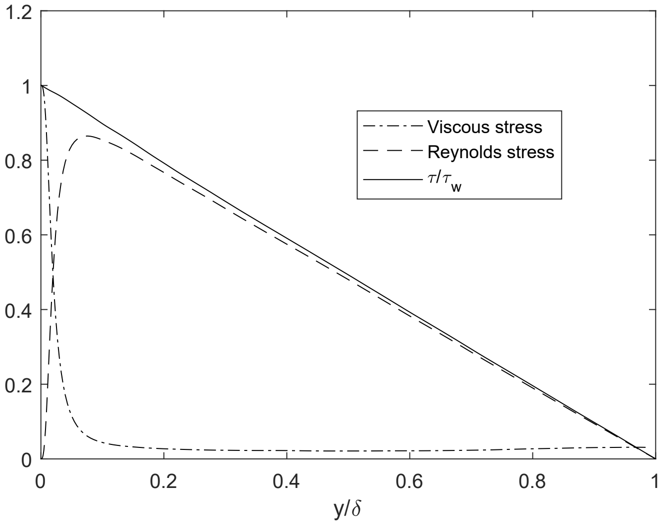
$$u_\tau = \sqrt{\frac{\tau_w}{\rho}} \tag{2.10}$$

and the *viscous lengthscale*

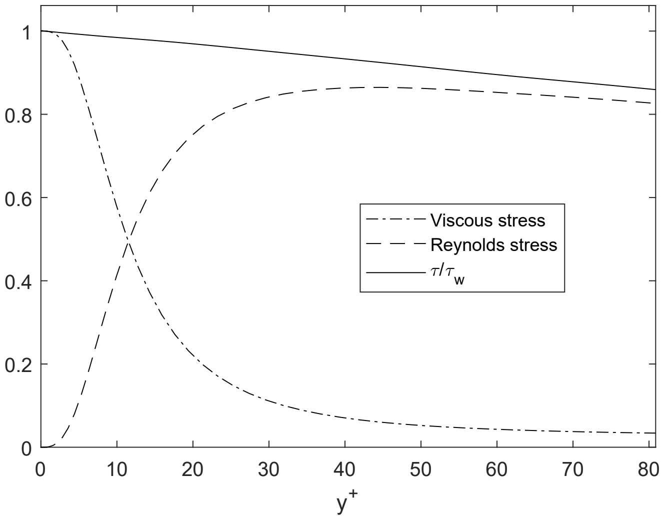
$$\delta_v = \frac{\nu}{u_\tau}. \tag{2.11}$$

Flow quantities normalized using these reference values are usually denoted with the apex $^+$. The friction Reynolds number is also defined as

$$Re_\tau = \frac{u_\tau \delta}{\nu}. \tag{2.12}$$



(a) Full channel



(b) Inner layer close-up

Figure 2.3: Viscous and Reynolds stresses, channel flow at $Re_\tau = 590$. DNS data from Moser, Kim, and Mansour [14]

As stated by Prandtl, at high Reynolds number and in close proximity of the wall, flow quantities normalized in wall units are independent of the external field and determined only by the viscous coordinate $y^+ = y/\delta_v$, which by its own definition is similar to a local Reynolds number, therefore indicating the prevalence of viscous or turbulent effects. Two distinct regions can therefore be identified in the inner layer of the channel flow, namely the *viscous sublayer*, where holds the relation

$$U^+ = y^+, \quad y^+ < 5 \tag{2.13}$$

and the *logarithmic region*, where viscosity can be neglected and the velocity profile obeys the following law

$$U^+ = \frac{1}{k} \log y^+ + C, \quad y^+ > 30, \tag{2.14}$$

where k and C are universal constants. In between those two regions a *buffer layer* is present, where neither viscosity nor turbulence is prevailing. Experimental and numerical data shows a good agreement with these theoretical predictions, as shown for example in figure 2.4, where the velocity profile from the channel DNS employed in this thesis is displayed.

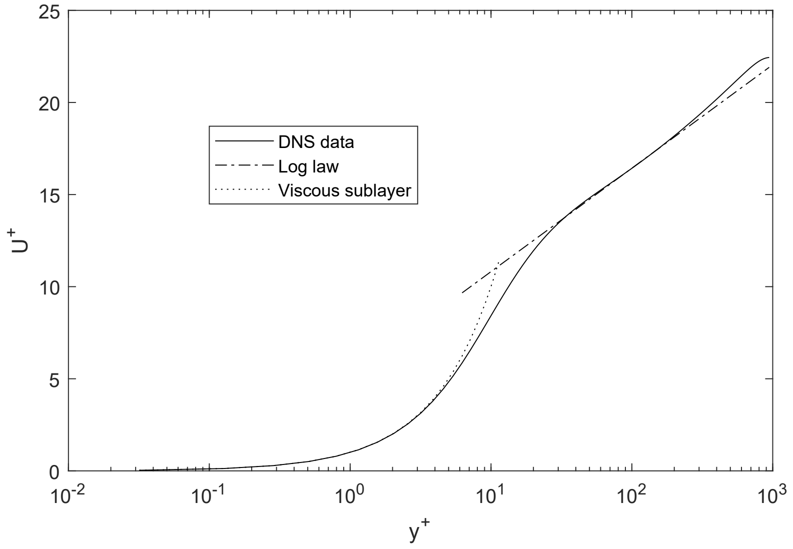


Figure 2.4: Velocity profile in a channel flow at $Re_\tau = 950$. DNS data from Kuerten and Brouwers [13]

Figure 2.5 on the following page shows Reynolds stresses $\langle u_i u_j \rangle$ for a channel flow DNS; these are zero at the wall and grow rapidly away from it. Production of turbulent kinetic energy has a peak located in the buffer layer, while viscous phenomena are the highest at the boundary and rapidly decay away from it.

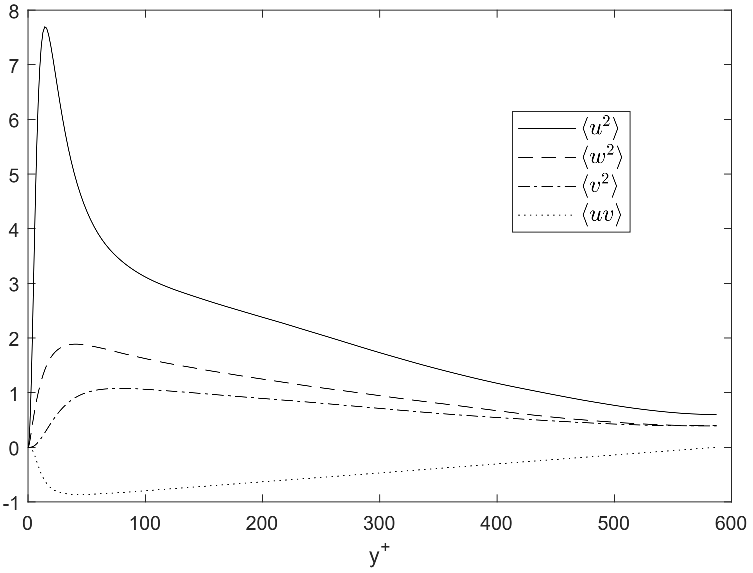


Figure 2.5: Reynolds stresses in a channel flow at $Re_\tau = 590$. DNS data from Moser, Kim, and Mansour [14]

2.2 Lagrangian approach and mixing

2.2.1 Trajectories

A natural approach for the treatment of diffusion and transport phenomena is the Lagrangian viewpoint, that relies on following the motion of infinitesimal particles that do not alter the state of the flow [29]. Having no inertia, these particles follow the underlying Eulerian filed, with their velocities matching the Eulerian one at any point in time and space. The instantaneous position is $\mathbf{x}\left(\mathbf{x}_0, t\right)$ and is a function of time and of the release coordinate $\mathbf{x}_0$ of the tracer. The evolution of the trajectory is simply given by

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{v}_L\left(\mathbf{x}_0, t\right), \tag{2.15}$$

where $\mathbf{v}_L\left(\mathbf{x}_0, t\right)$ is again the Lagrangian velocity of a particle released in $\mathbf{x}_0$ at time t. This is a purely Lagrangian relationship; in order for the trajectory to be coupled to the Eulerian velocity field $\mathbf{v}_E$, a relation must be provided between the two velocities. Since no inertial effects are involved, this is

$$\mathbf{v}_L(\mathbf{x}_0, t) = \mathbf{v}_E\left(\mathbf{x}\left(\mathbf{x}_0, t\right), t\right). \tag{2.16}$$

Both particle tracking experiments and direct numerical simulation (DNS) may be employed to obtain Lagrangian data. In particular in the case of numerical simulation, both an Eulerian computation of the velocity field on a grid and an interpolation scheme to obtain Lagrangian data everywhere else in the domain are needed. The latter is the method that has been used to obtain the trajectories studied in this thesis and will be discussed in deeper detail in chapter 4.

Figure 2.6 shows the projection on the xy and xz plane of the trajectory of a single particle in a turbulent channel flow. As can be seen the particle path is irregular, especially when it is located near the walls where it appears to be trapped in vortical, highly irregular structures.

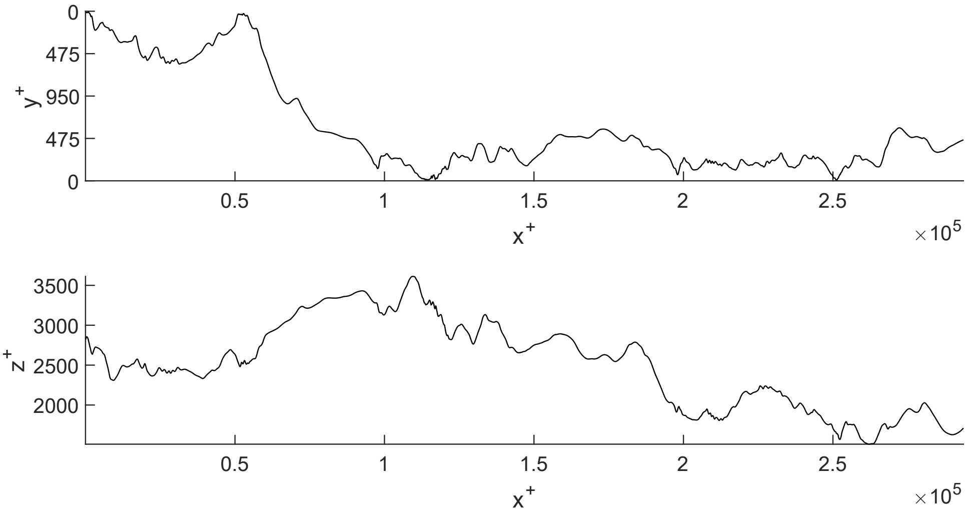


Figure 2.6: Trajectory of a single tracer in a turbulent channel. DNS data from Kuerten and Brouwers [13]

2.2.2 Acceleration and intermittency

Figure 2.7 shows again the trajectory of a single tracer with the modulus of its acceleration, calculated for each time instant with a second order central finite difference. The previously mentioned trapping of particles near the boundary can be clearly seen. The onset of strong vortical motion is associated with high instantaneous accelerations, with a magnitude even greater than 1000 $\mathrm{m/s^{-2}}$.

Intense turbulent fluctuations are present in Lagrangian time histories in a discontinuous manner, both in time and space; this behaviour is commonly addressed as *intermittency*. Intermittency is an inhomogeneous phenomenon and thus experimental and numerical evidence provides results that deviates somewhat strongly from Kolmogorov's prediction.

Lagrangian velocity structure functions of order p, defined as the p-th order statistical moment of velocity increments

$$D_p(\tau) = \langle |\mathbf{v}\left(\mathbf{x}_0, t+\tau\right) - \mathbf{v}\left(\mathbf{x}_0, t\right)|^p \rangle, \tag{2.17}$$

are predicted, as per Kolmogorov's theory, to scale as $(\varepsilon\tau)^{p/2}$. Instead experimental and numerical evidence shows that, in general, the scaling law is

$$D_p(\tau) \propto (\varepsilon\tau)^{\xi_p}$$

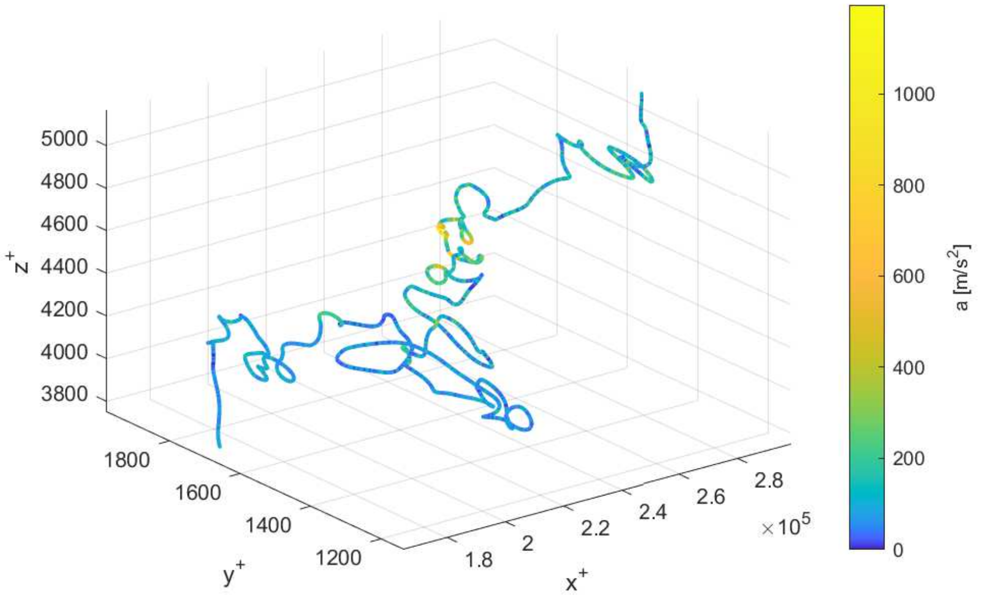


Figure 2.7: Trajectory of a single tracer with its acceleration superimposed. DNS data from Kuerten and Brouwers [13]

with $\xi_p \leqslant p/2$. In addition to that, probability distributions of velocity differences are increasingly non Gaussian as the time lag τ diminishes, due to intermittency; this is closely related to the intermittency of acceleration; indeed, acceleration is proportional to the velocity structure function in the limit of $\tau \to 0$.

2.2.3 Particle dispersion

Taylor [25] has shown that the dispersion of tracers in homogeneous turbulence is closely related to the Lagrangian integral timescale. This relation can be extended to inhomogeneous flows as follows [22]; the integration of (2.15) yields

$$\mathbf{x}\left(\mathbf{x}_0, t\right) = \int_{t_0}^{t} \mathbf{v}(\mathbf{x}_0, r)\mathrm{d}r + \mathbf{x}_0.$$

Being $\mathbf{x}$ a coordinate in cartesian space with x_i $i = 1, 2, 3$ components, the covariance matrix for a set of trajectories is

$$\langle x_i(\mathbf{x}_0, t) x_j(\mathbf{x}_0, t)\rangle = \int_{t_0}^{t}\int_{t_0}^{t} \langle v_i(\mathbf{x}_0, r) v_j(\mathbf{x}_0, s)\rangle \mathrm{d}r\mathrm{d}s, \tag{2.18}$$

where the mean is to be taken on a set of particles starting from the same $\mathbf{x}_0$; in the case of channel flow, the set is composed of tracers belonging to the same starting level y_0, since the x and z direction are statistically homogeneous. The normalized correlation coefficient can be introduced as

$$\rho_{ij}(\tau) = \frac{\langle v_i(\mathbf{x}_0, t_0) v_j(\mathbf{x}_0, t_0 + \tau)\rangle}{\sqrt{\langle v_i^2(\mathbf{x}_0, t_0)\rangle \langle v_j^2(\mathbf{x}_0, t_0 + \tau)\rangle}}, \tag{2.19}$$

whose integral

$$T_{ij}^L = \int_0^\infty \rho_{ij}(\tau)\mathrm{d}\tau \tag{2.20}$$

is the Lagrangian integral timescale, a measure of the time during which the velocities of a tracer are correlated. The normalization factor in (2.19) is the product of the standard deviations of velocity at times t_0 and $t_0 + \tau$ and will be referred to as $\sigma_{ij}(\tau)$. Equation (2.18) can then be rewritten as

$$\langle x_i(\mathbf{x}_0, t) x_j(\mathbf{x}_0, t)\rangle = \int_{t_0}^{t} \int_{t_0}^{r} \sigma_{ij}(s)\rho_{ij}(s)\mathrm{d}s\mathrm{d}r$$

which can be further manipulated to obtain

$$\langle x_i(\mathbf{x}_0, t) x_j(\mathbf{x}_0, t)\rangle = 2(t - t_0) \int_{t_0}^{t} \sigma_{ij}(s)\rho_{ij}(s)\mathrm{d}s - 2\int_{t_0}^{t} s\sigma_{ij}(s)\rho_{ij}(s)\mathrm{d}s. \tag{2.21}$$

For long time lags $\tau = t - t_0$ the velocity is uncorrelated ($\rho_{ij} \approx 0$) and (2.21) reduces to

$$\langle x_i(\mathbf{x}_0, t) x_j(\mathbf{x}_0, t)\rangle = 2(t - t_0) \int_{t_0}^{t} \sigma_{ij}(s)\rho_{ij}(s)\mathrm{d}s, \tag{2.22}$$

being the integral a constant term for large t. Differently from the classical procedure of Taylor for homogeneous flows, the normalization factor $\sigma_{ij}(\tau)$ may not be brought out of the integral since it is connected to the inhomogeneity of the flow.

Chapter 3

Turbulence and networks

3.1 Network theory

The definition of networks encompasses a huge amount of objects and their respective connections, from transportation infrastructures to social acquaintances. All of these very different subjects receive a formal treatment via graph theory. A graph G consists of two sets $\mathcal{N}$ and $\mathcal{L}$, one containing the interacting elements, or *nodes*, the other containing their respective connection, or *links*.

A single link, or edge, $l_{ij} \in \mathcal{L}$ is denoted by the two nodes, or vertices, v_i, $v_j \in \mathcal{N}$ which it joins. Graph theory does not allow self loops, so $l_{ii} = 0 \ \forall i$. If a connection has a directional information embedded in it (as, for example, in a predator-prey relationship) the network is said to be *directed* and $l_{ij} \neq l_{ji}$. Instead, if $l_{ij} = l_{ji}$ the network is *undirected*. In addition to that, a network is weighted if a real number $w_{ij} \in \mathbb{R}$ is associated to each link l_{ij}; weights may represent a variety of characteristic in different network, such as Euclidean distances or the number of cars in each road between cities. The number of nodes is the *order* of a graph, while the number of links is its *size*.

The definition for the links l_{ij} leads naturally to that of the *adjacency matrix*, that is the $N \times N$ matrix $\mathbf{A}$ which has $a_{ij} = 1$ if a link is present between nodes i and j and is zero otherwise. For an undirected network, $\mathbf{A}$ is symmetrical. Since self loops are excluded, the maximum number of links that can be formed in an undirected network is $N(N-1)/2$, while it is $N(N-1)$ in a directed network. The weight matrix $\mathbf{W}$ may be associated to the adjacency matrix.

3.1.1 Degree and strength

The degree k_i of the i-th node is the number of links incident to that node. As such, it is calculated as

$$k_i = \sum_{j \in \mathcal{N}} a_{ij}; \tag{3.1}$$

its most immediate characterization is the *degree distribution* $P(k)$, representing the probability of a given node has degree k or, alternatively, the fraction of nodes having degree k.

An additional distinction can be made for directed networks; it may indeed be useful to part ingoing and outgoing links, thus defining the ingoing degree k^{IN}

$$k_i^{\mathrm{IN}} = \sum_{j\in\mathcal{N}} a_{ji} \tag{3.2}$$

and the outgoing degree k^{OUT}

$$k_i^{\mathrm{OUT}} = \sum_{j\in\mathcal{N}} a_{ij}, \tag{3.3}$$

which correspond to column-wise and row-wise summation of the adjacency matrix respectively.

The degree distribution of a great number of real networks is not homogeneous, as earlier works expected, but instead follow a power law of the type

$$P(k) = Ck^{-\gamma}. \tag{3.4}$$

Such networks are called *scale free*, since relations like (3.4) retain the same functional form at all scales; similar behaviour are present in several natural phenomena, most notably fractals and the inertial range energy spectrum in turbulence (2.4). Scale free networks tend to have a small number of high degree nodes and a high number of nodes with few connections.

Weighted networks characterized by $\mathbf{A}$ and $\mathbf{W}$ do not only possess a degree distribution but also a node strength distribution. For each node i, its strength s_i is defined as

$$s_i = \sum_{j\in\mathcal{N}} w_{ij}. \tag{3.5}$$

For directed networks, the ingoing and outgoing strengths are defined as in (3.2) and (3.3).

3.1.2 Other matrices and graph spectrum

Other than via the adjacency matrix $\mathbf{A}$ the description of a directed graph can be done through the *directed incidence* matrix $\mathbf{B} \in \mathbb{R}^{n\times e}$, defined as

$$B_{ij} = \begin{cases} 1 \text{ if } v_i \text{ is the starting vertex of } l_j \\ -1 \text{ if } v_i \text{ is the ending vertex of } l_j \\ 0 \text{ otherwise} \end{cases} \tag{3.6}$$

From the incidence matrix, the *combinatorial Laplacian* matrix $\mathbf{\Lambda}$ can be derived[2]:

$$\mathbf{\Lambda} = \mathbf{B}\mathbf{B}^{\intercal}; \tag{3.7}$$

which is by its own definition symmetric and semi-definite positive. It may be shown that this matrix has on its diagonal the total number of edges incident on each vertex (*i.e.* the sum of the ingoing and outgoing degree) and that $\Lambda_{ij} = -1$ if there is a link between v_i and v_j ($\Lambda_{ij} = -2$ if there is a directed link in both directions). For undirected graphs the following relation holds:

$$\mathbf{\Lambda} = \mathbf{D} - \mathbf{A}, \tag{3.8}$$

where $\mathbf{D}$ is the diagonal matrix containing the degrees of each node ($D_{ii} = k_i$).

The spectral analysis of $\mathbf{\Lambda}$ is of particular importance for a graph. Let $\lambda_1 < \ldots < \lambda_i < \ldots < \lambda_n$ be the eigenvalues of $\mathbf{\Lambda}$; they are all real by the definition of $\mathbf{\Lambda}$. If a graph is connected (a path between any two vertices is always present) then it will have only one Laplacian eigenvalue equal to 0, that is λ_1. Instead, if the graph is disconnected in m subgraphs, the null eigenvalue will have multiplicity m. In the following the focus will be on connected graphs.

The first non trivial eigenvalue λ_2 has a greater importance and is strongly related to the possibility of partitioning a graph in two weakly connected subgraphs; the greater λ_2, the more difficult it will be to partition the graph. These facts will be addressed using as example the simple graph G depicted in figure 3.1, composed of two highly connected communities which interact between themselves only via a single link l_{16}.

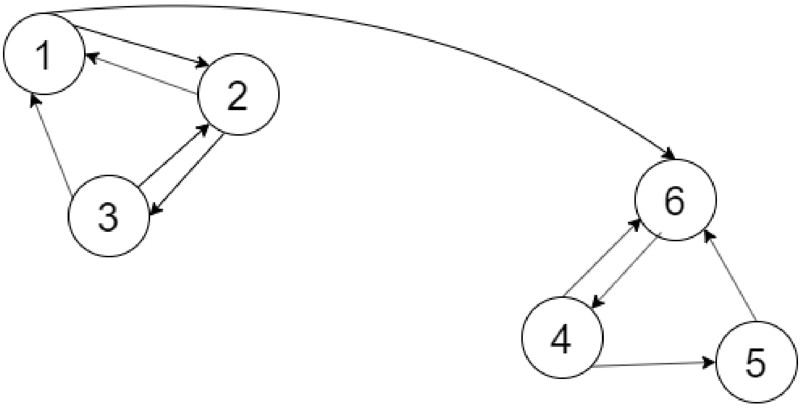


Figure 3.1: Example graph G

While the first eigenvalue of the Laplacian $\mathbf{\Lambda}(G)$ is as expected zero, the second is $\lambda_2 = 0.5$, which is quite small compared to the others and confirms the quite clear partition present in G. Further insight may come from the analysis of eigenvectors, which are able to properly separate nodes into communities [5]. The eigenvector related to the zero eigenvalue $\mathbf{v}_1$ is constant; since eigenvectors are orthogonal and the first is constant, the others must have a zero sum. For the example graph present earlier, the first non trivial eigenvector is

$$\mathbf{v}_2 = \begin{pmatrix} -0.298 \\ -0.437 \\ -0.468 \\ 0.428 \\ 0.480 \\ 0.296 \end{pmatrix}$$

As can be seen, $\mathbf{v}_2$ components associated to the first three nodes are negative, while the other three are positive. If the separation is less sharp, this distinction becomes milder, but taking into account successive eigenvectors may provide further definition. Figure 3.2 shows the components of the first eigenvector; red and blue marks correspond to the first and second community of G. The separation in eigenvector space is visible and quantifiable, although adding more dimensions to it may present the need to improve the measure of separation between components.

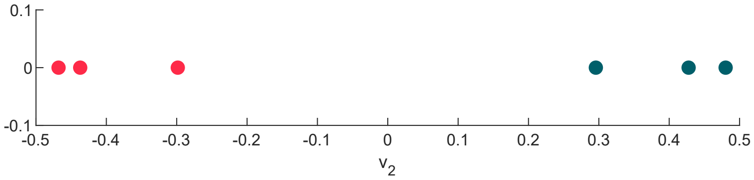


Figure 3.2: First non trivial eigenvector of G

3.2 Applications to turbulent flows

The following section will briefly examine previous results obtained by the research group in which this thesis was conceived, regarding the application of network theory to fluid dynamics and wall turbulence. Earlier works have focused on climate dynamics, exploiting the features of complex networks to highlight interconnections between different regions of the globe [18]. Recent works have instead shifted their attention to the study of high resolution turbulent flows, exploiting the ability of networks to describe with simple metrics phenomena that could otherwise be investigated only by statistical means.

The main issue in this kind of approach is providing a proper definition of the $\mathcal{N}$ and $\mathcal{L}$ sets, so that the usual network metrics provide useful insight on the evolution of the turbulent flow.

A first approach to transform a time series into a network is the *visibility graph*; different time instants are the nodes, while a link is established if two nodes are in mutual visibility, *i.e.* it is possible to draw a straight line between two data points without intersecting other data. So a link is present between times t_i and t_j for the time series $s(t)$ if

$$s(t_k) < s(t_j) + (s(t_i) - s(t_j))\,\frac{t_j - t_k}{t_j - t_i} \quad \forall k = i \dots j. \tag{3.9}$$

Peaks in the time series $s(t)$ usually are connected to a greater number of nodes, while time instants in their surroundings may have their visibility obstructed by the peak and thus have a reduced degree. Also irregularities in the time series are highlighted by similar behaviour of the visibility graph. This approach has been exploited to describe time series obtained from single points in space in a channel flow, considered from the Eulerian viewpoint [7].

Also in the Eulerian framework, the definition of a link between single points in space could stem from physical features of those points, such as kinetic energy correlation (employed in homogeneous isotropic turbulence [17]) and velocity correlations (for wall bounded turbulence [9]). For all of these methods, also limits on link distance were imposed to exclude non physical correlation from forming network links; the Eulerian framework naturally provides characteristic lengths (such as the Taylor microscale λ for the homogeneous turbulence case) to limit the spatial length of interactions.

The application to Lagrangian data allows to use single tracers or groups of them as network nodes. A definition of links based on spatial proximity [8] allows to measure the evolution of interactions between tracers through time and to analyse diffusion and dispersion. The development of additional methods based on Lagrangian data is the subject of this thesis.

Chapter 4

Channel flow dataset

In order to obtain Lagrangian data and statistics for a turbulent flow, a robust and effective method to generate trajectories is needed. The numerical resolution of (2.15) implies the knowledge of the Eulerian velocity field everywhere in the fluid domain, which has been obtained with direct numerical simulation at a friction Reynolds number Re_τ equal to 950. Other than presenting the DNS tools, this chapter will provide a brief description of the tracers' trajectories dataset, analysing distance, velocity and acceleration statistics.

4.1 Channel flow DNS

The ensemble of trajectories has been obtained through the work of Kuerten and Brouwers [13], which comprises both a channel flow DNS and an interpolation and integration scheme. The Navier-Stokes equations (2.1) - (2.2) are solved in their *rotation* form

$$\nabla \cdot \mathbf{v} = 0 \tag{4.1}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla P = \mathbf{f} - \boldsymbol{\omega} \times \mathbf{v} + \nu \Delta \mathbf{v} \tag{4.2}$$

where the vector identity

$$(\mathbf{v} \cdot \nabla)\, \mathbf{v} = \nabla \left(|\mathbf{v}|^2 \right) /2 - \boldsymbol{\omega} \times \mathbf{v}$$

has been employed; $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ is the vorticity and $P = p + \rho|\mathbf{v}|^2/2$ is the total pressure. The Navier-Stokes equations are made non dimensional using the friction velocity u_τ, the channel half-height δ and density ρ. Re_τ (2.12) is kept fixed to 950 by imposing an x-directed driving force per unit mass $\mathbf{f}$; the bulk Reynolds number based on the centreline velocity is thus not fixed.

Using a procedure pioneered by Kim, Moin, and Moser [10], equations (4.1) - (4.2) are solved in a box volume with height equal to 2δ, length $2\pi\delta$ in the streamwise x direction and $\pi\delta$ in the spanwise z one. A periodic boundary condition is imposed on the x and z directions, while the no slip condition $\mathbf{v} = 0$ is imposed on the channel walls. Incompressibility is satisfied by means of a convenient choice of dependent variables.

A pseudo-spectral method is used, with a Fourier-Galerkin approach in the homogeneous directions and a Chebyshev-Tau method in the wall normal one. Non linear terms are calculated in physical space and are explicitly integrated in time with the second order Runge-Kutta method, while linear terms are advanced in time through an implicit Crank-Nicolson scheme.

4.2 Trajectories integration

The DNS is run until all Eulerian quantities achieve statistical convergence; after that, particle are seeded in the simulated channel and their paths integrated. $N_p = 10^4$ particles are inserted at $x = 0$, organized in $N_l = 100$ equally spaced levels in the wall normal direction; the first level is located at $y^+ = 9.5$ and their spacing is $\Delta y^+ = 19$. Tracers are also homogeneously distributed along z with $\Delta z^+ = 29.85$. Each level therefore contains $N_{p1} = 100$ particles.

At each time step the velocity of the tracer is found applying trilinear interpolation to the computational grid. Although other methods may be more accurate in computing single trajectories, statistical accuracy on the entire set of particles was not affected [13]; still, spurious oscillations in velocity and acceleration of particles may derive, modifying probability distributions [4] [27]. For improved precision the velocity field was employed after the transformation in real space, where the number of grid points is increased by a factor of 3/2 for anti-aliasing.

Advancement in time was performed using the same explicit second order Runge-Kutta method employed in the DNS. The trajectories were calculated for a total time equal to $T^+ = 15\,200$, which is a time greater than that needed for vertical mixing and for the onset of the asymptotic Taylor dispersion regime; the time step of the integration is $\Delta t^+ = 0.475$, this resulting in $N_t = 32\,000$ time instants.

The dataset was provided in the form of a single precision floating point Matlab matrix of size $N_p \times 3 \times N_t$, where along the second dimensions the three spatial coordinates of each particles are stored. Since the size of this matrix is considerable (> 3 GB), some precautions had to be taken during post processing in order to perform a fast and memory efficient computation. On the other hand, being the data stored in single precision, some numerical accuracy may be lost, especially when performing certain operations such as derivatives; particular care has been employed when computing velocities and accelerations[†].

4.3 Distance statistics

Starting from the trajectories dataset, a first natural approach is to examine the evolution of the mutual distance of tracers. Owing to the dispersive action of turbulent flows, an overall diffusive action is to be expected; still, different phases, each with its peculiar behaviour, can be anticipated (as has been displayed in section 2.2.3 on page 12) and the analysis of higher order statistical moments may provide useful insight on the behaviour of particles.

[†] Most notably in the calculation of derivatives, spatial coordinates were divided by the time step Δt^+ *before* subtracting them, in order to mitigate the effects of numerical cancellation

A matrix $\mathbf{D}$ containing mutual distances of all particles can thus be defined as

$$D_{ij} = |\mathbf{x}_i^+ - \mathbf{x}_j^+|, \tag{4.3}$$

where $\mathbf{x}_i^+$ is the position of the i-th particle in viscous coordinates. $\mathbf{D}$ is varying with time, following the motion of tracers. Since the channel has an inhomogeneous direction y, it could reasonably be expected that particles behave differently depending on their wall normal release coordinate, *i.e.* y_0^+. It is then useful to consider a reduced matrix $\mathbf{D}^l$, which retains only the distance of particles belonging to level l with the rest of the dataset. The matrix has size $N_{p1} \times N_p$ ($100 \times 10\,000$) and has the form

$$\begin{array}{cc} & \text{All particles} \\ \text{Single level} & \begin{pmatrix} D_{11} & D_{12} & \dots & D_{1,N_p} \\ \cdots & \cdots & \cdots & \cdots \\ D_{N_{p1},1} & D_{N_{p1},2} & \dots & D_{N_{p1},N_p} \end{pmatrix} \end{array} \tag{4.4}$$

Figures 4.1 - 4.4 show the evolution of statistical moments of mutual distances between particles, for both the whole set and selected levels, denoted by their respective y_0^+. In figure 4.1 an asymptotic regime can be noted, where the mean distance increases linearly with time; this is somewhat similar to (2.22), although the formulations are not equivalent.

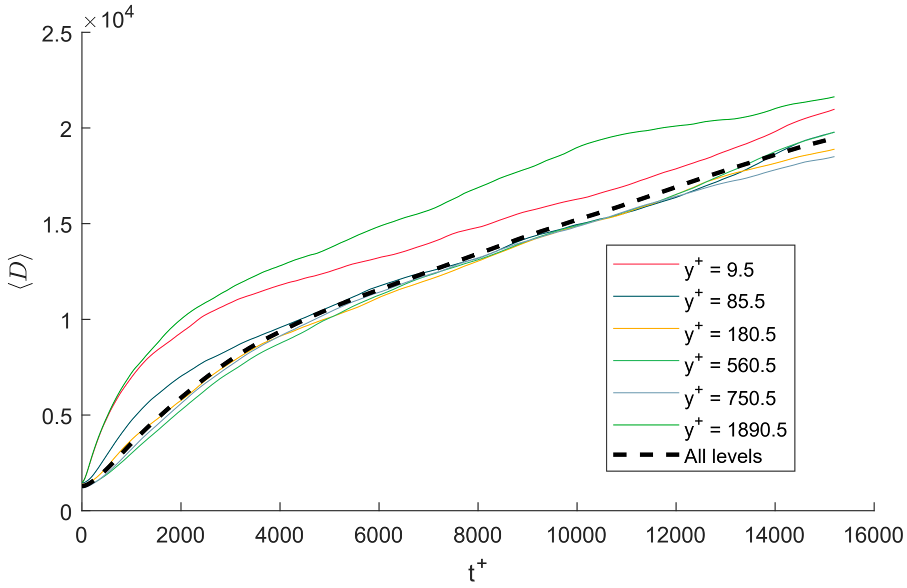


Figure 4.1: Mean distance of particles

The initial transient lasts until $t^+ \approx 4000$, a separation that stands consistently through many analysis in the following of this work and thus can be considered a transition between different regimes, namely the Taylor dispersion regime and the previous, non asymptotic, one. The mean distance of levels near the walls is higher at $t^+ = 0$ (because of how particles are arranged at their release) and also grows faster then the rest of the dataset; this could be because tracers embedded in the inner part of the boundary layer move at a slower velocity and are thus rapidly parted from the rest of the set. On the long run mixing is effective and all curves tend to collapse on the mean for all particles, although the process appear to be very slow for tracers whose y_0^+ is very small.

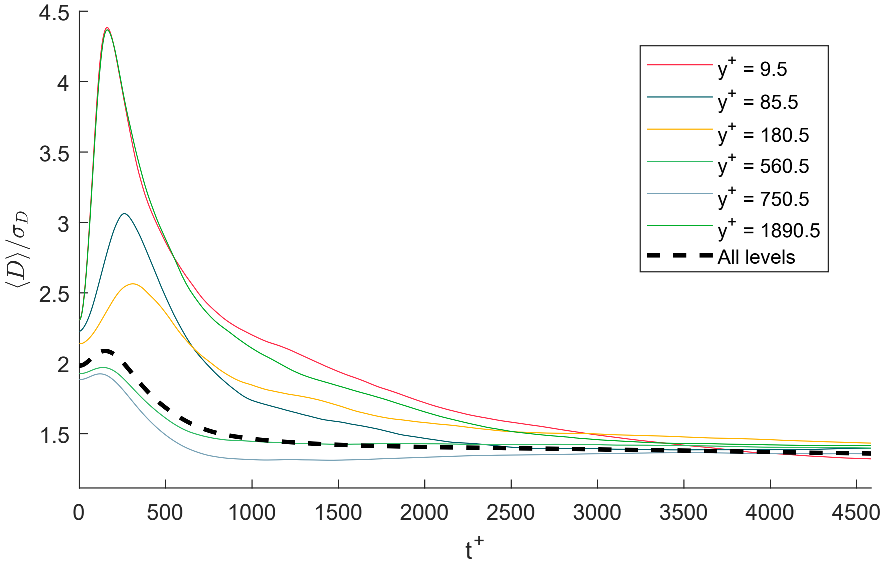


Figure 4.2: Mean distance of particles divided by its standard deviation

Figure 4.2 shows the ratio between mean distance and its standard deviation σ_D. In the asymptotic regime those two quantities appear to be proportional for all levels, while in the starting transient they behave considerably different. In particular lower levels increase their mean distance considerably faster than their standard deviation, this because mean distance increases and standard deviation initially decreases; both these phenomena are probably due to particles being initially trapped in the region closest to the wall, as already stated.

Figure 4.3 shows the skewness, which is proportional to the third statistical moment of the distribution of mutual distances. It samples the asymmetry of the distribution, so that a positive skewness indicates that values smaller than the mean are more frequent. All levels reach a nearly asymptotic positive value by $t^+ \approx 4000$, although being quite irregular, showing that low distances between particles are more common than high ones. Only levels near the wall for a brief time exhibit a negative skewness, nearly at the same time for which the ratio $\langle D \rangle / \sigma_D$ is maximum.

Finally in figure 4.4 the kurtosis of the distance distribution is shows. Kurtosis is a measure of the relative importance of outliers in a distribution; in particular the kurtosis of a normal distribution is 3, with higher values corresponding to *tighter* distributions. The kurtosis for all levels appears to be always higher than three, although no clear trend can be easily identified other than an initial growth. This could be due to the relatively low number of distances on which this moment is calculated.

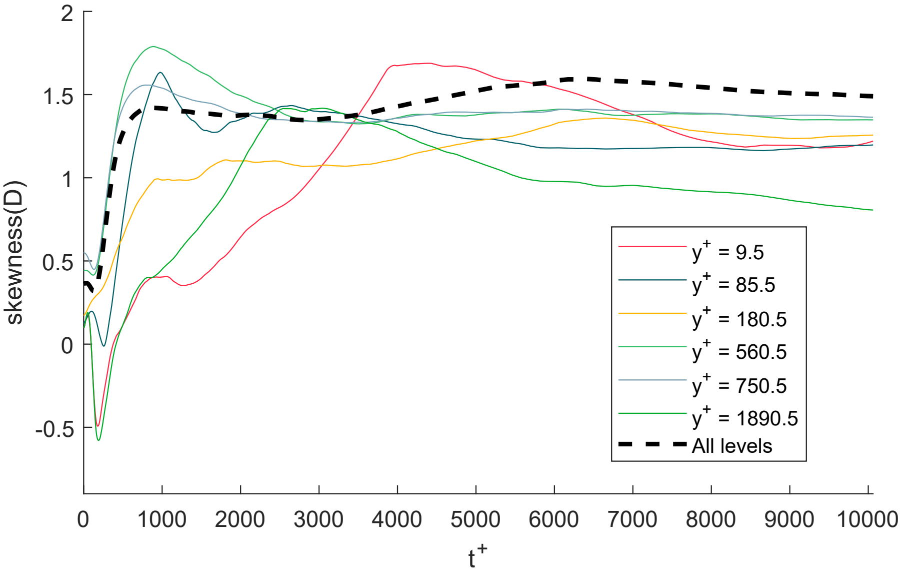


Figure 4.3: Skewness of the distance of particles

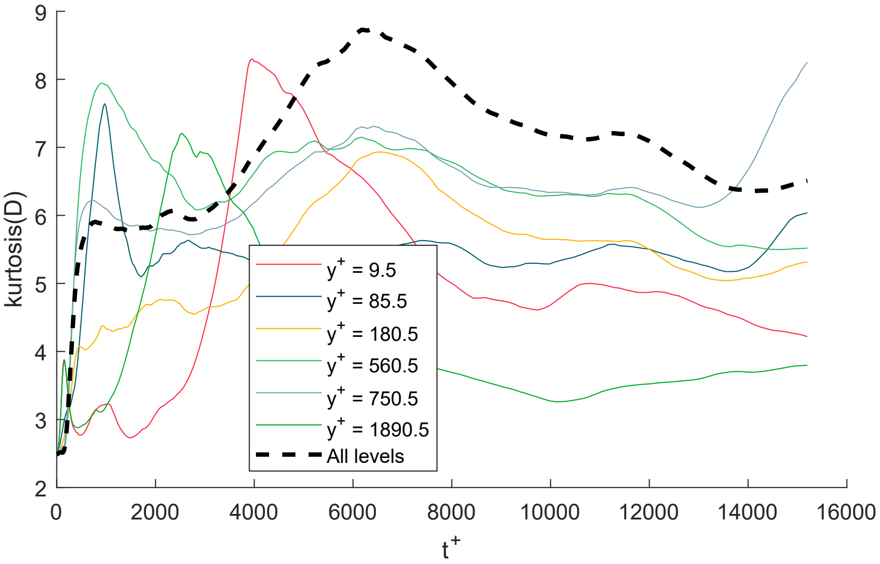


Figure 4.4: Kurtosis of the distance of particles

4.4 Velocity statistics and Lagrangian timescale

Starting from the trajectories, a velocity time series for each particle has been calculated using a second order central finite difference scheme

$$\mathbf{v}^+\left(\mathbf{x}_0^+, t^k\right) = \frac{\mathbf{x}^+\left(\mathbf{x}_0^+, t^{k+1}\right)}{\Delta t^+} - \frac{\mathbf{x}^+\left(\mathbf{x}_0^+, t^{k-1}\right)}{\Delta t^+} \tag{4.5}$$

where t^k is the discretized time and $\Delta t^+ = 0.475$ the constant time step; as already stated, the division has been performed before the subtraction to alleviate the effects of numerical cancellation.

4.4.1 Probability distributions

Figures 4.5 shows probability distributions of the streamwise component of velocity for particles originating from different levels at a time instant in the starting transient. It can be noticed how particles have clearly different velocities depending on their y_0^+, since at low t^+ these are still well inside their release region. Also particles near the walls appear to have a greater variance, possibly meaning that velocity fluctuations are more intense.

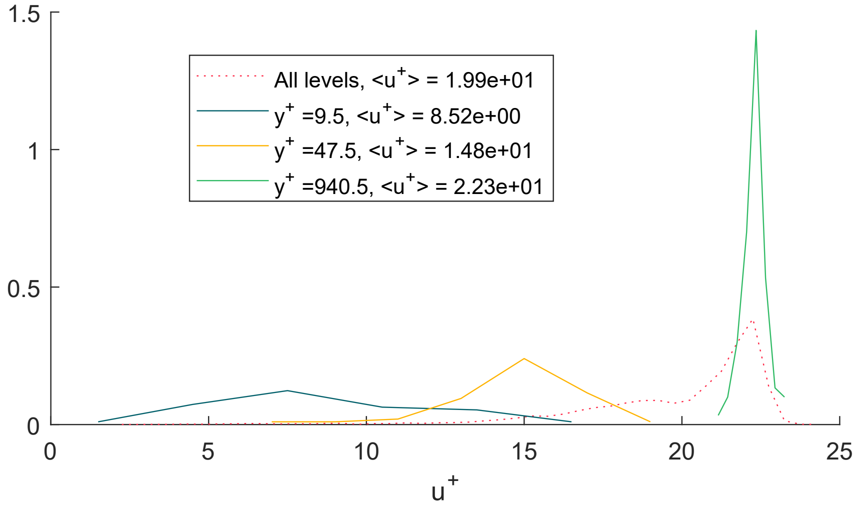


Figure 4.5: Probability distribution of u^+, $t^+ = 100$

Figure 4.6 shows the evolution of the mean streamwise velocity. As already noticed before, a clear shift in flow properties happens at $t^+ \approx 4000$, where the mean velocity of particles with different y_0^+ become equal; this is due to tracers becoming thoroughly diffused in the channel height. After $t^+ \approx 4000$ trajectories have lost all information regarding their origin coordinate and so their statistics are mostly independent of the starting level.

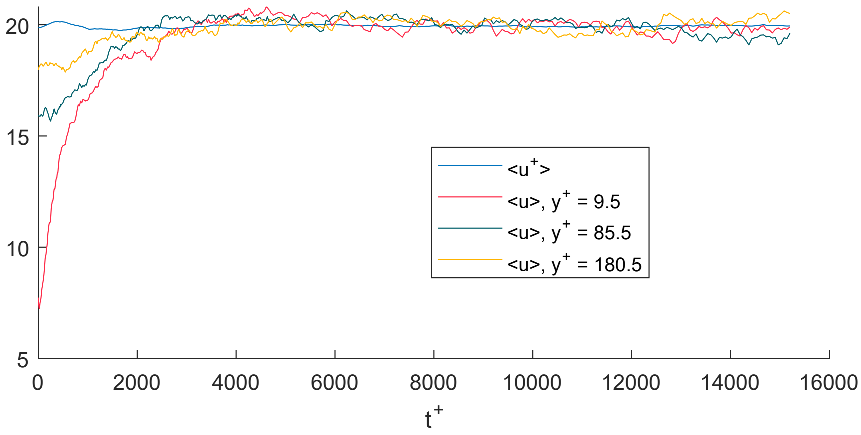


Figure 4.6: Mean axial velocity for different y_0^+

4.4.2 Velocity correlations and timescales

Another measure of the persistence of flow structures are velocity correlations and their integrals, which in the Lagrangian viewpoint are the characteristic timescales of turbulent motions. The timescales are the typical times of large eddies embedded in the channel, calculated from the motion of tracers inside it; they are thus depending on the inhomogeneous release coordinate y_0^+.

Velocity correlations are calculated from perturbed velocity, *i.e.* $\mathbf{v}' = \mathbf{v} - \langle \mathbf{v} \rangle$, where $\langle \mathbf{v} \rangle$ is the mean velocity and has only the streamwise component ($\mathbf{V} = (U,\,0,\,0)$). U is obtained from Eulerian DNS data, averaged for different axial coordinates and times. The normalized cross correlation function, already defined in (2.19) on page 13, is

$$\rho_{ij}(\tau) = \frac{\langle v_i'(\mathbf{x}_0, t_0) v_j'(\mathbf{x}_0, t_0 + \tau) \rangle}{\sqrt{\langle v_i'^2(\mathbf{x}_0, t_0) \rangle \langle v_j'^2(\mathbf{x}_0, t_0 + \tau) \rangle}},$$

and it depends on $\mathbf{x}_0$, t_0 and τ; the mean is taken on the set of particles sharing the same y_0^+. It should be noted that due to the smallness of the dataset (100 tracers for each level) the velocity fluctuations have still a non zero mean, so their mean has been removed and ρ_{ij} is better defined as a *covariance* matrix. For $i = j$, ρ_{ii} is the autocorrelation function. The dependence on t_0 is tightly linked to that on $\mathbf{x}_0$, since changing the starting time the particle will be moved to another starting coordinate and its correlation function will reflect that. The autocorrelation function is bounded between 1 and -1, with the unitary value corresponding to perfectly correlated variables; ρ_{ii} for a turbulent velocity is equal to 1 only for $\tau = 0$.

Figure 4.7 shows the autocorrelation function for the axial component of the velocity u', for two different starting levels. It can be immediately noted that particles originating near the boundary have their velocities correlated for a much shorter time than the others. This signals that structures in the inner part of the boundary layer have a lesser lifespan than those near the centreline.

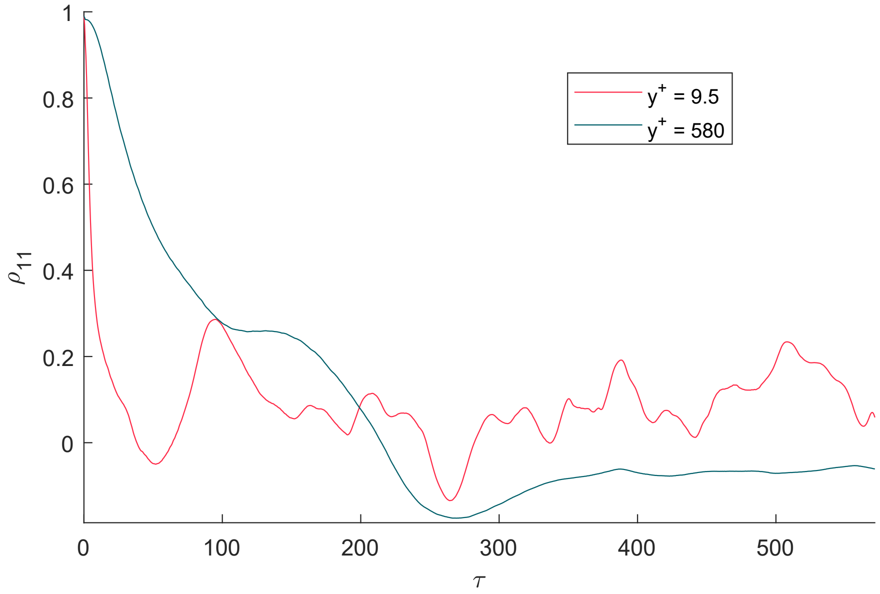


Figure 4.7: Autocorrelation ρ_{11} for two different y_0^+

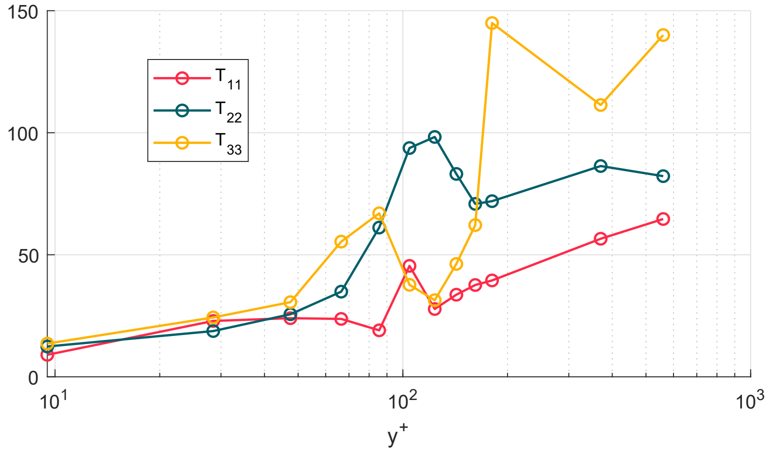
The autocorrelation function is expected to decrease exponentially, still the finiteness of the set makes this true only for short time lags τ; particular care has then to be taken in order to properly integrate the autocorrelation functions. The autocorrelation integral is the timescale T_{ij} as defined in (2.20) on page 13; since integration for all times is not practical nor it yields significant results, two different approaches have been used:

- Integration up to the first zero crossing of ρ_{ii}, neglecting negative autocorrelations and the subsequent oscillations. The results are shown in figure 4.8a.
- Exponential fitting for small time lags τ. The integral timescale is set as the time τ for which $\rho_{ij} = 1/\mathrm{e}$; the fitting curve is then $\mathrm{e}^{-\tau/T_{ii}}$. The results are shown in figure 4.8b.

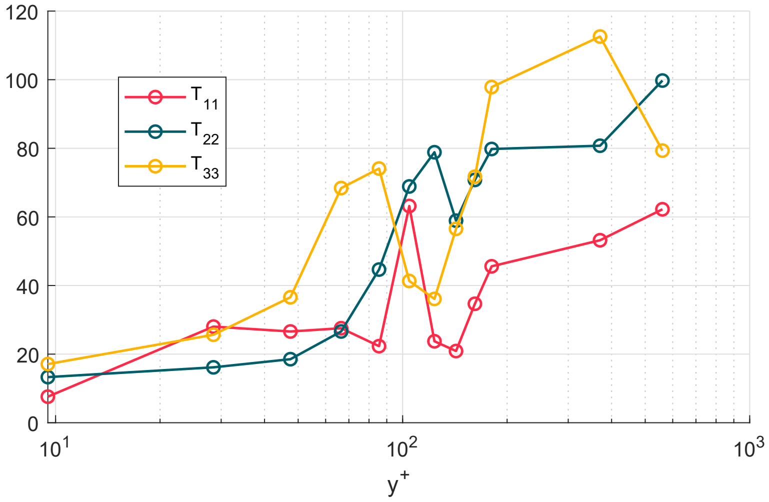
Both methods yields similar trends with small differences in magnitude. Although data from the viscous sublayer is missing because no tracers originate from y_0^+ lower than 9.5, it can clearly be seen that the integral timescale is increasing for all components of velocity moving away from the wall. This is linked to the fact that structures near the boundary have shorter lifespans. The trend is quite irregular inside the logarithmic law region, especially near $y_0^+ = 100$. Axial velocity fluctuations are correlated for a shorter time than the other two components.

4.4.3 Taylor microscale

Starting from the autocorrelation function, also the temporal Taylor microscale λ_i can be calculated. It is equal to the intercept of the osculating parabola to ρ_{ii} at $\tau = 0$, as



(a) Integral method



(b) Fitting method

Figure 4.8: Velocity timescales T_{ii}^{+} in wall time units

portrayed in figure 4.9. It is a scale representative of small but still dynamically significant eddies in the flow belonging to the inertial range. Its value is

$$\lambda_i^2 = -\frac{2}{\left(\frac{\partial^2 \rho_{ii}}{\partial \tau^2}\right)_0}, \tag{4.6}$$

which is equal to

$$\lambda_i^2 = \frac{2\langle u_i'^2 \rangle}{\left(\frac{\partial u_i'}{\partial t}\right)^2}.$$

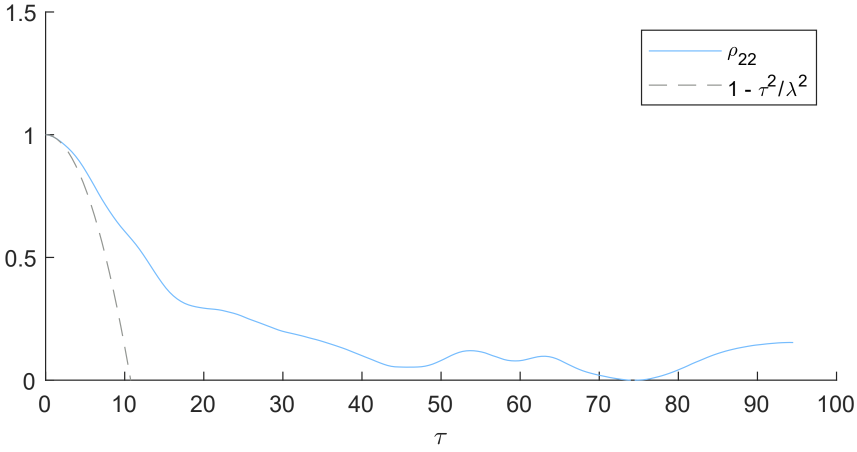


Figure 4.9: Graphical depiction of the temporal Taylor microscale

Values of the Taylor microscale for different y_0^+ are shown in figure 4.10. It should be noted that this scale is much larger than the Kolmogorov characteristic time τ_η (in a similar fashion as the spatial Taylor scale and the Kolmogorov length scale). It can be easily noted that the Taylor microscales λ_2 and λ_3 increase in a similar manner and with values close one to the other, while λ_1 attains a nearly constant value across the height of the channel.

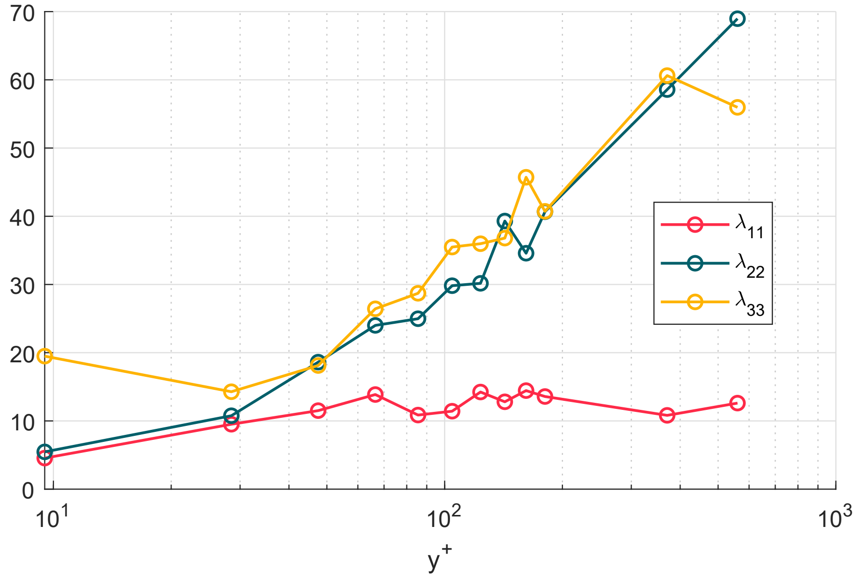


Figure 4.10: Taylor scale λ_i for different particle release coordinates

4.5 Acceleration statistics

In this section statistics on the acceleration of tracers are presented. The acceleration is calculated as the second order central finite difference of the position

$$\mathbf{a}^+\left(\mathbf{x}_0^+, t^k\right) = \frac{\mathbf{x}^+\left(\mathbf{x}_0^+, t^{k+1}\right)}{\left(\Delta t^+\right)^2} - 2\frac{\mathbf{x}^+\left(\mathbf{x}_0^+, t^k\right)}{\left(\Delta t^+\right)^2} + \frac{\mathbf{x}^+\left(\mathbf{x}_0^+, t^{k-1}\right)}{\left(\Delta t^+\right)^2}. \tag{4.7}$$

The normalization coefficient of the acceleration in wall units is u_τ^3/ν. Since typical values for $\|\mathbf{a}^+\|$ are in the order of unity, very high values of physical acceleration can be attained, as already shown in figure 2.7 on page 12. The probability distribution function of a component of the acceleration, normalized by its standard deviation, is shown in figure 4.11; other components show a similar behaviour for different time instants. As can be seen there is no significant deviation between particles originating from different levels; the distribution is non Gaussian and exhibits long tails, which is a sign of the presence of strong bursts in the acceleration and thus intermittency. Still, results may be biased by the inaccuracies in the computation of the acceleration and by the low number of particles in each level.

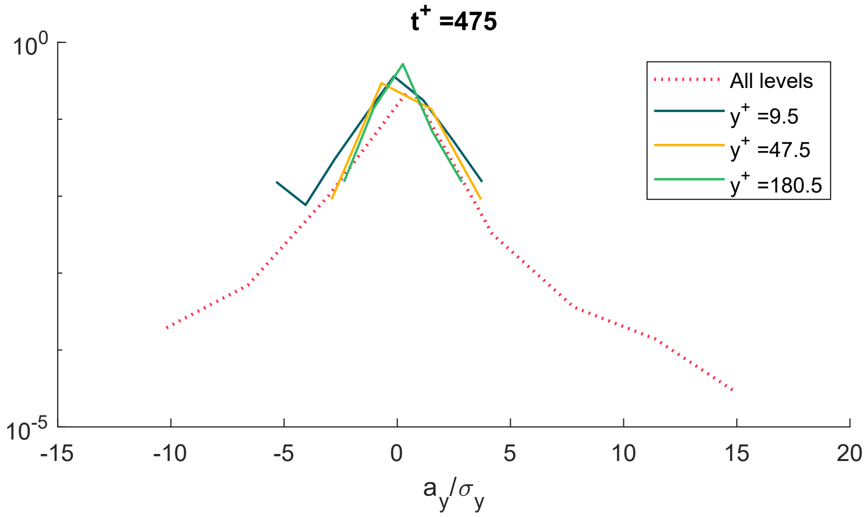


Figure 4.11: Probability distribution function of a_y^+/σ

4.5.1 Acceleration timescale

In a manner similar to subsection 4.4.2, acceleration integral timescales may be calculated. The results are reported in figure 4.12 and appear coherent with those from [23], except for the spike in T_{11} at $y_0^+ \approx 100$. Acceleration components in general stay correlated for a shorter time than velocity ones, as shown in 4.13; their timescale also does not increase much with wall distance.

The timescale has been calculated using the integral method since the quality of fitting was lower due to the small time lags involved.

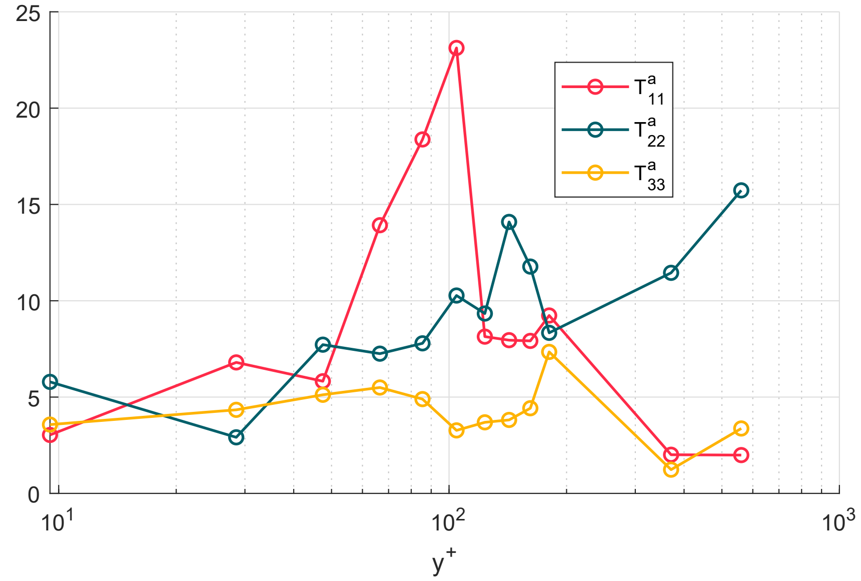


Figure 4.12: Acceleration timescale in wall units of a_x, a_y and a_z for different y_0^+

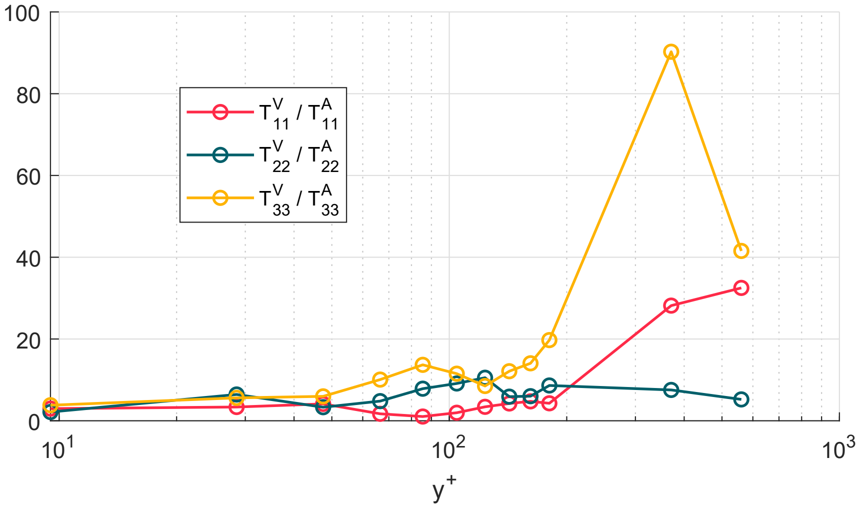


Figure 4.13: Ratio T^v_{ii}/T^a_{ii} between velocity and acceleration timescales

Chapter 5

Clustering network

5.1 Definitions

Tracer particles are driven further and further away one from the other thanks to the diffusive action of turbulence; nonetheless it could be reasonably expected that turbulence has an inverse effect at least locally, regrouping tracers in certain spatial and temporal coordinates while the overall trend is still a diffusive one. For the scope of this chapter, this behaviour will be addressed as *clustering*.

The mutual distances matrix for all particles $\mathbf{D}$, already defined in (4.3) for each particle i, j as

$$D_{ij} = \|\mathbf{x}_i^+ - \mathbf{x}_j^+\|$$

contains in each row the euclidean distance of the particle i (ranging from 1 to the number of particles N_p) with the j-th particle. $\mathbf{D}$ is therefore a square symmetric matrix; only half of it can be calculated in order to speed up the process. Since $\mathbf{D}$ changes with time following the motion of particles, also its shift in time $\mathbf{\Delta D}$ can be defined as

$$\mathbf{\Delta D}^k = \mathbf{D}^{k+1} - \mathbf{D}^k, \tag{5.1}$$

where the superscript denotes the time instant at which the distance matrix is taken. It should be noted that, since the time lag between consecutive times $\Delta t = t^{k+1} - t^k$ is constant, (5.1) is a multiple of the ensemble of particles' mutual velocities, calculated with a forward first order finite difference.

Aiming to define a network of clustering particles, positive elements in $\mathbf{\Delta D}$ are discarded; in addition to that, only particles whose distance at t^k is lower than a certain threshold value d_t are retained. This is done in order to exclude approaching motions that do not arise from the belonging of the particles to the same flow structure but rather to different structures with separate behaviour. More detail on the choice of this threshold will be given in the following. A modified version of $\mathbf{\Delta D}$ can thus be defined as

$$\widetilde{\Delta D}_{ij}^k = \begin{cases} \Delta D_{ij}^k, \text{ if } \Delta D_{ij}^k < 0 \wedge D_{ij}^k < d_t^+ \\ 0, \text{ otherwise} \end{cases} \tag{5.2}$$

for each couple i, j of particles. Figure 5.1 shows non zero elements of $\widetilde{\mathbf{\Delta D}}$ and their magnitude; again, only the upper half of the matrix has been calculated since it is symmetric. Taking this matrix as (half of) a graph weight matrix $\mathbf{W}$, a network can be built. Each entry $\widetilde{\Delta D}_{ij}$ then corresponds to a link between particles i and j. Substituting non zero entries with ones, also the adjacency matrix $\mathbf{A}$ for this network can be built.

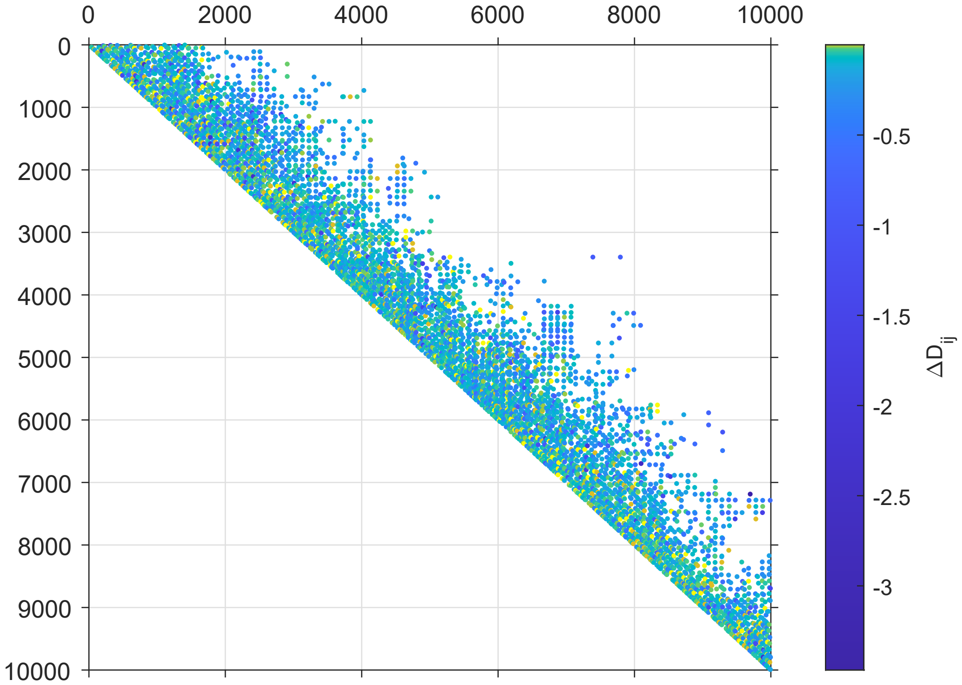


Figure 5.1: Network matrix $\widetilde{\mathbf{\Delta D}}$, $t^+ = 475$, $d_t^+ = 100$

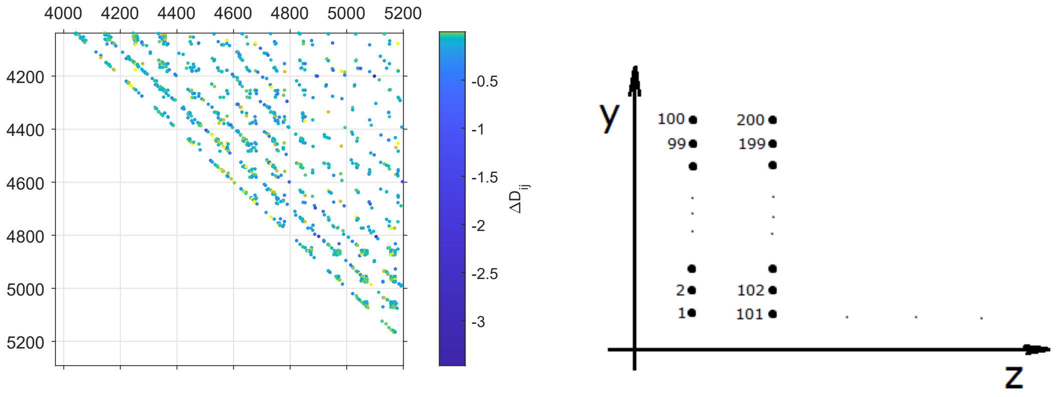
The degree of each node, being the number of links incident to that node, is the sum of each row (or column, independently) of $\mathbf{A}$. It is useful to remark that this definition includes also isolated vertices, *i.e.* nodes that to do not participate in any link. In the following also degree distributions of only nodes with links (so with $k \geqslant 1$) will be given.

It is also useful to define the *centroid coordinate* of each link as the mean position of particles forming a link, that is

$$\mathbf{x}_{c,ij} = \frac{\mathbf{x}_i^k + \mathbf{x}_j^k}{2}, \tag{5.3}$$

where the apex ^{k} signals that positions are calculated before the clustering happens. Other metrics pertinent to links are their weight value $\widetilde{\Delta D}_{ij}$ and the mutual distance r of the two tracers before clustering, simply calculated using the mutual distances matrix $\mathbf{D}$ at time t^k.

Figure 5.2a shows a close-up of the weight matrix; links are present only between neighbouring particles; therefore, non zero entries are present only near the main diagonal and the upper (and lower, due to symmetry) 100th, 200th,…diagonals. This happens because of the disposition of tracers at their release, which is schematised in figure 5.2b. Each tracer i is close, at $t^+ = 0$, not only to $i + 1$ and $i - 1$ (moving in the y direction), but also to $i + 100$ and $i - 100$ (moving in the z direction). The effects of the original disposition of particles persist for some time, but is subsequently lost when tracers are completely mixed.



(a) Network matrix $\widetilde{\mathbf{\Delta D}}$, $t^+ = 475$, $d_t^+ = 100$ (b) Tracers' arrangement at $t^+ = 0$

Figure 5.2: Detail of the structure of $\widetilde{\mathbf{\Delta D}}$

5.2 Single time clustering

Figure 5.3 shows both the swarm of particles (represented as grey dots) and blue links, corresponding to the network shown in figure 5.1; the thickness of these lines is proportional to the magnitude of distance reduction.

At a quick glance, some main features of this network can be identified:

- Links are strongly inhomogeneous in physical space; this mostly reflects the uneven concentration of particles. This matter will be addressed in section 5.5.3.
- Near the centreline of the channel a greater number of links is present than at the walls. The number of links of each node is, in the network formalism, the degree k; figure 5.4 shows the joint probability distribution of k and of the distance from the nearest wall of each particle; also the marginal distributions and a trendline are shown[†]. The marginal *pdf* of k is shown in a logarithmic plot to highlight its exponential behaviour. Such a distribution implies that there are a few of highly

---

[†] The trendline is equivalent to a moving average and has been calculated partitioning the x axis and taking the average of the y values of point belonging to that interval. The correlation coefficient is shown in the caption.

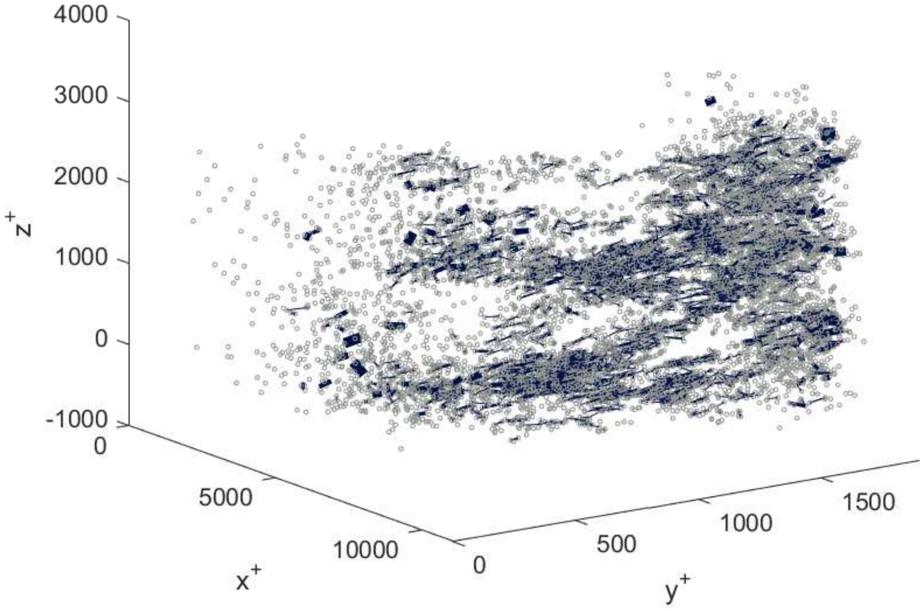


Figure 5.3: Particle snapshot and clustering links at $t^+ = 475$, $d_t^+ = 100$

connected nodes, while the majority of tracers is clustering with a small number of other particles.

- Particles near the wall get closer one to the other with more intensity than particles near the centreline. This is shown in figure 5.5 on page 38, where the joint probability distribution of the entity of distance decrease and the distance of the centroid y_c^+ from the nearest wall is shown. The y_c^+ marginal *pdf* clearly shows that while particles are thoroughly distributed across the channel (as from the y^+ marginal *pdf* of 5.4), links are concentrated near the centreline (at least at this time). Also it can be seen that low intensity links are way more frequent than high intensity ones, which are in turn located near the walls. This highlights the fact that near wall structures present stronger turbulent motions, which in turn leads to zones where tracers reduce their mutual distance.

- A dependency between particle mutual distance and link strength (normalized by mutual distance) is also highlighted in their joint probability distribution (figure 5.6 on page 38). It may be deducted that (non-normalized) link strength grows with mutual distance; still, the mean normalized link strength (shown in the trendline) is slightly higher for small r^+ links. This may lead to the conclusion that small scale interactions are relatively more intense than those on larger scales. This fact remarks the influence of flow structures on the behaviour of clustering links. Since the threshold is way below the channel height, all links considered here belong to physically significant scales. The marginal *pdf* of r^+ shows how longer range links are more frequent overall.

- Figure 5.7 shows the joint *pdf* of $\Delta D^+/r^+$ and y_c^+; its appearance is quite similar to that of figure 5.5. Instead figure 5.8 shows the joint *pdf* of r^+ and y_c^+, which appears to be uncorrelated. Therefore the increment in $\Delta D^+/r^+$ is to be attributed to the increase of link strength rather than the decrease of the typical range of action of clustering tracers.
- The main direction in which clustering happens is y, *i.e.* particle pairs tend to decrease their distance in the y component much more frequently and intensively than x or z. This is probably due to particles being constrained by the channel walls. Section 5.3 will deal with single direction clustering.

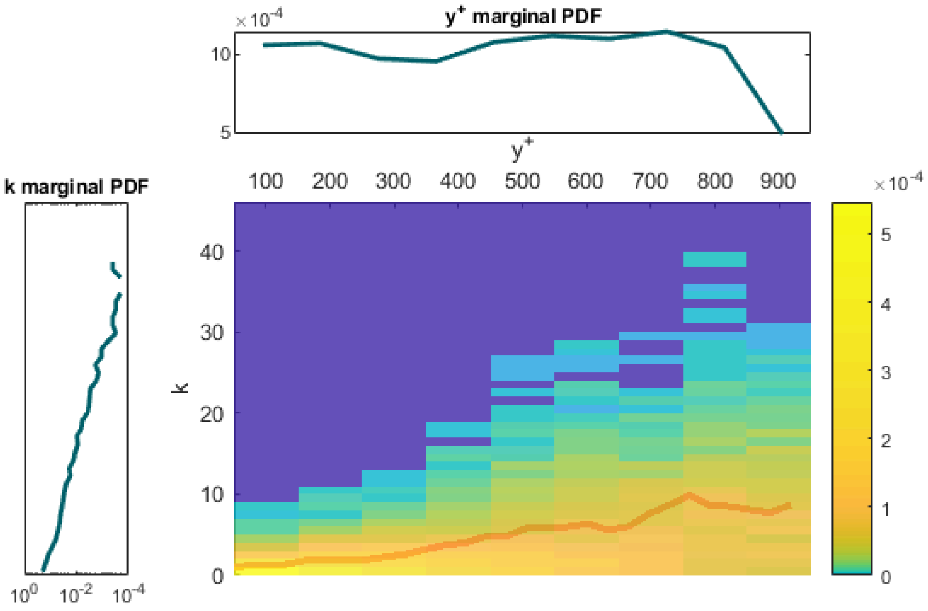


Figure 5.4: Joint probability distribution of degree and wall normal coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$. The correlation coefficient between y^+ and k is $\rho = 0.50$

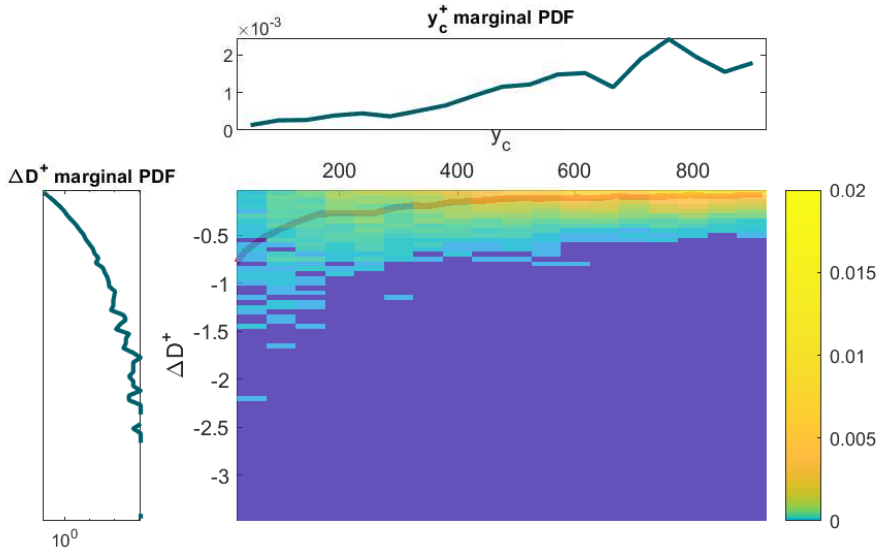


Figure 5.5: Joint probability distribution of distance decrease and centroid wall normal coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$. The correlation coefficient between y_c^+ and ΔD^+ is $\rho = 0.42$

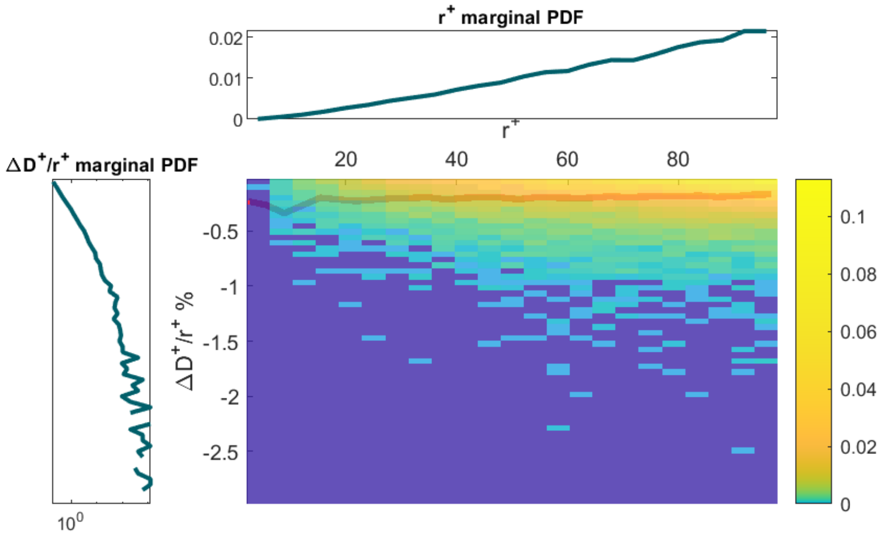


Figure 5.6: Joint probability distribution of distance decrease normalized by mutual distance and mutual distance with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$. The correlation coefficient between r^+ and $\Delta D^+/r^+$ is $\rho = 0.05$

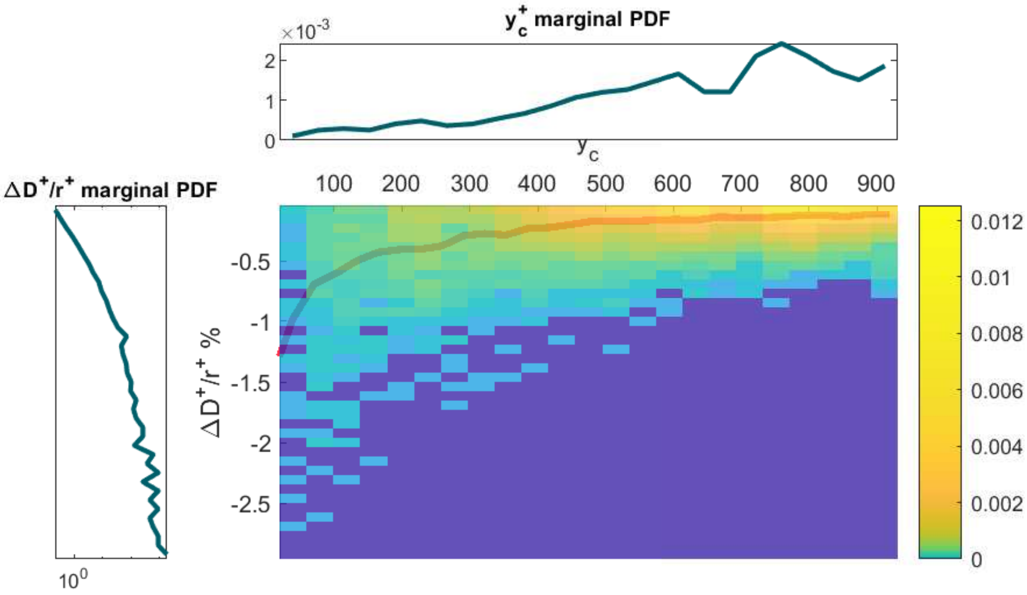


Figure 5.7: Joint probability distribution of distance decrease normalized by mutual distance and centroid coordinate with its marginal distributions and trendline, $t^+ = 475$, $d_t^+ = 100$. The correlation coefficient between y_c^+ and $\Delta D^+/r^+$ is $\rho = 0.41$

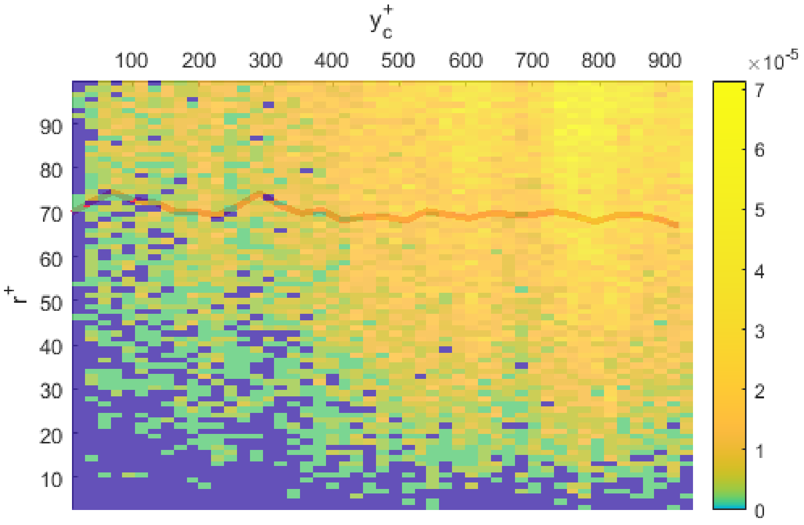


Figure 5.8: Joint probability distribution of y_c^+ and r^+, with its trendline, $t^+ = 475$, $d_t^+ = 100$. The correlation coefficient between y_c^+ and r^+ is $\rho = -0.03$

5.3 Clustering along the x and y direction

5.3.1 x-clustering

While clustering happens mostly in the y direction because of the boundedness of the flow, it may be of interest to analyse this behaviour in the streamwise direction, that is x. In order to do that, a modified mutual distance matrix, $\mathbf{D}_x$ is introduced:

$$D_{x,ij} = |x_i^+ - x_j^+|. \tag{5.4}$$

Since imposing the threshold only on the x direction would result in particles too far away being linked, the threshold will still be imposed on the Euclidean distance, as was done before.

For a given time and threshold, the number of clustering links along x is always greater than the number calculated using the Euclidean distance. This happens probably because links are established also between particles whose overall distance is increasing while their streamwise distance is decreasing. As an example, the number of links in the case presented in section 5.2 was 10 792, while with the same time instant and threshold it is now 11 363.

Figures 5.9 and 5.10 show the distributions of degree and distance decrease versus y^+; they are qualitatively very similar to the Euclidean clustering case, with the degree increasing towards the centreline and the link strength increasing towards the channel walls. Again the k marginal distribution appears to follow a similar trend to the Euclidean case, albeit slightly more irregular.

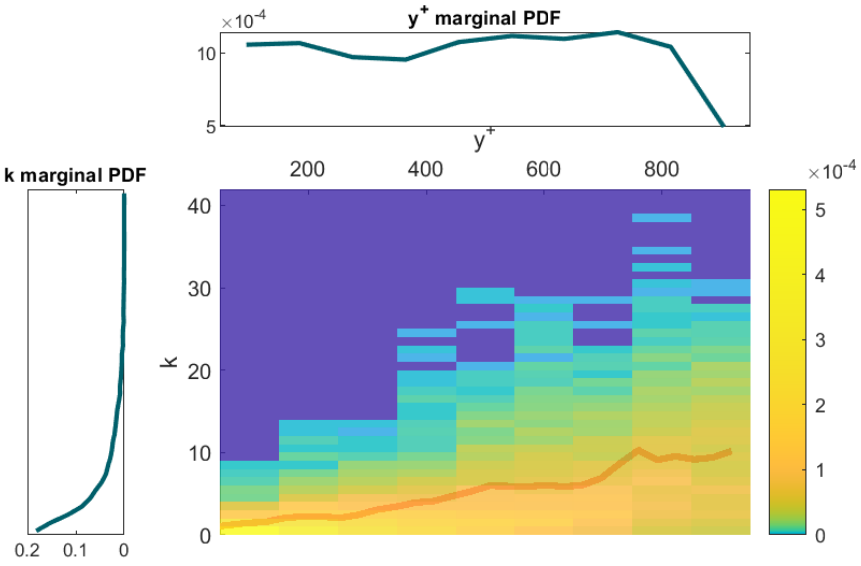


Figure 5.9: Joint probability distribution of degree and wall normal coordinate, x-clustering, $t^+ = 475$, $d_t^+ = 100$

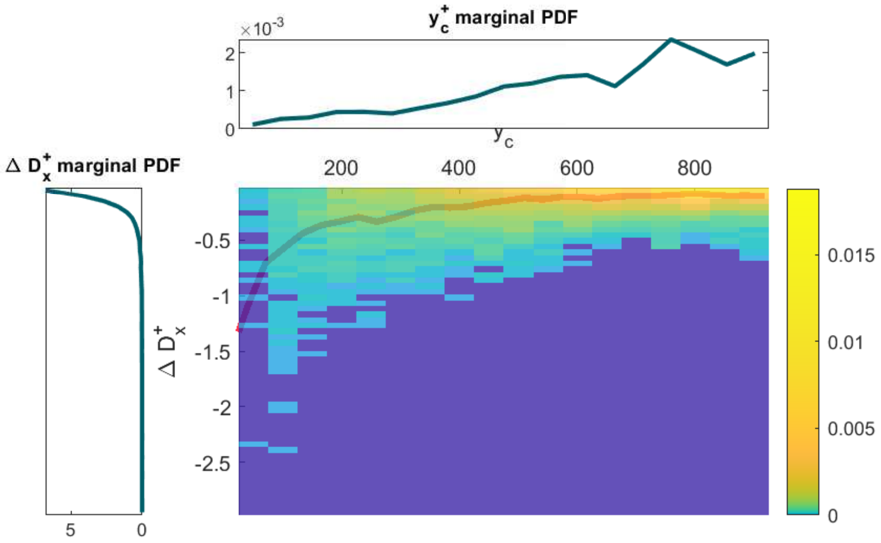


Figure 5.10: Joint probability distribution of distance decrease and centroid wall normal coordinate, x-clustering, $t^+ = 475$, $d_t^+ = 100$

Figure 5.11 shows instead the joint *pdf* of r^+ and ΔD_x^+, where r^+ is as before the Euclidean distance between particles before clustering. Again, this distribution is similar to the previous case; this may lead to conclude that taking into account single direction clustering does not provide significantly different results.

5.3.2 y-clustering

In an analogous way clustering along the y direction can be analysed. The y distance matrix $\mathbf{D}_y$, similarly to (5.4), can be defined:

$$D_{y,ij} = |y_i^+ - y_j^+|. \tag{5.5}$$

Figure 5.12 show the joint probability distribution of the distance decrease and mutual distance; as the other joint probability distributions is very similar to the case of Euclidean clustering. The only noticeable difference (as for the x-clustering case) is that the mean link intensity is slightly higher.

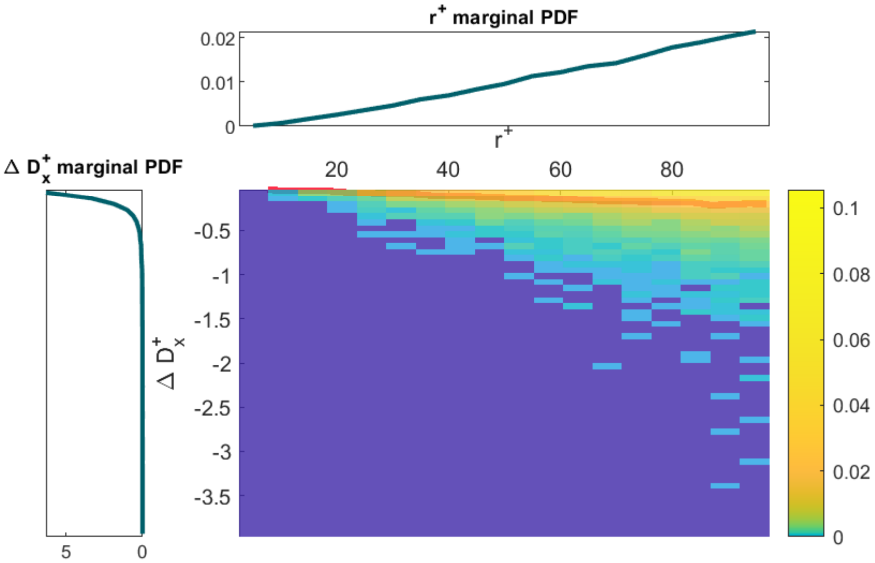


Figure 5.11: Joint probability distribution of distance decrease and mutual distance, x-clustering, $t^+ = 475$, $d_t^+ = 100$

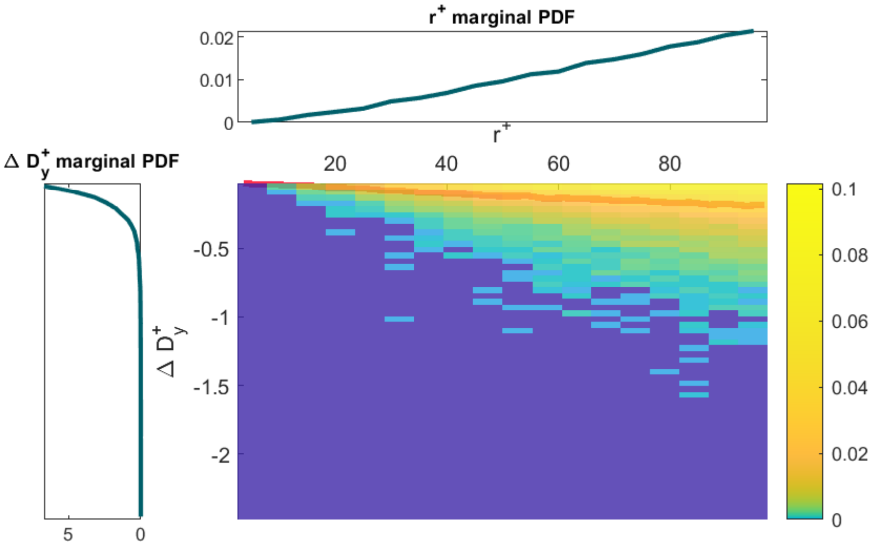


Figure 5.12: Joint probability distribution of distance decrease and mutual distance, y-clustering, $t^+ = 475$, $d_t^+ = 100$

5.4 Threshold value and degree evolution

Two approaches have been employed to define the threshold value, with different underlying motivations:

- A variable threshold, bound to the mean distance $\langle D\rangle$ between all particles at any given time. This is done to account for tracers getting farther away and retaining a considerable amount of them into the network.
- A fixed threshold for any instant of time. This is perhaps the most physically correct approach, since flow structures do not change size throughout the evolution of the flow; because of this, the number of tracers present in each structure gradually reduces with time because of diffusion. Nonetheless, it is rather difficult to estimate characteristic length scale of the flow starting from Lagrangian data alone; in addition to that, the low number of tracers may lead to difficulties in computing statistical quantities of interest. Said threshold may be tied to a given percentage of the initial or final mean distance between particles, that is $\langle D^+\rangle(t^+ = 0) = 1298$ or $\langle D^+\rangle(t^+ = 15\,200) = 19\,476$.

5.4.1 Variable threshold

Figures 5.13 and 5.14 show both the mean and standard deviation of the degree k of the network of clustering tracers, for different percentage values of the mean distance $\langle D\rangle$. The mean degree is exactly proportional to the number of connections; this value, except for an initial transient, stays nearly constant during the evolution of the flow. This could mean that the average distance $\langle D\rangle$ reflects quite well the probability for particles to interact one with the other. For large times, the threshold becomes greater than the channel height 2δ, so there is no correspondence to physical scales present in the flow.

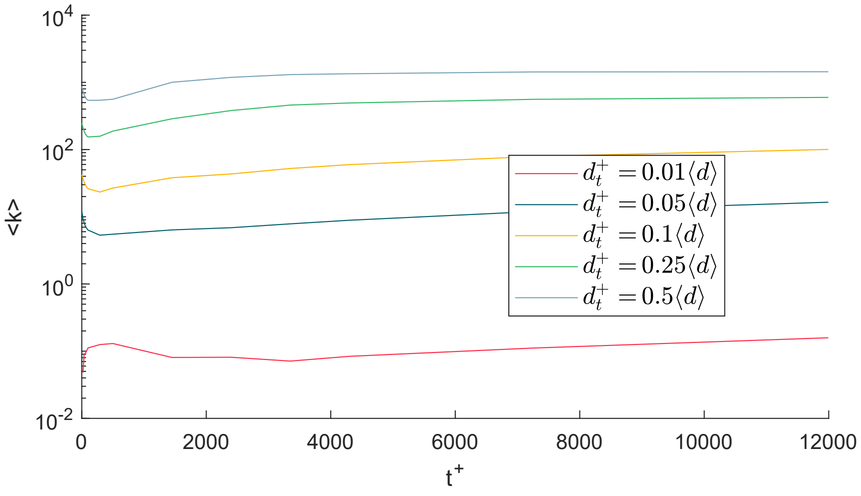


Figure 5.13: Mean degree, variable threshold

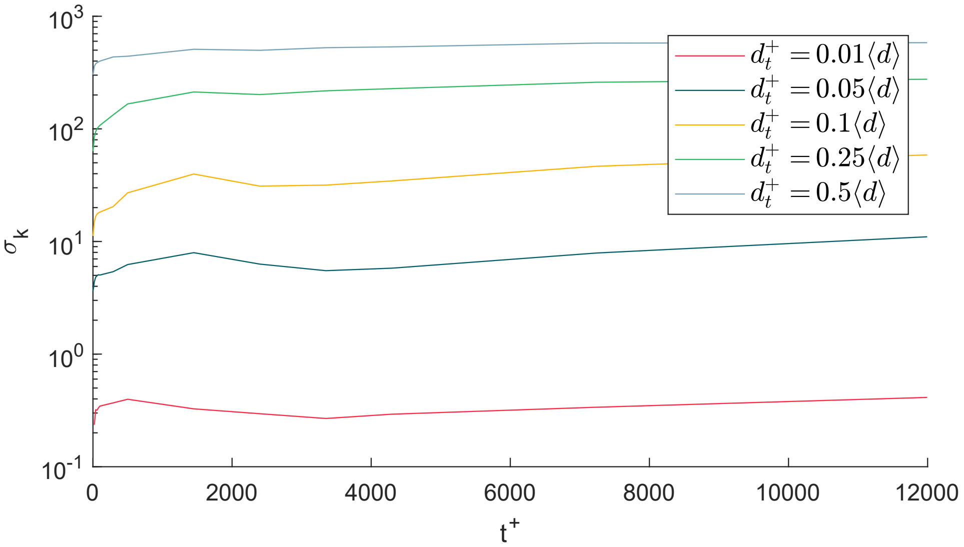


Figure 5.14: Degree standard deviation, variable threshold

5.4.2 Fixed threshold

Figures 5.15 to 5.16 show the evolution of the degree for different values of a fixed threshold d_t. The mean degree (and thus the total number of connections) is diminishing through time, as was expected. The standard deviation of k has an initial growth, then follows the trend of $\langle k\rangle$. The mean degree normalized by its standard deviation (figure 5.17) shows an initial transient stage in which curves for different thresholds are indistinguishable; a similar behaviour is shown in figure 5.18 for the kurtosis of $\langle k\rangle$; its relatively high values are due to the presence of a significant number of outliers in the degree distribution. The fixed thresholds are respectively equal to $(0.02, 0.04, 0.08, 0.12, 0.20, 0.28)$ times the initial mean distance or, alternatively, to $(1.3, 2.6, 5.2, 7.8, 13.0, 20.8) \cdot 10^{-3}$ times the final mean distance.

It should be noted that statistics for the lowest thresholds are computed on a low number of tracers and therefore may be unreliable. On the other hand, the greatest threshold is $d_t^+ = 350$, which is smaller than half channel height δ and thus may include relevant structures in the turbulent flow. The evolution of the degree statistics has a turning point at $t^+ \approx 4000$, a time at which other changes in the behaviour of the flow happened (see figure 4.1 on page 23); after this point an asymptotic regime in dispersion may be identified.

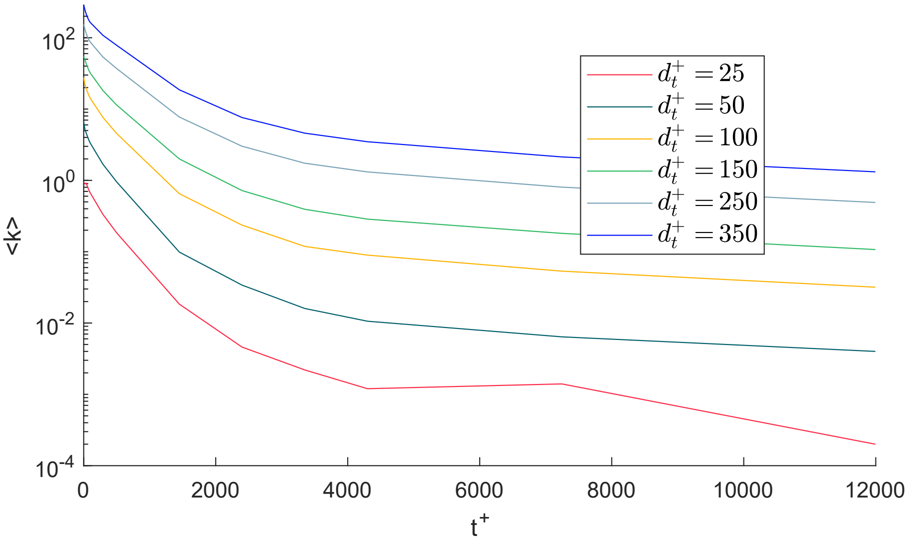
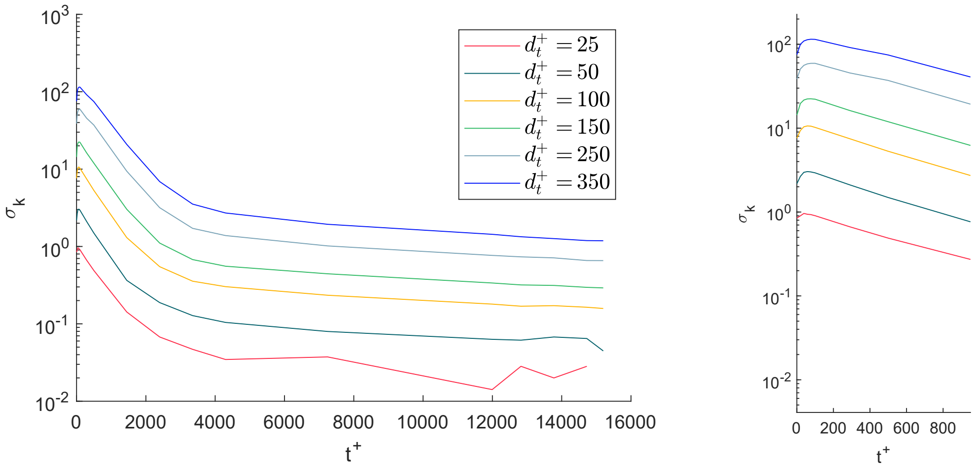


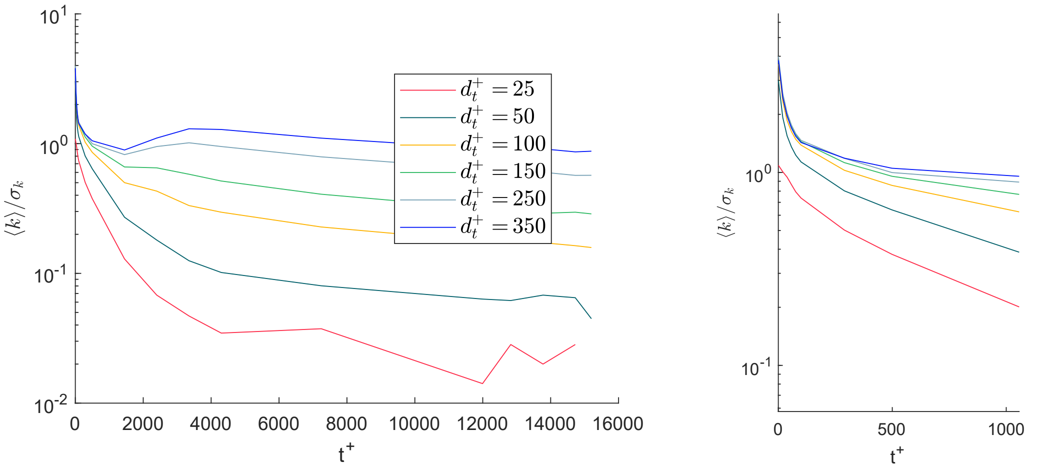
Figure 5.15: Mean degree, fixed threshold



(a)

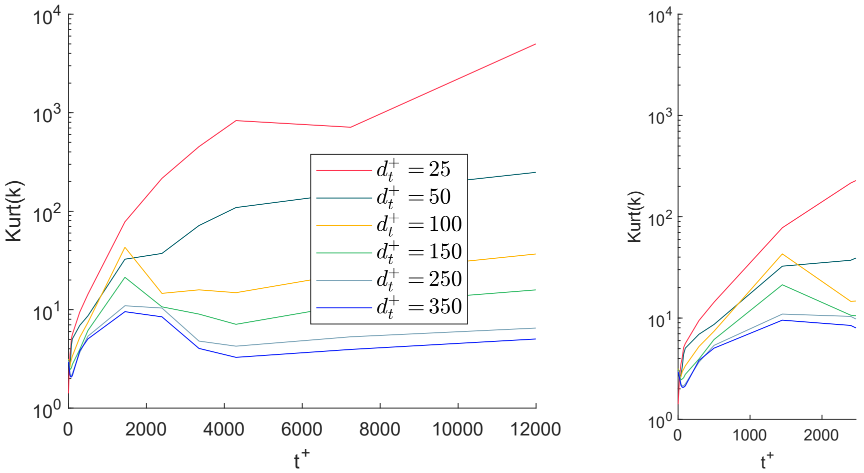
(b) Starting transient

Figure 5.16: Degree standard deviation, fixed threshold



(a) (b) Starting transient

Figure 5.17: Mean degree normalized by its standard deviation, fixed threshold



(a) (b) Starting transient

Figure 5.18: Kurtosis of k, fixed threshold

As can be seen, for certain threshold values the mean degree drops (and sometimes even starts from) under 1. This happens because tracers without links are considered too in its computation. It may be useful, in the framework of a constant threshold, to evaluate degree statistics only for linked nodes, therefore excluding tracers whose degree is equal to zero. In the following distribution of this modified degree will be indicated by the apex *. The evolution of $\langle k^* \rangle$ is shown in figure 5.19; as expected it does not drop below one. The double logarithmic plot again highlights the presence of different phases in the evolution of $\langle k^* \rangle$, although a bit fuzzier than it was for $\langle k \rangle$. Again by $t^+ \approx 4000$ the shift in regime appears to be concluded. This representation highlights the fact that, even taking into account only the nodes which cluster, a reduction of connectivity with time takes place.

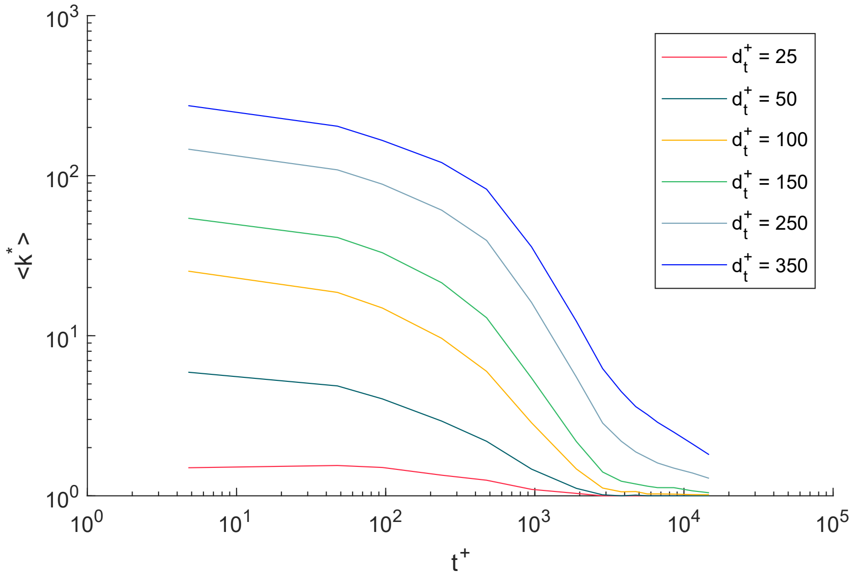


Figure 5.19: Evolution of $\langle k^* \rangle$ for different threshold values

The evolution of the degree of tracers is shown, decomposed according to the particles' wall normal coordinate, is shown in figure 5.20. As can be seen, most links are formed near the centreline and the centrality value decreases rapidly everywhere in the domain.

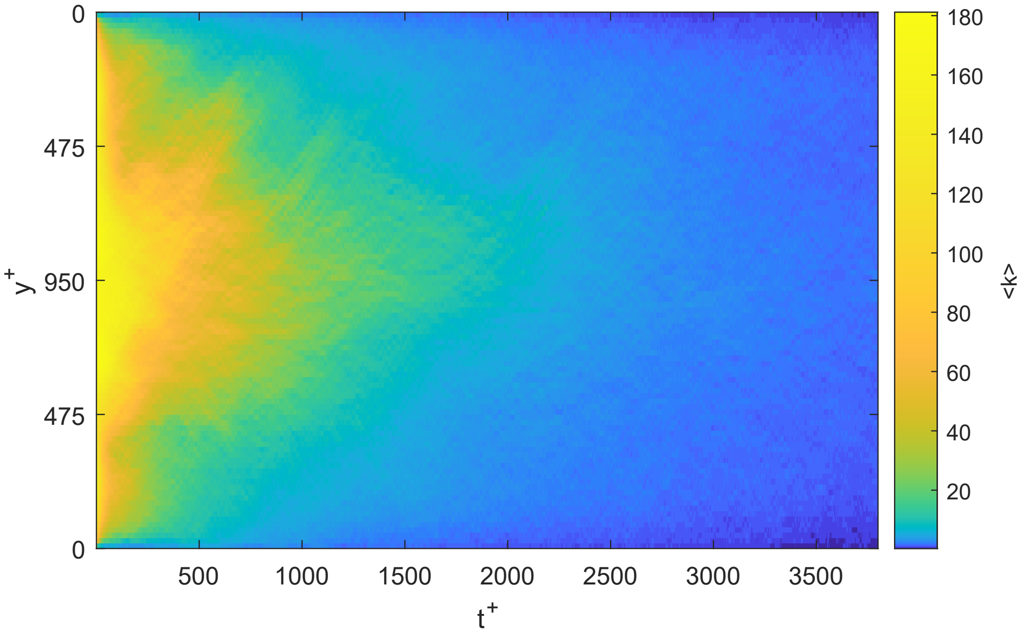


Figure 5.20: Degree of nodes

5.5 Temporal and spatial evolution

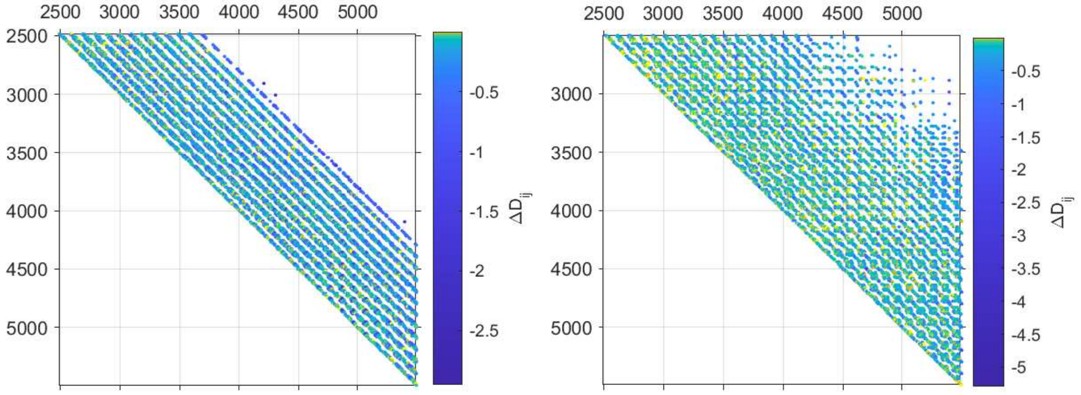
This section will deal with the dynamical and spatial characterization of the clustering network; indeed, links and nodes are distributed in Euclidean space and evolve with time in non trivial ways which require proper investigation. In particular, aspects such as the evolution of the probability distribution presented in section 5.2 on page 35, the onset and duration of clustering links and their inhomogeneous distribution in space will be discussed.

A first, simple step is the analysis of the evolution of the adjacency and strength matrices.

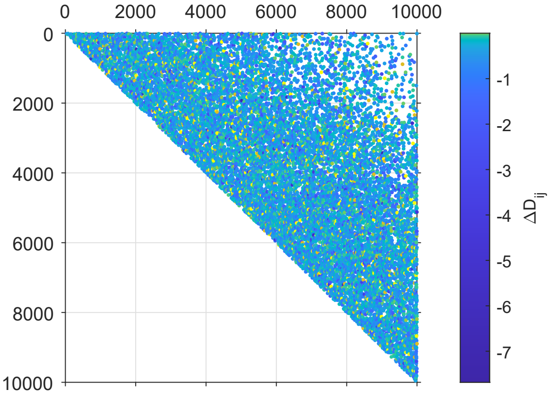
Figure 5.21 shows part of the adjacency matrix $\widetilde{\mathbf{\Delta D}}$ for different times. At the start of the simulation only neighbouring particles form clustering links; thus only the main diagonal and the 100th, 200th… upper diagonals (the lower diagonals are not computed due to matrix symmetry) have non-zero elements, due to the distribution of particles in the dataset matrix. Over time particles move away from their original location and are able to create new link with nodes that were originally too far away. In the end the distribution of non-zero elements in $\widetilde{\mathbf{\Delta D}}$ becomes nearly homogeneous since all information regarding the starting coordinate is lost.

5.5.1 Evolution of marginal probability density functions

Joint *pdf* for relevant quantities in the flow network are shown in section 5.2 on page 35 and following, coupled with their marginals; quantities deserving further analysis are



(a) $t^+ = 9.5$ (b) $t^+ = 237$



(c) $t^+ = 3990$

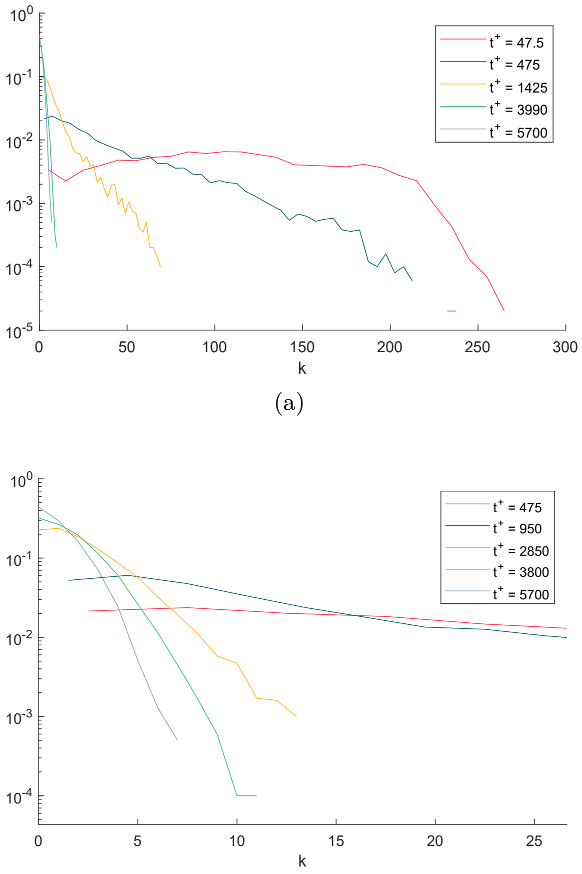
Figure 5.21: Non-zero elements of $\widetilde{\mathbf{\Delta D}}$, weight superimposed (constant threshold $d_t^+ = 350$)

- The degree distribution $P(k)$
- The node strength distribution $P(s)$
- The spatial distribution of links across the height of the channel, by means of the distribution of the centroid of linked tracers $P(y_c^+)$
- The mutual distances distribution of linked particles $P(r^+)$

In the following the probability distribution functions of the aforementioned quantities are calculated for different times. A fixed threshold $d_t^+ = 250$ is employed. The probability distribution of k, shown in figure 5.26, exhibits an exponential behaviour as was already noted previously; this behaviour is more marked for long times, while for very short times ($t^+ < 400$) the distribution is not exponential. As time goes the maximum connectivity decreases and the distribution becomes narrower. The exponential distribution has equation

$$P(k) = \mathrm{e}^{-\lambda k}; \tag{5.6}$$

The exponential coefficient λ can be calculated finding the slope coefficient of the k-$\log(P(k))$ line; the trend of λ is shown in figure 5.23. As expected from the probability density function plots, λ increases with time. An asymptotic value is not to be reached since the diffusion of tracers is still ongoing and the distribution is expected to become more and more narrow.



(a)

(b)

Figure 5.22: Probability distribution function of k for different times

Figure 5.24 shows the evolution of $P(y_c^+)$, that is the centroid coordinate of links. Links are concentrated outside the inner part of the boundary layer, so at $y^+ \geqslant 100$. For small times (approximately $t^+ < 2000$) there is a smaller number of links in the first half of the channel height δ ($y^+ < 475$) than near the centreline of the channel. For larger times, links are more or less equally distributed across the channel (but still outside the inner part of the boundary layer).

Figure 5.25 shows the distribution of the mutual euclidean distances of linked nodes. As was shown before, the distribution is oriented towards large values of r^+ (of course

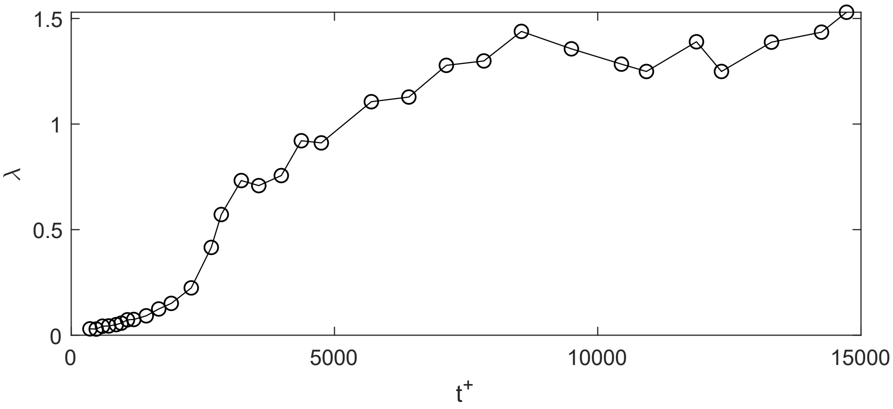


Figure 5.23: Exponential distribution coefficient λ

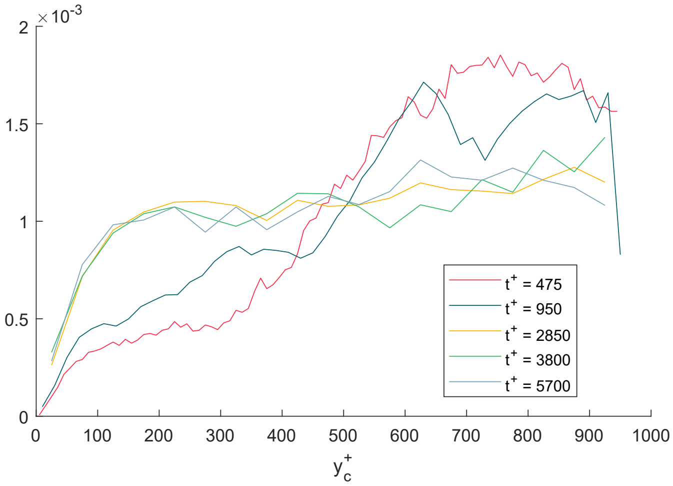


Figure 5.24: Probability distribution function of y_c^+ for different times

limited by the threshold value imposed in this computation). As already shown in figure 5.6 for relatively small times the distribution appears to be linear, while growing faster for larger times. This points that, as time goes, links are formed in prevalence between farther particles; this is an obvious consequence of the mean distance between tracers getting bigger and the imposition of a fixed threshold, so that short distance links form with a smaller probability with time growing. The adoption of a variable threshold makes this probability distribution function linear for different times.

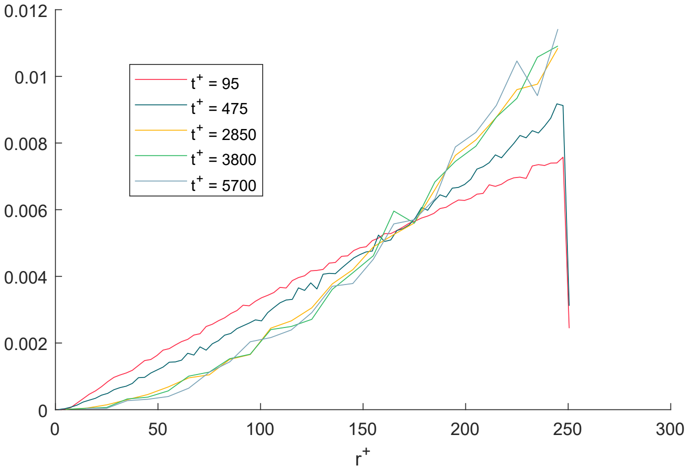


Figure 5.25: Probability distribution function of r^+ for different times

At last the distribution of the strength of the nodes is shown in figure 5.26. Clustering strength ΔD^+ again appears to follow an exponential distribution for different time instants; as time increases, stronger links become more prevalent. Recalling from the joint probability distribution 5.6 on page 38 that r^+ and ΔD^+ are positively correlated, it can be noticed how the prevalence of large r^+ links leads to an increase of their strength.

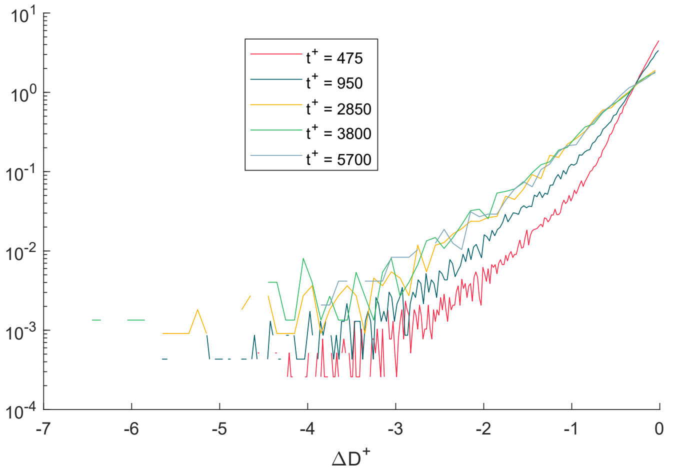


Figure 5.26: Probability distribution function of ΔD^+ for different times

Finally, evolution of mean quantities (excluding the degree, already shown in section 5.4.2) will be presented in the following. Figure 5.27 shows the evolution of $\langle y_c^+ \rangle$ for four different threshold values; the threshold appears to have no significant influence on the behaviour of $\langle y_c^+ \rangle$, which reaches first a maximum ($y_c^+ \approx 650$, about a third of channel height) and then an asymptotic value ($y_c^+ \approx 500$) for $t^+ > 2700$; the time by which an asymptotic value is attained is significantly lower than the time instant observed before as a shifting point in flow behaviour (which was $t^+ \approx 4000$).

Figure 5.28 shows the evolution of $\langle r^+ \rangle$; normalization by the threshold make different curves collapse into one, except for very small times. Furthermore, as already stated, the mean distance increases since tracers get farther away one from the other and, with a constant threshold, this make longer range links more likely. Nonetheless, the trend seems to show an asymptote starting at $t^+ \approx 4000$ and equal to about $\langle r^+ \rangle / d_t^+ = 0.75$

Figures 5.29 and 5.30 show the evolution of $\langle \Delta D^+ \rangle$, both non-normalized and normalized by the linked tracers mutual distance. At a single point in time, higher r^+ values were correlated with stronger links of clustering nodes; this is still the case if evolution in time is considered: the value of $\langle r^+ \rangle$ and so does the (absolute) value of $\langle \Delta D^+ \rangle$. Again, an asymptote is reached at $t^+ \approx 4000$. The evolution of $\langle \Delta D^+ \rangle$ is the main factor that shapes the evolution of $\langle \Delta D^+ \rangle / d_t^+$.

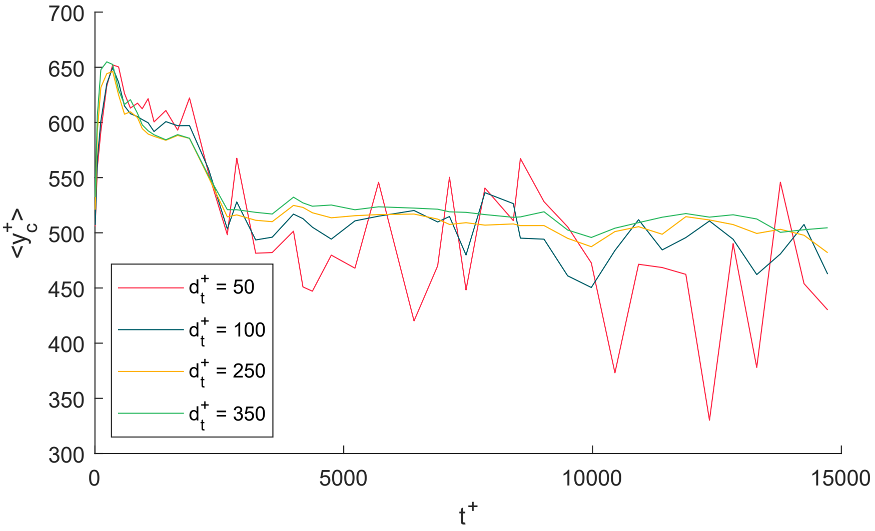


Figure 5.27: Mean centroid wall normal coordinate

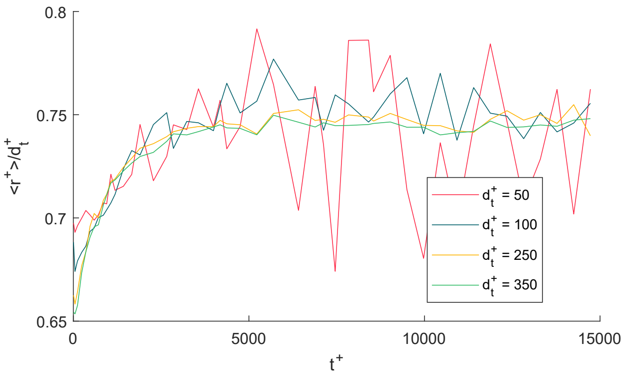


Figure 5.28: Mean mutual distance normalized by threshold value

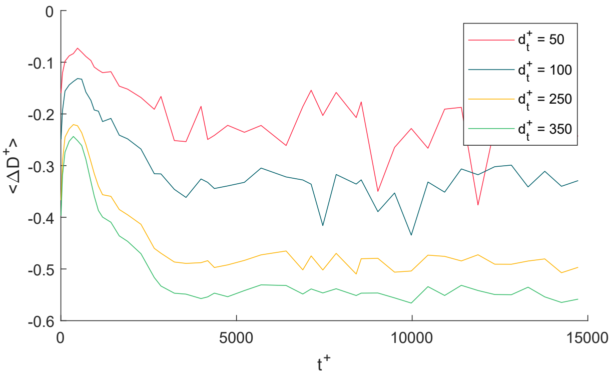


Figure 5.29: Mean link strength

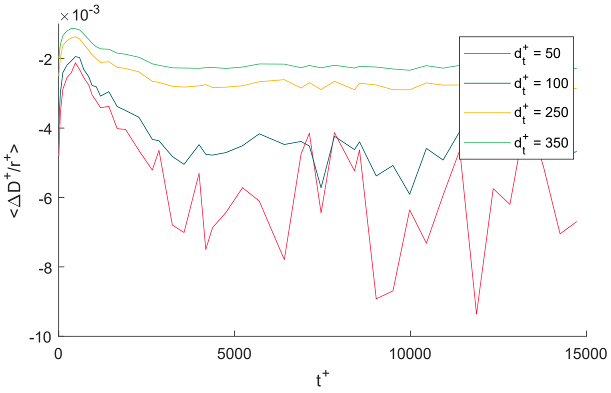


Figure 5.30: Mean link strength normalized by mutual distance

5.5.2 Onset and persistence of clustering

It is reasonable to ask whether clustering between individual tracers is a long lasting phenomenon or not. The duration and the average intensity of clustering can yield information about the intensity of structures in which tracers are contained. Indeed, if a structure is collapsing rapidly, the duration of clustering will be reduced for sure, since the maximum relative distance a pair of tracer can cover is bounded. This means that structures with a smaller characteristic timescale (*i.e.* those closer to the walls, for this particular case) will produce clustering links of shorter duration and *vice versa*. Also, larger velocity fluctuations, which are again found near the walls, may reduce the duration of links.

This hypothesis has been put to test calculating the mean link duration $\langle T_l^+ \rangle$ and the associated total reduction in distance for several time instants and wall normal coordinates. The link duration is the amount of time in which a pair of tracers is getting closer one to the other, that is the amount of time in which their relative velocity is negative. Instead the mean total reduction in distance $\langle \Delta D_l^+ \rangle$ is the difference between the relative distance when clustering ends and when it starts. To highlight differences due to the wall normal coordinate of the link centroid (as defined in (5.3)), links have been also ordered according to their belonging to three different distances from the wall: $y^+ \in [0\ 250)$, $y^+ \in [250\ 475)$ and $y^+ \in [475\ 950]$. A threshold value of $d_t^+ = 250$ has been chosen, being smaller than the smallest y^+ partition.

Figure 5.31 shows the evolution of the mean link time; as expected, it is larger in general for links farther from the wall. This, along with the data for $\langle \Delta D_l^+ \rangle$ (figure 5.32), confirms that structures leading to clustering near the walls are in general more intense than those at the centreline.

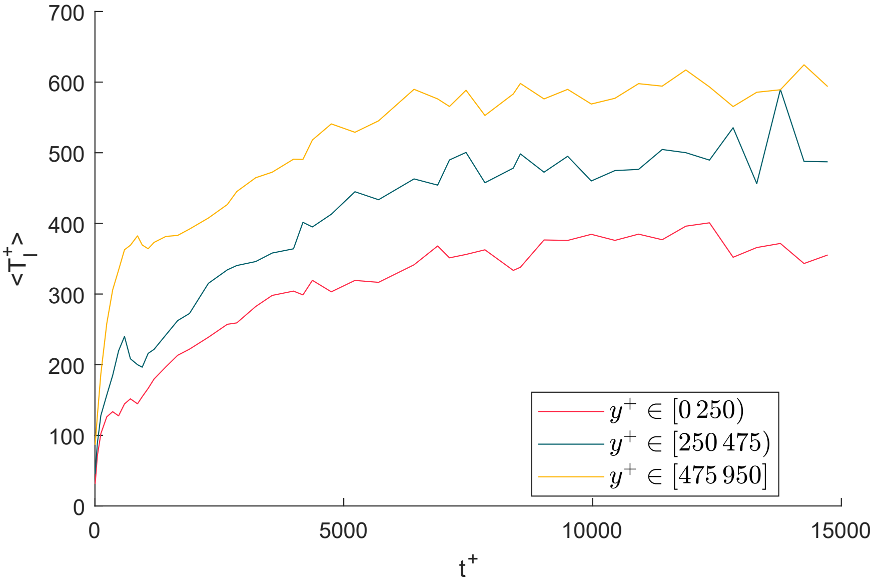


Figure 5.31: Mean duration of links $\langle T_l^+ \rangle$

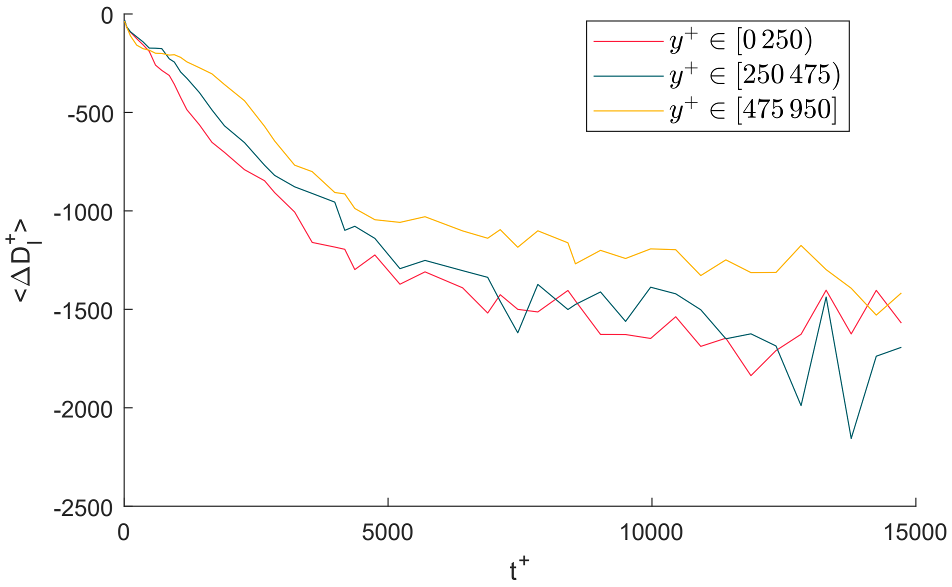


Figure 5.32: Mean distance reduction for full links $\langle \Delta D_l^+ \rangle$

It is also worth noting that the mean duration of links is of the order of the hundreds; thus clustering is not an instantaneous phenomenon and its persistence may be associated to that of the flow structures that generate it. Values of the total distance reduction greater than d_t^+ are obtained since, when moving backwards in time for each trajectory to calculate the starting time of the clustering event, also time instants where the mutual distance was greater than the threshold were included. Figure 5.33 shows an example of that: while the mutual distance is lower than the threshold at the time at which the clustering network matrix is computed (blue circle), it is far larger at the time at which the clustering begins. This leads to large values of clustering times and total strength.

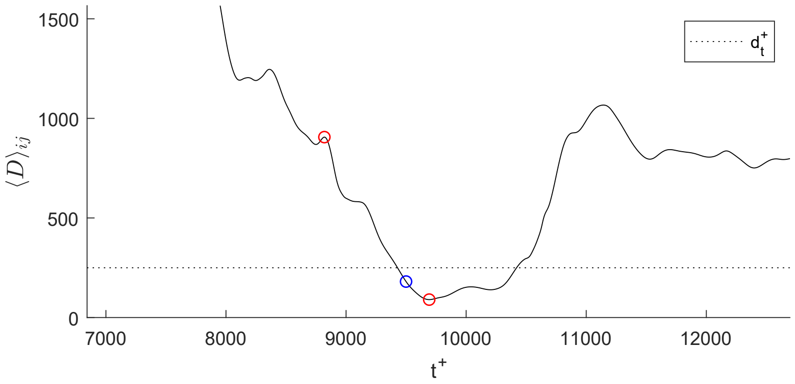


Figure 5.33: Mutual distance of a pair of particles. Red circles: start and end of the clustering event; blue circle: $t^+ = 9500$

Figures 5.34 and 5.35 show the evolution of both mean link duration and mean distance reduction with the threshold. Even if limited by the threshold, values of the duration are still quite large and are comparable both in magnitude and trend with Lagrangian velocity timescales.

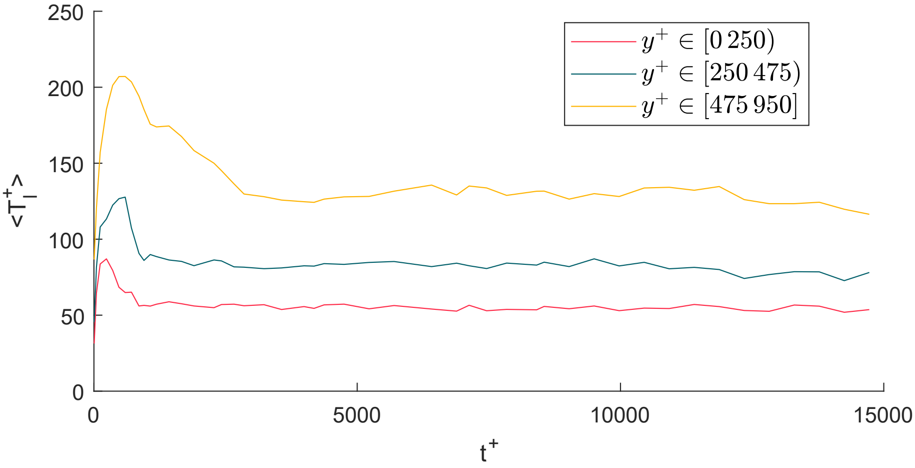


Figure 5.34: Mean duration of links $\langle T_l^+ \rangle$, threshold included

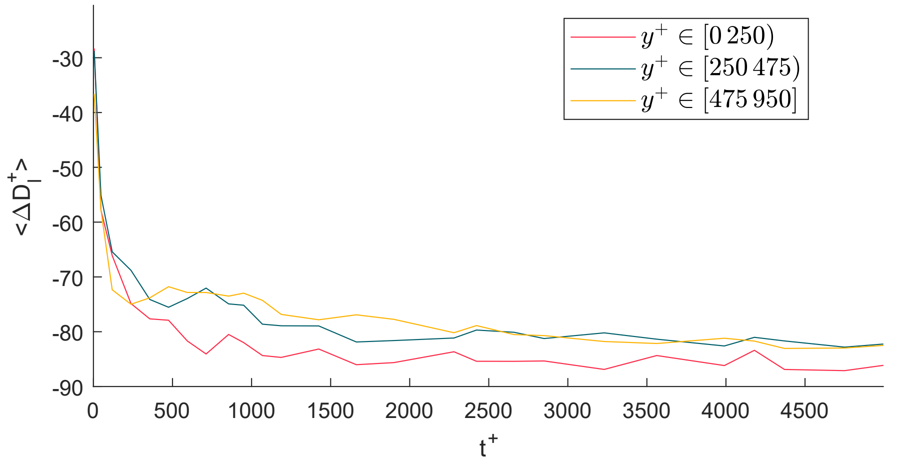


Figure 5.35: Mean distance reduction for full links $\langle \Delta D_l^+ \rangle$, threshold included

5.5.3 Spatial inhomogeneity

As already noted in the previous section, clustering appears not to be homogeneous in space. The following analysis will try to characterize in deeper detail this behaviour. Figure 5.36 shows the distribution of particles in the channel as seen from the streamwise direction, in the form of a joint probability distribution and its marginal *pdf*s. The distribution is not homogeneous, especially along the z direction where it can be noted that particles appear to be diffusing away from the mean z^+ coordinate; in the y coordinate particles are instead bounded.

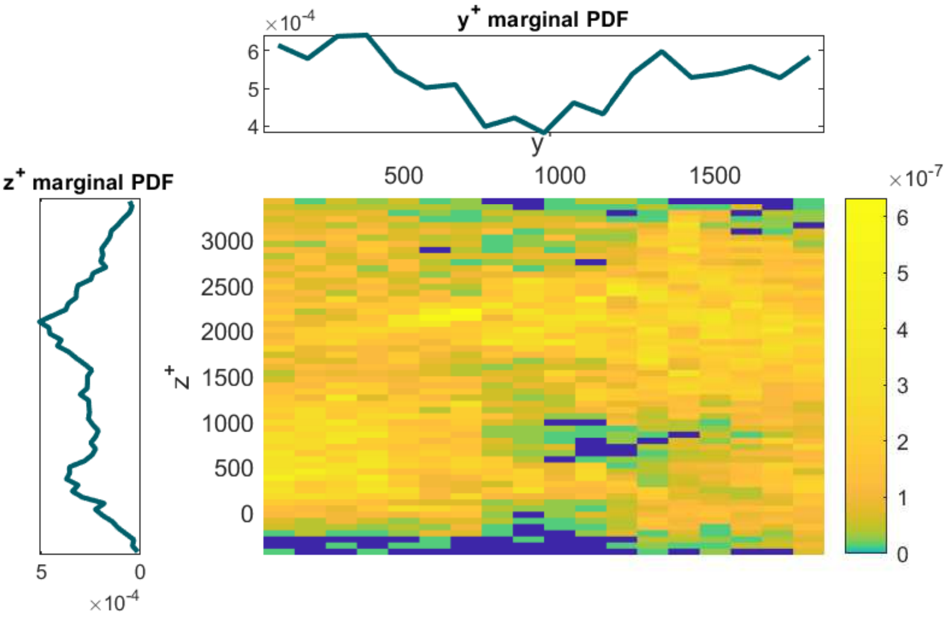


Figure 5.36: Distribution of tracers in the yz plane, $t^+ = 950$, $d_t^+ = 350$

Figure 5.37 shows instead the distribution of link centroid coordinates $\mathbf{x}_{c,ij}$, again projected onto the yz plane; these coordinates are calculated as the mean position of two linked particles, that is, as defined in (5.3),

$$\mathbf{x}_{c,ij} = \frac{\mathbf{x}_i + \mathbf{x}_j}{2}.$$

As can be seen, the link coordinate distribution somewhat follows that of tracers, but is more definite and limited to a reduced number of zones in the flow. In particular links are not present at all in zones where there is a reduced number of tracers, or at high (or low) z^+. Also, the borders of the link distribution are somewhat jagged and irregular. The latter could be explained by the fact that particle are diffusing away from the mean z^+ coordinate and are thus less prone to clustering.

As was already noted before, the y_c^+ marginal distribution is quite different from the y^+ one, since links tend to arise away from the wall while tracers are scattered throughout the channel. In addition to that, the centreline of the channel is somewhat devoid of tracers; this is certainly the same behaviour as in figure 6.3 on page 66. The z_c^+ and z^+ distributions are more similar.

Furthermore it should be noted that for this case ($t^+ = 950$, $d_t^+ = 350$) while particles are (as always) 10 000, clustering links are about $2{\cdot}10^5$; so the distributions of the $_c$ quantities are more detailed.

Aside from visual inspection of distributions, it may be useful to provide some metrics to describe the spatial inhomogeneity of clustering links. To a certain degree, an object with a highly irregular distribution and indented boundaries could be assimilated to a fractal, keeping in mind that the irregularities in a fractal are not inferiorly bounded since they are purely mathematical objects, while their physical counterparts are not (be it turbulent eddies, the distribution of clustering links or broccoli [6]). In this particular case the inferior bound is not the Kolmogorov scale η as in turbulence, but rather the definition provided by tracers, which can become very low as diffusion goes on. Nonetheless in a certain regime a fractal behaviour can be observed (as in the inertial scale of the energy cascade).

Fractals are, from the mathematical standpoint, objects characterized by a non-integer dimension $\mathcal{D}$; if the fractal is contained in a plane, then $1 < \mathcal{D} < 2$ (as for the Koch's fractal curve, figure 5.38). The calculation of $\mathcal{D}$ is not straightforward and relies on different definitions. The one used on the following for its practicality in application is the *Minkowski–Bouligand dimension*, or *box counting* dimension [1] [19]. In a non fractal planar object ($\mathcal{D} = 2$) the number of squares of side ε required to cover the object will, say, quadruple if ε is halved. This does not hold for a fractal and instead the following definition can be introduced:

$$\mathcal{D} = \lim_{\varepsilon \to 0} -\frac{\log N(\varepsilon)}{\log \varepsilon}, \tag{5.7}$$

where $N(\varepsilon)$ is the number of boxes of side ε which contain part of the object.

In this case, the fractal object is the distribution of clustering links in the yz plane,

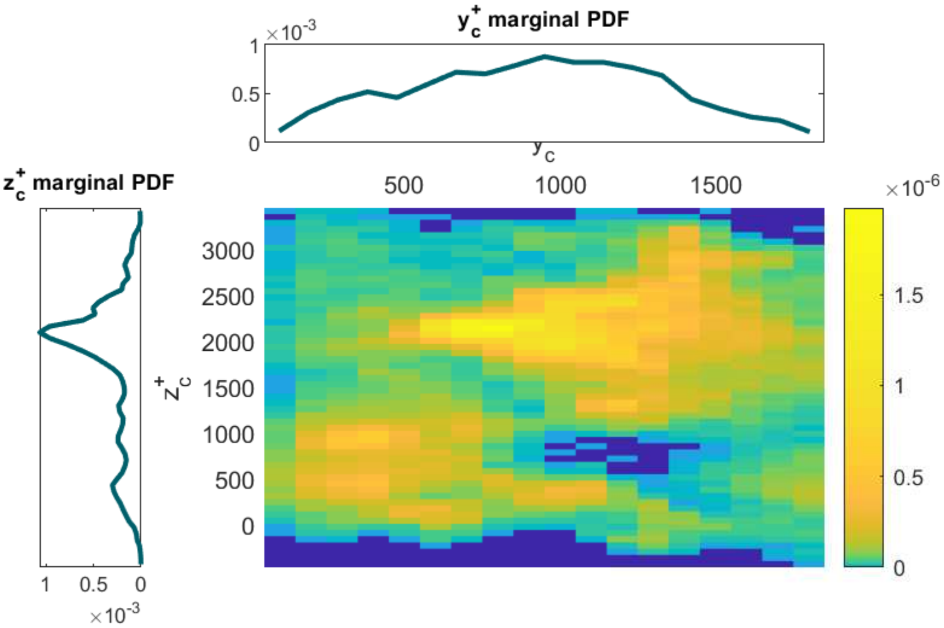


Figure 5.37: Distribution of links in the yz plane, $t^+ = 950$, $d_t^+ = 350$

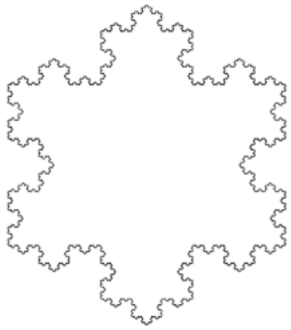


Figure 5.38: Koch's snowflake, $\mathcal{D} \approx 1.26$

which has been partitioned in an increasing number of rectangular boxes of side

$$\varepsilon = \sqrt{l_y l_z}$$

(the introduction of a different box shape does not alter the result; circles have been used too). The projection of the distribution of links on the yz plane has been chosen both because it seems that it may represent well the global inhomogeneity of the distribution and because a three dimensional partition does not appear to be computationally feasible; nonetheless, information on the three dimensional distribution may be inferred [21]. The limit operation cannot be properly executed for the already mentioned problems in the dataset resolution; a linear fitting has been used to calculate the dimension $\mathcal{D}$ as the slope of the $N(\varepsilon)$ versus ε curve in a double logarithmic plot.

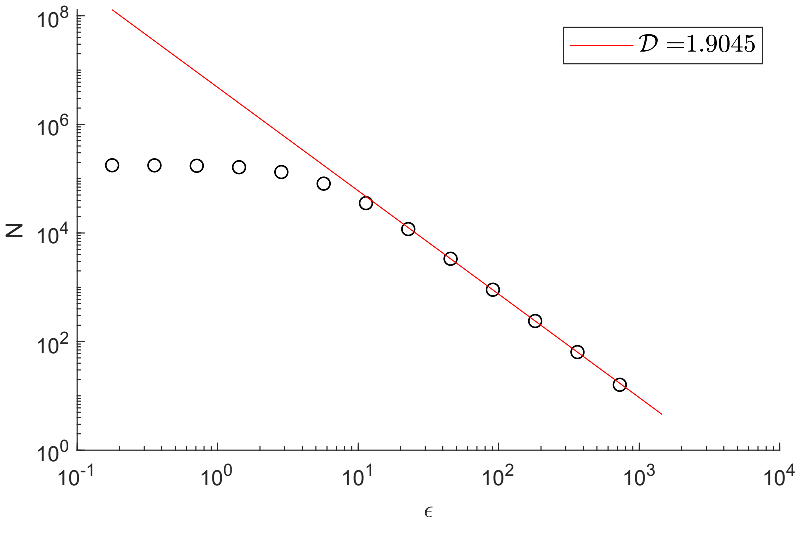


Figure 5.39: $N(\varepsilon)$ versus ε and fractal dimension, $t^+ = 950$, $d_t^+ = 350$

The results are reported, for a single time, in figure 5.39 and the calculated fractal dimension is $\mathcal{D} = 1.9$; it can be noted that the slope of the curve (ans so the dimension) tends to zero as $\varepsilon \to 0$, since increasing the number of boxes does not increase the number of links which is always about $2 \cdot 10^5$ (which is also the maximum value attainable by $N(\varepsilon)$). The value of $\mathcal{D}$ is close to 2, so a true fractal behaviour is not observed.

The evolution of $\mathcal{D}$ in time is shown in figure 5.40; as can be seen the Minkowski dimension of the distribution of links is smaller than 2 and decreases with time, indicating that the distribution becomes more irregular. However, this behaviour may as well be due to the diminishing number and to sparsity of links leading to numerical inaccuracies.

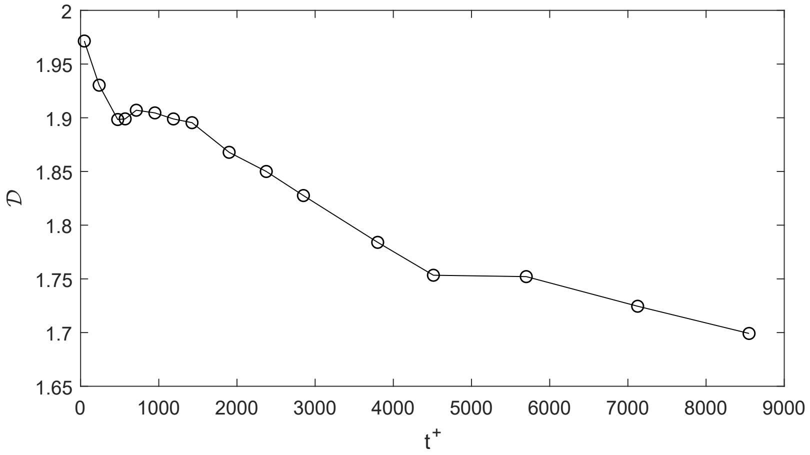


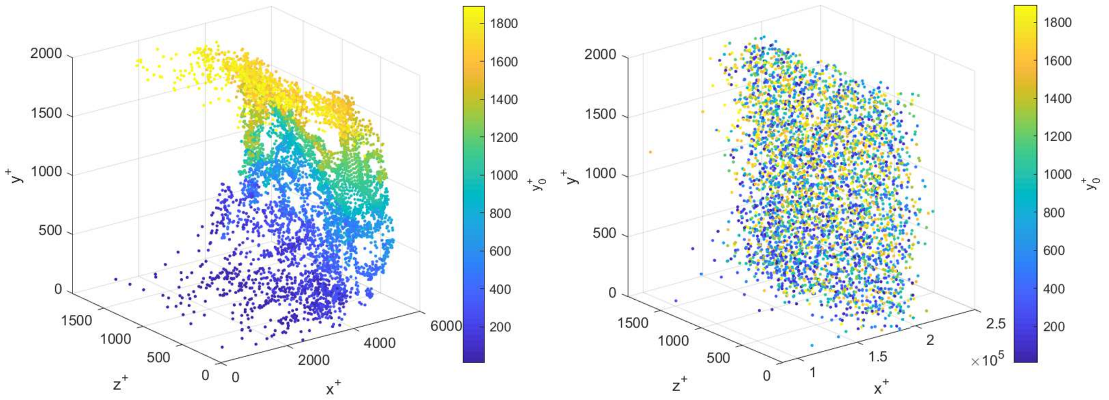
Figure 5.40: Fractal dimension $\mathcal{D}$ versus time, $d_t^+ = 350$

Chapter 6

Flow map network

6.1 Transport mapping

A conceptually simple way of analysing transport properties of turbulent flows is to follow the motion of an advected scalar (or a tracer particle, for the scope of this work) between different, separate regions of the fluid domain [20]. The dataset employed in this thesis is particularly suitable for this approach, since the fluid domain can be partitioned in the only inhomogeneous direction (the wall normal one) and the measure of particles belonging to each partition is straightforward. Figure 6.1 shows the distribution of particles and their starting level, which will be characterized by means of complex networks in the following. For short times (figure 6.1a) tracers are disposed orderly in the channel, close to their starting level; instead, for long times (figure 6.1b) they are diffused thoroughly and information about their origin is lost.



(a) $t^+ = 240$

(b) $t^+ = 9500$

Figure 6.1: Distribution of particles in the channel; the starting level is color coded

The channel is divided in one hundred equally spaced sections, each one with height $\Delta y^+ = 19$. Since the tracer particles are also released starting from one hundred different levels at $y^+ = 9.5, 28.5 \ldots$, each section will contain, at $t^+ = 0$, exactly 100 tracers.

A transport matrix $\mathbf{P}^k$ for each time instant t^k can be defined as

$$P_{ij}^k = \frac{N_{i\to j}^k}{N_i^0} \tag{6.1}$$

where $N_{i\to j}^k$ is the number of particle that at time t^k have moved from the i-th level (which they occupied at $t = 0$) to the j-th. Being the number of particles originally present in each level, N_i^0 is always equal to 100. The matrix shifts from an initial state in which only diagonal terms are present, to one where particles are diffused through the channel height and the matrix is almost filled.

Figure 6.2 shows the evolution of $\mathbf{P}$; along with the aforementioned diffusion of particles, it can be seen that particles near the walls diffuse slower than the rest of the set; this may be due to their inherently lower velocity.

$\mathbf{P}$ is asymmetric and has the following features:

- the sum of the elements of each row is always equal to 1, since it is the number of particles originally belonging to the i-th level normalized by itself (both are equal to 100). So
$$\sum_{i=1}^{N_l} P_{ij} = 1, \ \forall j$$
with N_l being the number of levels;
- the sum of each column is equal to N_j^k/N_j^0, namely the number of tracers contained in the j-th level at $t = t^k$ normalized by 100. It is therefore a measure of the tracer *density* of each level.

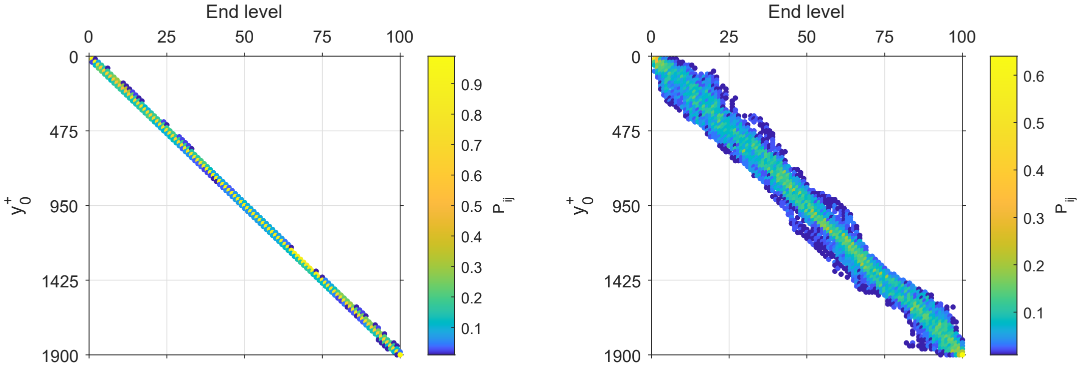
Figure 6.3 shows the difference, evolving in time, between the number of particles present in each level and the starting number (always equal to 100). Therefore, this plot is always inferiorly bounded by -100, which corresponds to an emptied level. On the x and y axis both the advancement in time and the y^+ coordinate of each level are represented.

It can be seen that the distribution of tracers in the wall normal direction becomes immediately inhomogeneous, with zones having a greater concentration than others. In particular, for $t^+ < 500$ tracers appear to move towards the centreline; shortly afterwards, the concentration near the walls increases. At the end, the distribution of particles becomes nearly homogeneous (but not constant).

6.2 Network building and an example

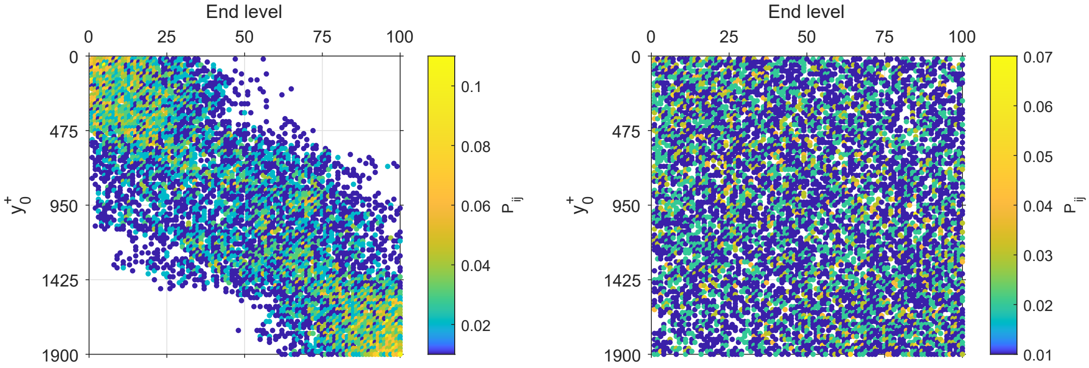
The flow map matrix can be easily interpreted as a network weight matrix; being $\mathbf{P}$ asymmetric the corresponding network will be directed. It is also necessary to ignore the terms p_{ii} on the diagonal since nodes do not have links to themselves; doing that, only the net transfers in and out of levels are considered.

In this section a reduced example of a flow network is presented, in order to further clarify concepts related to its temporal evolution. The example consists of a fictitious three



(a) $t^+ \simeq 10$

(b) $t^+ \simeq 100$



(c) $t^+ \simeq 1000$

(d) $t^+ \simeq 10000$

Figure 6.2: Evolution of $\mathbf{P}$

layered channel ($\{l_1, l_2, l_3\}$, corresponding to the nodes). Each layer at $t = 0$ contains 10 particles. Two consecutive time instants t_1 and t_2 are defined with their particle dynamics and respective graphs in order to exemplify the main features. The distribution of particles, together with their respective starting level, is tabulated in table 6.1.

Table 6.1: Particle distribution for the example flow network. The subscript denotes the starting layer of particles

	t_0	t_1	t_2
l_1	10_1	$8_1 + 1_2 = 9$	$6_1 + 4_2 + 1_3 = 11$
l_2	10_2	$9_2 + 2_1 + 1_3 = 12$	$6_2 + 3_1 + 1_3 = 10$
l_3	10_3	9_3	$8_3 + 1_1 = 10$
N_p	30	30	30

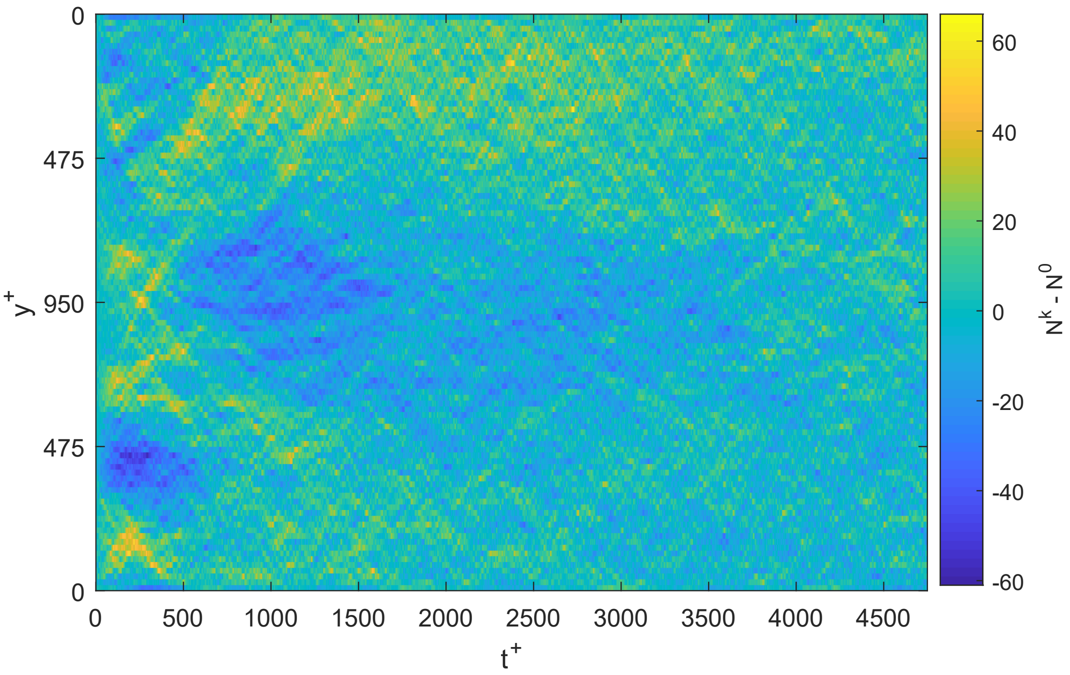


Figure 6.3: Balance between N^k and N^0, initial transient

The construction of the transport matrix $\mathbf{P}^0$ is trivial and so is its associated graph, which has order three and size zero. For $t = t_1$ instead, the motion of particles leads to a non-trivial graph; its transport matrix is

$$\mathbf{P}^1 = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.1 & 0.9 & 0 \\ 0 & 0.1 & 0.9 \end{pmatrix} .$$

It can be clearly verified that the sum of rows is always equal to 1, which is the initial number of tracers in each level (10) divided by itself. Also the sum of columns is equal to the number of particles present in each level at t_1, again divided by 10. Cancelling out the diagonal of $\mathbf{P}$, a weight matrix $\mathbf{W}$ is obtained; the corresponding network is shown in figure 6.4, with the number of moving tracers superimposed to links. For small times particles only diffuse in nearby levels, so non zero terms in $\mathbf{W}$ are concentrated near the main diagonal.

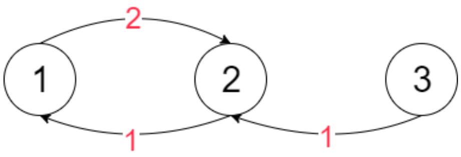


Figure 6.4: Graph of the example flow network at $t = t_1$

Further diffusion drives particles away from their starting location; at $t = t_2$ the transport matrix is

$$\mathbf{P}^2 = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.6 & 0 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}.$$

As time goes, links are created between particles originally separated by a greater distance. This is also evident since almost all terms in the **P** and corresponding **W** matrices are non zero. The network is graphically represented in 6.5.

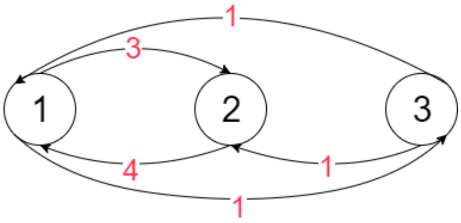


Figure 6.5: Graph of the example flow network at $t = t_2$

It is important to stress that transport matrix $\mathbf{P}^k$ shows the evolution of the network between t^0 and t^k and does not fully characterize the motion of tracers between any two consecutive time instants. As an example, the description of the motion of particles from t_1 to t_2 is not possible starting only from $\mathbf{P}^1$ and $\mathbf{P}^2$ (recalling that such matrices all have their starting point at t_0).

6.2.1 Strength and degree for the transport network

An usual metric for weighted networks is the strength of nodes s as defined in equation (3.5); since this is also a directed network, it is possible to define both a measure for outgoing links s^{OUT} and for ingoing links s^{IN}. The outgoing strength is calculated for each node i as

$$s_i^{\mathrm{OUT}} = \sum_{j=1}^{N_l} P_{ij} - P_{ii}, \tag{6.2}$$

which is the sum of each row of **P** excluding the diagonal. This value corresponds to the number of tracers that had left the i-th level at a certain time; adding the value on the matrix diagonal P_{ii} the value is always 1 for all i. Outgoing strength for both the example graphs are, respectively:

$$s^{\mathrm{OUT}}(t_1) = \begin{pmatrix} 0.2 \\ 0.1 \\ 0.1 \end{pmatrix} \quad s^{\mathrm{OUT}}(t_2) = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.2 \end{pmatrix}$$

showing that the first node donates the greatest number of particles in the network. In a similar fashion, the ingoing strength is

$$s_i^{\mathrm{IN}} = \sum_{j=1}^{N_l} P_{ji} - P_{ii}, \tag{6.3}$$

where inside the summation indexes i and j are inverted to perform the addition on columns instead of rows. This value is the number of tracers that came into the i-th level (still divided by the number of those originally present in the level); adding P_{ii} the number of particles present in the level at any time is found. The ingoing strength for the example flow is here displayed, showing that the second layer is the most active at first, but for $t = t_2$ the first node receives more tracers:

$$s^{\mathrm{IN}}(t_1) = \begin{pmatrix} 0.1 \\ 0.3 \\ 0 \end{pmatrix} \quad s^{\mathrm{IN}}(t_2) = \begin{pmatrix} 0.5 \\ 0.4 \\ 0.1 \end{pmatrix}$$

Another important metric to evaluate is the degree k, *i.e.* the number of links incident in a node; again, since this is a directed network, a distinction has to be made between the ingoing degree k^{IN} and the outgoing k^{OUT}. In order to compute the degree, the adjacency matrix $\mathbf{A}$ has to be introduced, defined as in chapter 3 on page 15; having already the network matrix, it is simply a rewriting of $\mathbf{P}$ with ones substituted to non zero elements, that is

$$A_{ij} = \begin{cases} 1 \text{ if } P_{ij} \neq 0 \\ 0 \text{ if } P_{ij} = 0 \\ 0 \text{ if } i = j \end{cases} \tag{6.4}$$

The ingoing degree k^{IN} is, as already defined in (3.2) on page 16,

$$k_i^{\mathrm{IN}} = \sum_{j=1}^{N_l} A_{ji}$$

and is the number of nodes which have given one or more tracers to the i-th level at a certain time; for the example network, the ingoing degrees are

$$k^{\mathrm{IN}}(t_1) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad k^{\mathrm{IN}}(t_2) = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

In the same way, the outgoing degree

$$k_i^{\mathrm{OUT}} = \sum_{j=1}^{N_l} A_{ij}$$

is the number of nodes which received a particle from the i-th level; again their values are

$$k^{\mathrm{OUT}}(t_1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad k^{\mathrm{OUT}}(t_2) = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

6.3 Network evolution

Using metrics and definitions introduced in the previous sections, the evolution of the transport network for the turbulent channel flow dataset will now be analysed. At first, probability distribution functions and temporal evolution of the first statistical moments for the strength s and the degree k will be shown; the evolution of s and k will also be shown using maps representing these quantities both in space and time, thus highlighting additional features.

6.3.1 Strength and degree evolution

The evolution in time of s^{IN} and s^{OUT} is shown (in the form of their probability distributions) in figures 6.6 and 6.7. The distribution of s^{OUT} rapidly becomes very narrow and centred on the unity value; this means that most of tracers originally present in a level quite rapidly leave it and diffuse throughout the channel. On the other hand, the distribution of s^{IN} also tends to an almost unitary mean (since its mean must be the same of that of s^{OUT}), but a larger deviation from the average is present. This higher variance may be related to the fact that, while particles quickly diffuse out of their starting level, the concentration of particles in the y direction (which is related with s^{IN}) is uneven at any time. All distributions have been tested for normality using the Kolmogorov-Smirnov test with a threshold p-value of 0.05; all ingoing probability distributions are Gaussian, while the outgoing are not.

Precisely the asymptote for the mean strength is situated at 0.99, as will also be stated in table 6.2 on page 74. This happens because, being tracers homogeneously distributed throughout the channel height, each level still has on average $N_i^0/N_l = 1$ particles inside, thus the asymptote cannot be 1.

The evolution of the standard deviation of s is shown in figure 6.8; the standard deviation of s^{OUT} is always smaller and reaches immediately its asymptote, while that of s^{IN} oscillates three times and reaches its asymptote at $t^+ \approx 4000$ (a time instant that was already identified previously as a turning point in the evolution of the flow).

This behaviour is shown in more detail in figure 6.9, which shows the evolution of the ingoing strength of all nodes. It follows closely the balance between ingoing and outgoing tracers (figure 6.3); after a transient stage, the strength distribution holds a constant mean, but as can be seen is quite inhomogeneous through different levels. Since s^{IN} is proportional to the number of particles each level has received, it can be seen how the tracer's concentration is fluctuating quite intensely around its mean. On the other hand, the distribution of the outgoing strength (not shown here) almost immediately reaches its asymptote and shows no fluctuations both in time and space.

Figures 6.10 and 6.11 show probability distribution function of both the ingoing and outgoing degree for several time instants. Distributions have been again tested for normality using the Kolmogorov-Smirnov test, after being subtracted their mean and divided by their standard deviation; p-values are shown in the legend and the threshold was $p = 0.05$. The degree has also been normalized by its maximum attainable value for this network, that is 99. Both distributions are quite irregular because of the low number of samples. They exhibit a slower shift towards an asymptotic value than the weight distributions, probably

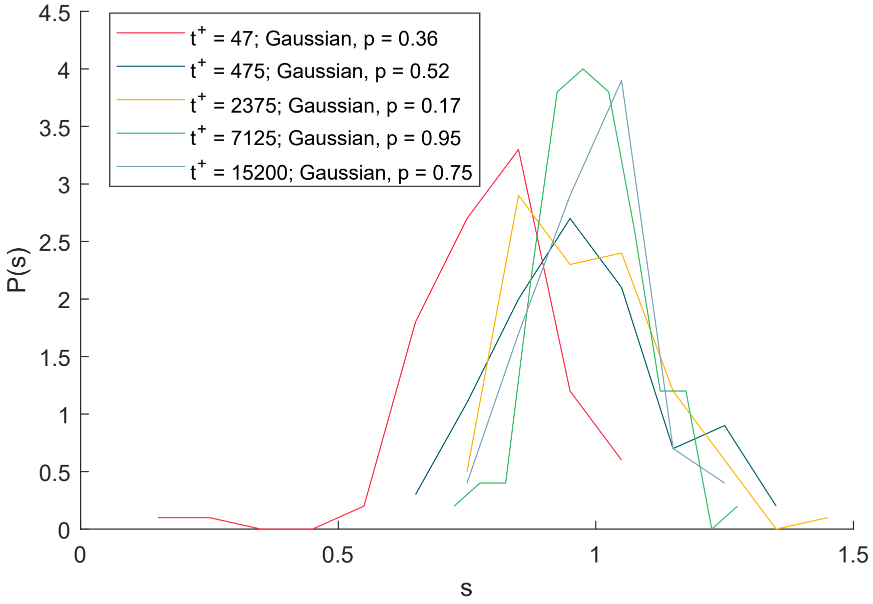


Figure 6.6: Probability distribution of the ingoing strength

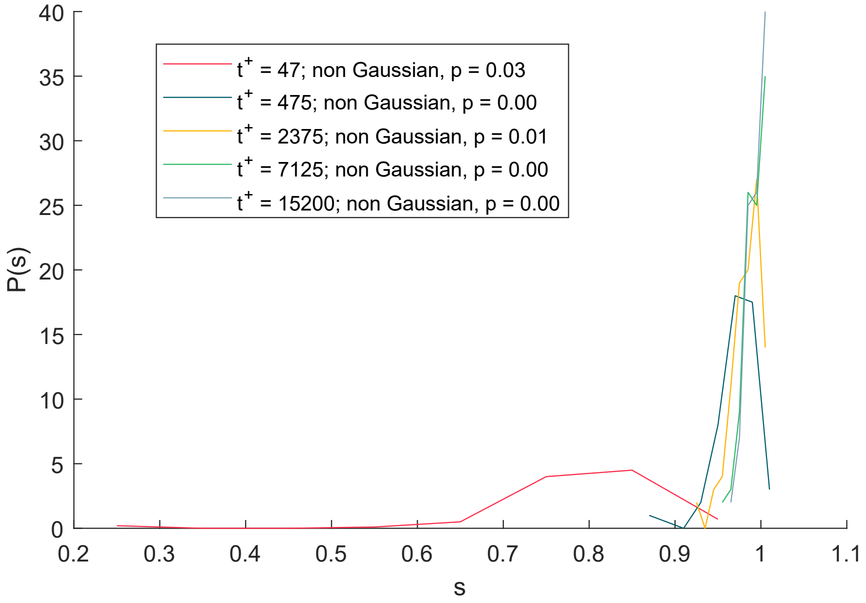


Figure 6.7: Probability distribution of the outgoing strength

meaning that the transfer of particles is faster than the establishment of connections between levels.

The evolution of both ingoing and outgoing degree is shown in figures 6.12 and 6.13.

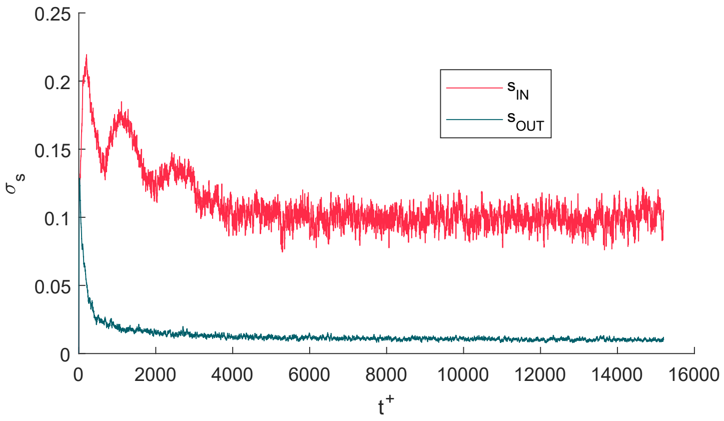


Figure 6.8: Standard deviation of s, both ingoing and outgoing

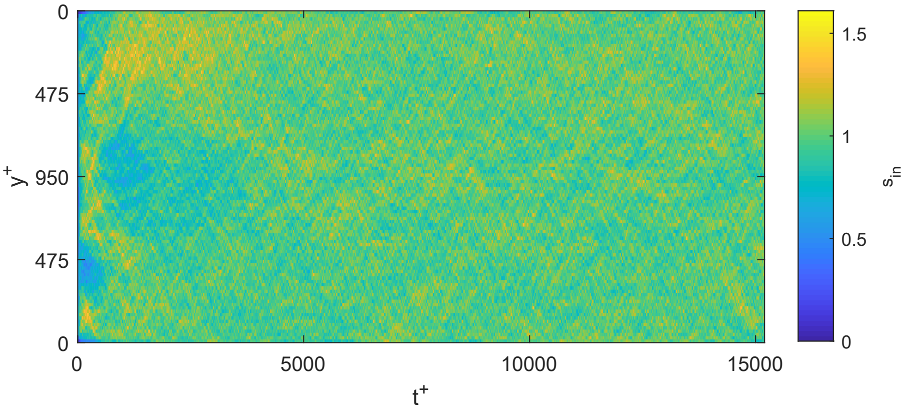


Figure 6.9: Evolution of the ingoing strength

The ingoing degree evolution shows that at the beginning the levels near the centreline of the channel receive tracers from more levels than those near the walls. After a transient period the distribution becomes more homogeneous. The outgoing degree evolution shows how levels near the wall do not communicate much with the outside (or, precisely, do not send particles to a great number of levels). Levels near the centreline instead have a greater outgoing degree; there is a well marked separation between these two zones, located near $y^+ \simeq 475$, which is one quarter channel height.

Figures 6.14 and 6.15 show the evolution in time of the mean and standard deviation of k. As already noted, $\langle k \rangle$ grows monotonically towards an asymptotic value. Also the standard deviation has an asymptote for large times, but reaches a peak and subsequently

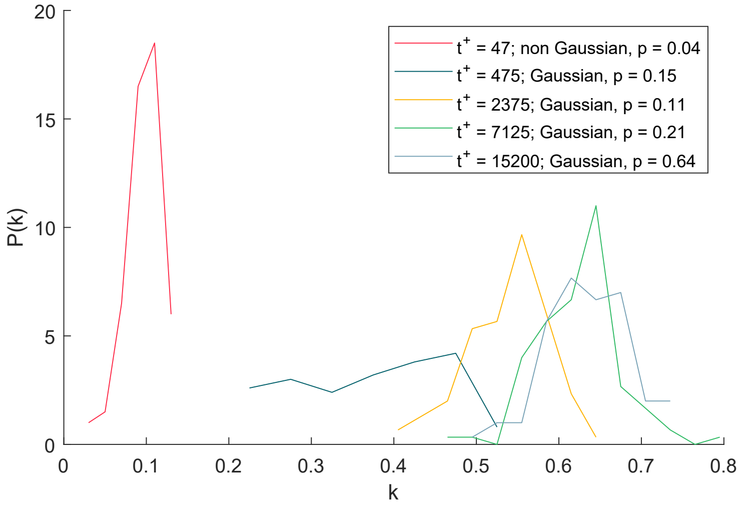


Figure 6.10: Probability distribution of k^{IN}

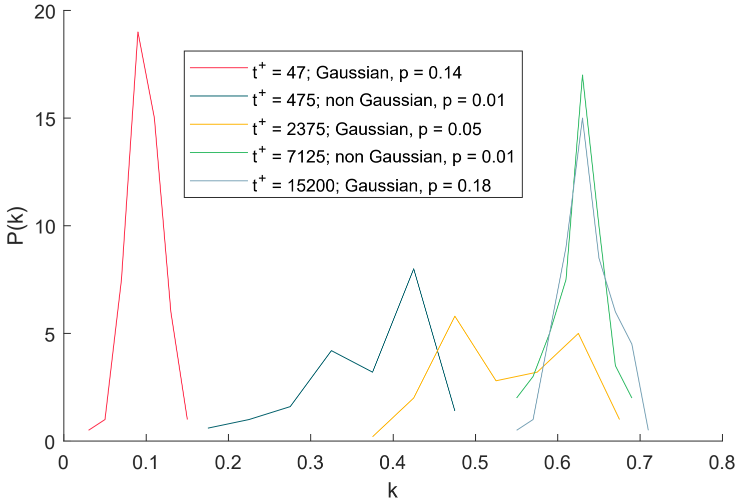


Figure 6.11: Probability distribution of k^{OUT}

decreases. The peak appears to be related with the period of inhomogeneity also visible in figures 6.12 and 6.13.

Asymptotic values for statistical quantities of the strengths and degrees are reported in table 6.2. The kurtosis has been computed using a range of consecutive time instants to provide an adequate number of samples. The kurtosis is mostly near 3, which is the value attained by the normal distribution; this means that the number of outliers in the degree and strength distributions is not large. The value of standard deviation for the $^{\mathrm{OUT}}$ quantities are smaller than those of the $^{\mathrm{IN}}$ ones.

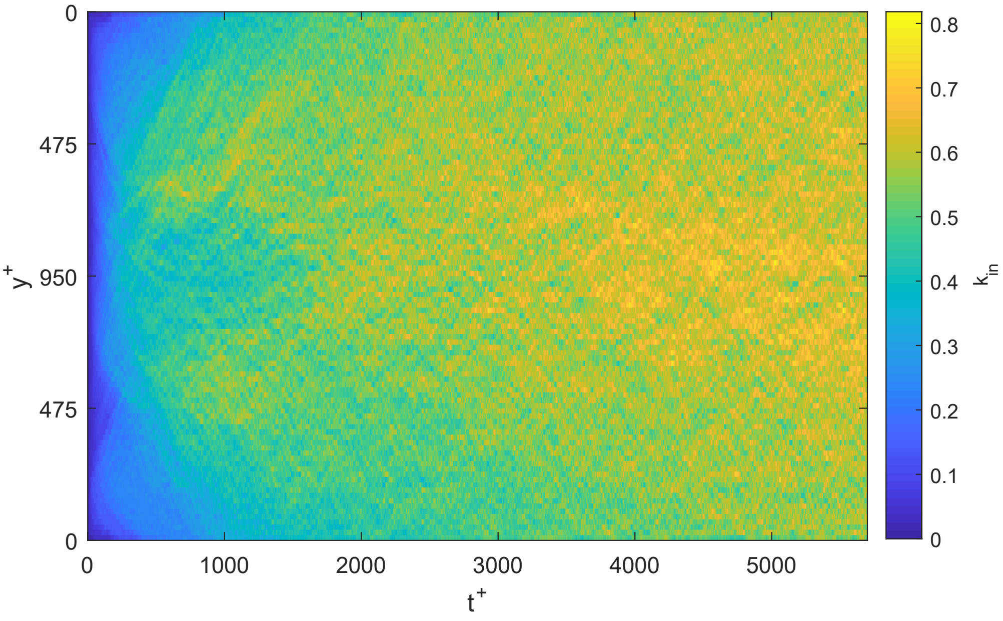


Figure 6.12: Evolution of the ingoing degree

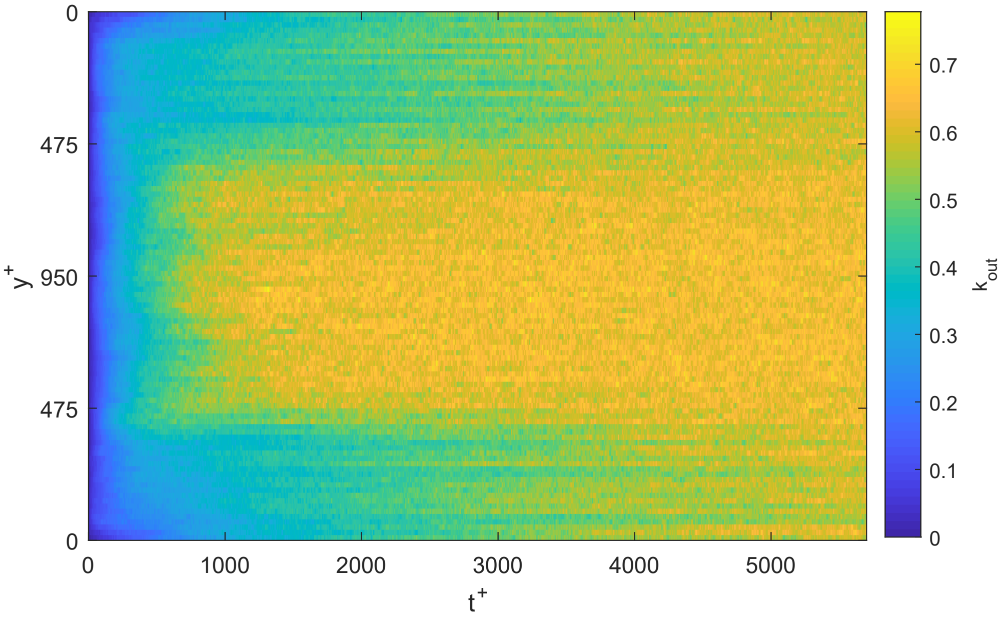


Figure 6.13: Evolution of the outgoing degree

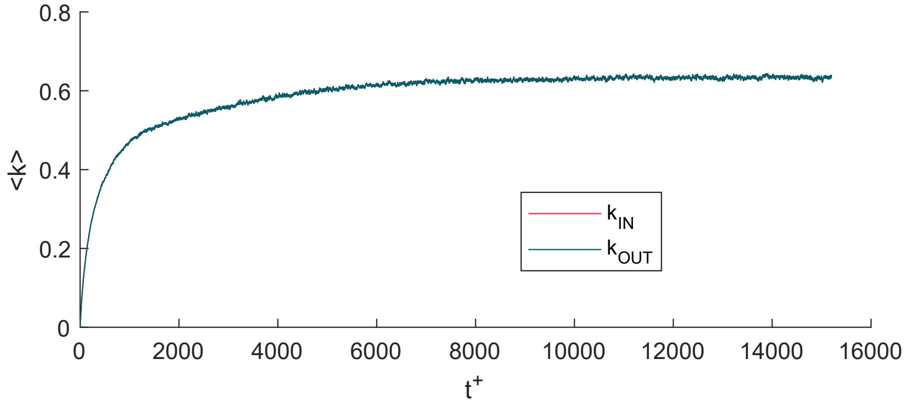


Figure 6.14: Mean degree $\langle k \rangle$

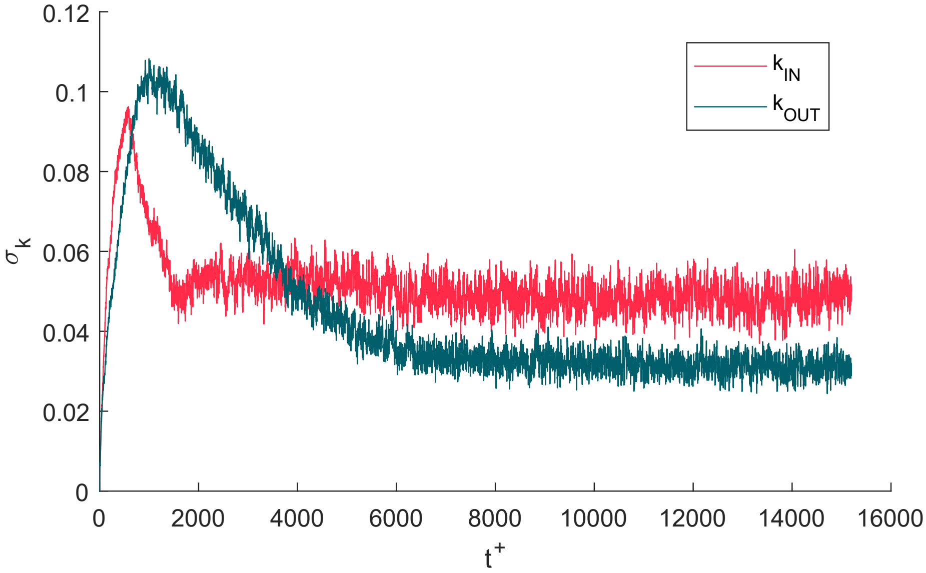


Figure 6.15: Degree standard deviation

Table 6.2: Statistical moments of s and k, asymptotic values

	s^{IN}	s^{OUT}	k^{IN}	k^{OUT}
Mean	0.99	0.99	0.63	0.63
Standard deviation	0.1	0.01	0.0475	0.0313
Kurtosis	2.7	3.8	2.9	2.9

Since the flow transport network is embedded in physical space, an additional characterization can be provided through the link distance, that is the y distance of linked nodes. Its non weighted value is calculated as

$$\langle l \rangle_{ij} = \Delta y^{+} |i - j| A_{ij}, \tag{6.5}$$

with $\Delta y^{+} = 19$ and where $\mathbf{A}$ is the adjacency matrix. In a similar fashion, its weighted variant is

$$\langle l \rangle_{w,ij} = \Delta y^{+} |i - j| W_{ij}, \tag{6.6}$$

which also retains some information about link strength. The mean evolution of both quantities is reported in figure 6.16; the weighted mean link length is shown multiplied by 100. As can be seen, both mean link lengths grow with time, signalling that longer range connections are formed. The weighted length is always larger since weights (which are here multiplied by 100) are always greater than 1.

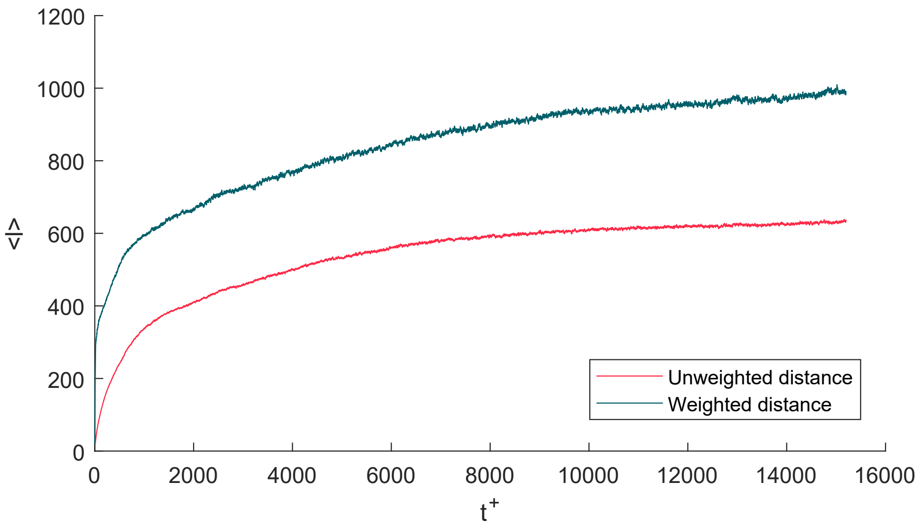


Figure 6.16: Mean link length

A further and more detailed analysis, where the mean link length is shown decomposed for each starting level, is presented in figure 6.17 and refers to outgoing links. As can be seen, as time goes the levels near the walls form longer distance links than the other levels. Instead, levels near the centreline form shorter distance links and the mean length grows slower; this is particularly true for levels comprised between $y^{+} \approx 200$ and $y^{+} \approx 450$. Besides, figure 6.17 appears to be streaked along the time axis; this indicates levels that form links with a consistently greater length than their neighbours. The cause for these *streaking* levels may be that tracers remain trapped for some time in flow structures located far away from their original y_0^{+}, thus increasing the mean link distance for some time.

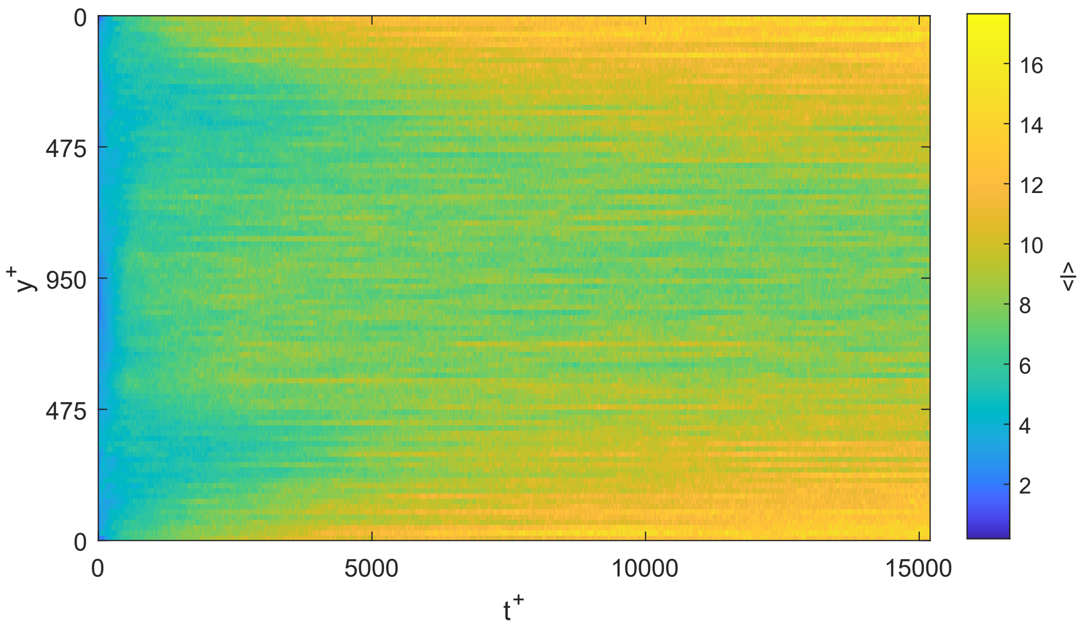


Figure 6.17: Mean outgoing weighted link length

6.3.2 Predominant links

After describing the behaviour of the flow network at single points in time and analysing the evolution of simple metrics, it may be useful to search the network for persistent links and preferential paths followed by tracers in their motion. To achieve that, a simple approach is to analyse a *cumulative* network matrix, calculated as the summation of flow maps at consecutive times

$$\mathbf{P}^{\mathbf{c}} = \frac{\sum_{k=k_0}^{k_f} \mathbf{P}^k}{k_f - k_0 + 1}, \tag{6.7}$$

where the normalization factor keeps the row-wise sum always equal to 1. k_0 and k_f are the starting and ending instants of the summation.

It stems directly from this definition that the relative importance of persistent links will be enhanced, while that of short duration ones lessened. The cumulative network matrix for the example network presented in section 6.2 is

$$\mathbf{P}^c = \mathbf{P}^0 + \mathbf{P}^1 + \mathbf{P}^2 = \begin{pmatrix} 24 & 5 & 1 \\ 5 & 25 & 0 \\ 1 & 2 & 27 \end{pmatrix}$$

and its associated graph visualization is shown in figure 6.18; both the matrix and the graph have not been normalized. The links between nodes 1 and 2 are more intense due to the combination of intense exchanges in the discrete time instants.

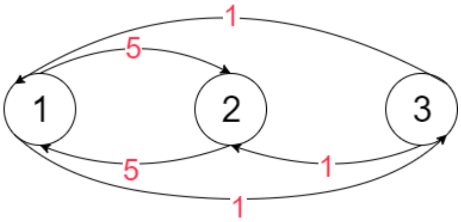


Figure 6.18: Graph of the example flow network, cumulative representation

For large t_0 - t_f spans, the degree distribution becomes trivial, since most nodes appear to be connected and the (normalized) degree reaches the value of 1 with little to no deviation. On the other hand, the strength of nodes quantifies their tendency to receive or donate particles from other levels. The strength distribution over y^+ for the cumulative network comprising all time instants ($t^+ \in [0\ 15\,200]$) is shown in figure 6.19. Contrary to what may be expected, the behaviour of s^{IN} is strongly asymmetric across the channel height; levels near the lower wall of the channel are more inclined to form ingoing links than those on the upper wall. This appears to be at odds with the fact that tracers are distributed homogeneously across the channel height even at large times. The outgoing strength is instead more homogeneous, while still being slightly bigger near the centreline.

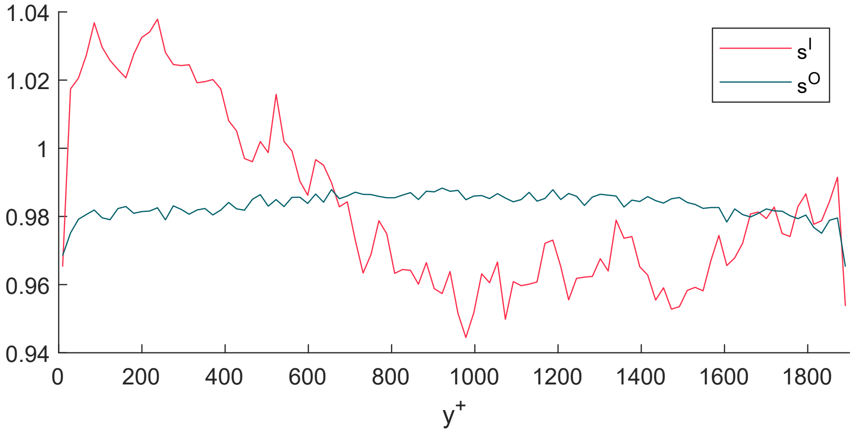


Figure 6.19: Strength distribution of the cumulative network, $t_0^+ = 0$, $t_f^+ = 15\,200$

The cumulative transport matrix for the previous case is shown in figure 6.20. Higher values correspond to persistent and intense links; for example, if a single particle moved immediately from level, say, l to level m and stayed there throughout the temporal evolution of the flow, then the value of P^c_{lm} would be 0.01, which is quite high considering the mean magnitude of P^c_{ij}. Inspection of the matrix shows that levels are more likely to bound for long times if they are physically close one to the other, but if they are located near the boundaries then they may also form longer range links.

In addition to the P^c_{ij} values, also the maximum of each row (blue diamond) and column (red circle) are marked, calculated excluding the main, first lower and first upper diagonals.

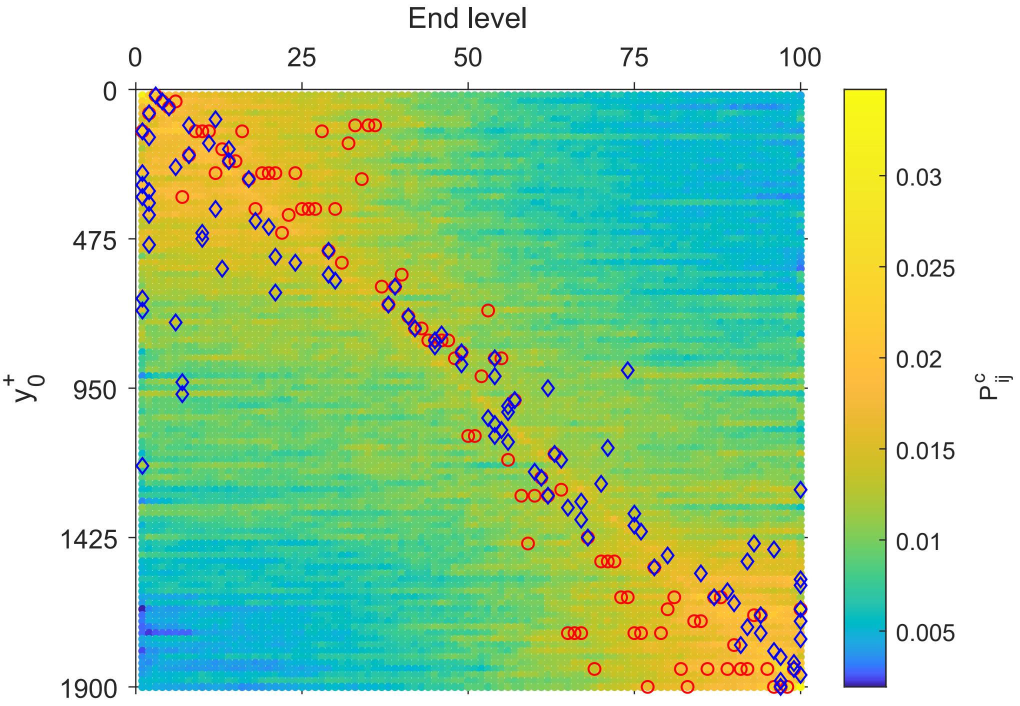


Figure 6.20: Cumulative transport matrix $\mathbf{P}^c$, $t_0^+ = 0$, $t_f^+ = 15\,200$.

The reason for this is to identify the maximum *receiver* level of particles starting from a given level (row maximum) and to identify the maximum *donor* of tracers to a certain level (column maximum).

So for each y_0^+, the level to which most outgoing links were formed is marked by the blue dot. Instead the red circle marks, for each ending level, the starting y_0^+ of the majority of its ingoing links. The main diagonal has been eliminated since the focus is on particles which change level. Furthermore, the first lower and upper diagonals have been elided from the computation of the maxima to eliminate short distance links formed by contiguous levels. While this enhances the visibility of some long distance links, the appearance of the $\mathbf{P}^c$ is not dramatically altered since in some cases long distance links are predominant over shorter ones.

The analysis of maxima reveals some additional features:

- As already noted before, levels near the walls form more long distance links than those near the centre of the channel. In particular, near the centreline, the main donor and receiver of levels are close one to the other (usually no more than 10 levels, or 190 wall units, apart). In the quarters of the channel near the boundaries, instead, maxima are farther away.

- Starting levels close to the boundaries have several of their (row) maxima, marked by blue diamonds, at levels 1 and 100 (and some of their closest neighbours) depending on the channel side. This means that particles from several levels are transferred to the levels closest to the wall and that these particles stay there for quite long times. This behaviour may be related to the presence of vortical structures in the lowest part of the boundary layer that trap tracers (the first level is comprised between $y^+ = 0$ and $y^+ = 19$, so it is below the logarithmic law region).

In addition to that, row maxima near the boundaries are present almost only under the matrix diagonal for the lower wall (up, in the matrix view) and over the diagonal for the upper wall. This means that the main (cumulative) outgoing link for each level is directed towards a lower y^+ coordinate. This is well exemplified by the particles which become trapped in the wall-adjacent zones.

- Column maxima (*i.e.* the red circles) are disposed in a symmetrical way to row maxima. Since column maxima represent the main donor of tracer to a given (ending) level, this means that the main ingoing link for each ending level is coming from a lower y_0^+ coordinate. It also appears that column maxima pop up in contiguous pairs or triplets belonging to the same starting level; this could be related to tracers moving in small structures belonging to these groups of levels for a long time.

All of these observations lead to the identification of preferential paths in the motion of tracers between y^+ levels and to the recognition of regions spanning several levels which have little interchange while still being very active within them.

6.3.3 Network partitioning

The objective of this section will be to provide some insight on the presence of partitions in the flow network, that is zones with a reduced number of links between them, using some results from spectral graph theory [2] [5] [3].

The procedure described in section 3.1.2 will be employed, using the depiction of the flow network graph via its Laplacian eigenvectors. However, since the flow network is weighted other than directed, the definition of the Laplacian matrix will be extended accordingly [24]. First a diagonal weight matrix $\mathbf{W}^d \in \mathbb{R}^{e \times e}$ is introduced, defined as

$$W_{ii}^d = w(l_i) \tag{6.8}$$

where $w(\cdot)$ stands for the weight of a given link and edges l_i are ordered in the same way as the directed incidence matrix $\mathbf{B}$. The weighted Laplacian is then defined as

$$\mathbf{\Lambda}^w = \mathbf{B}\mathbf{W}^d\mathbf{B}^\intercal \tag{6.9}$$

The introduction of a weighted measure is of particular importance when trying to find communities in the cumulative graph defined in the previous section, since for large spans all nodes become connected and no communities can be identified without considering the weight.

Figure 6.21 shows the distribution of the components of the first three non trivial eigenvectors of $\mathbf{\Lambda}^w$. As can be noticed, no clear separation between groups of levels is present, rather components follow an ordered distribution. Still, some grouping takes place,

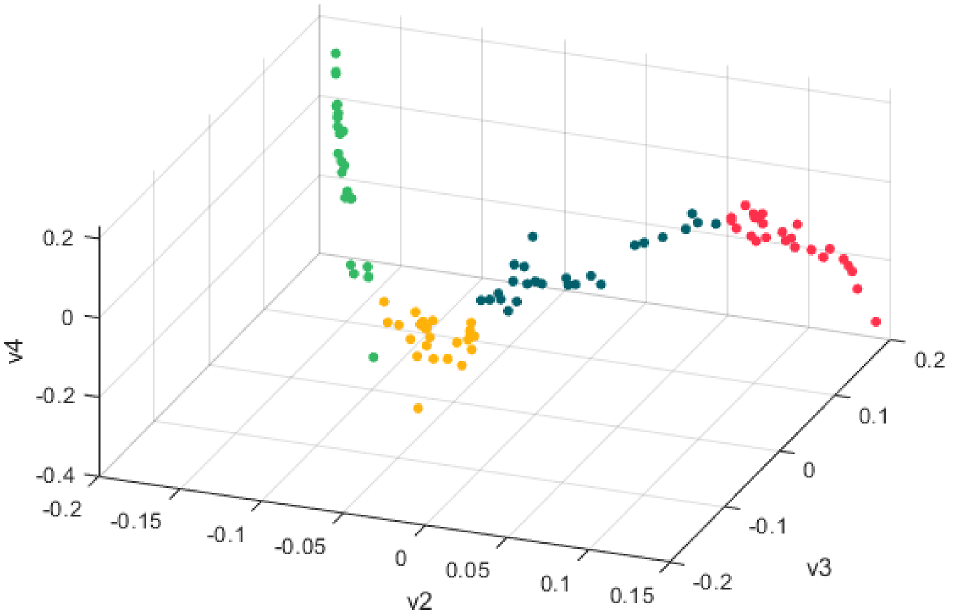


Figure 6.21: Eigenvector plot, non cumulative flow network at $t^+ = 712$. The colouring identifies the origin zones; red: $0 < y^+ < \delta/2$, blue: $\delta/2 < y^+ < \delta$, yellow: $\delta < y^+ < 3\delta/2$, green: $3\delta/2 < y^+ < 2\delta$

especially between the levels of the upper wall region (green dots); anyway levels from the same origin zone stay generally closer one to another than the rest of the set. It should be noted that the main difference in component position is given by the first eigenvectors and as distinction between communities vanish, the number of eigenvectors showing non trivial behaviour is reduced. Figure 6.22 shows the eigenvector plot for $t^+ = 9000$, a time at which most of the original distinction between levels has faded. The fourth eigenvector $\mathbf{v}_4$ has not been shown since it did not provide any additional information on subnetworks separation. The overall separation is reduced and also the ordered disposition of components has gone away, but still some trace of the original disposition of levels is retained. Instead in the eigenvector plot in figure 6.23 all information about the starting position of levels is lost and no community at all can be detected.

The eigenvector plot for the cumulative flow network with $t_0^+ = 0$ and $t_f^+ = 15\,200$ is shown in figure 6.24. Again the distinction between zones is quite marked, even if not totally defined; this is in agreement with the presence of isolated regions which only exchanged tracers between them, as already noted in the previous section.

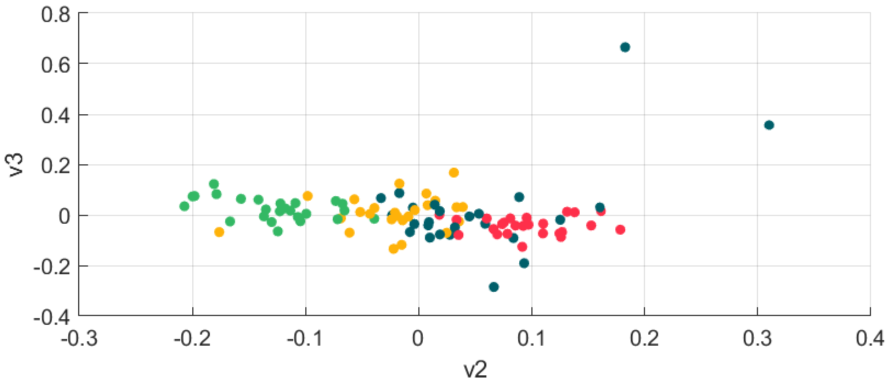


Figure 6.22: Eigenvector plot, non cumulative flow network at $t^+ = 9000$. Colour codes as in 6.21

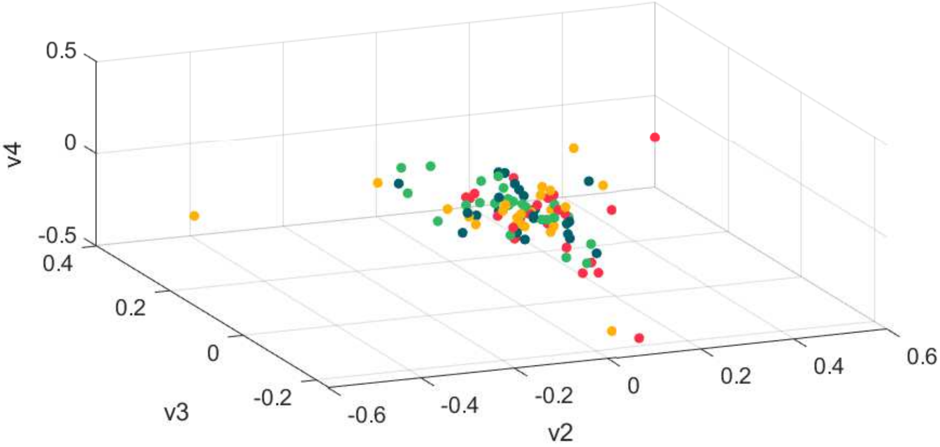


Figure 6.23: Eigenvector plot, non cumulative flow network at $t^+ = 15\,200$. Colour codes as in 6.21

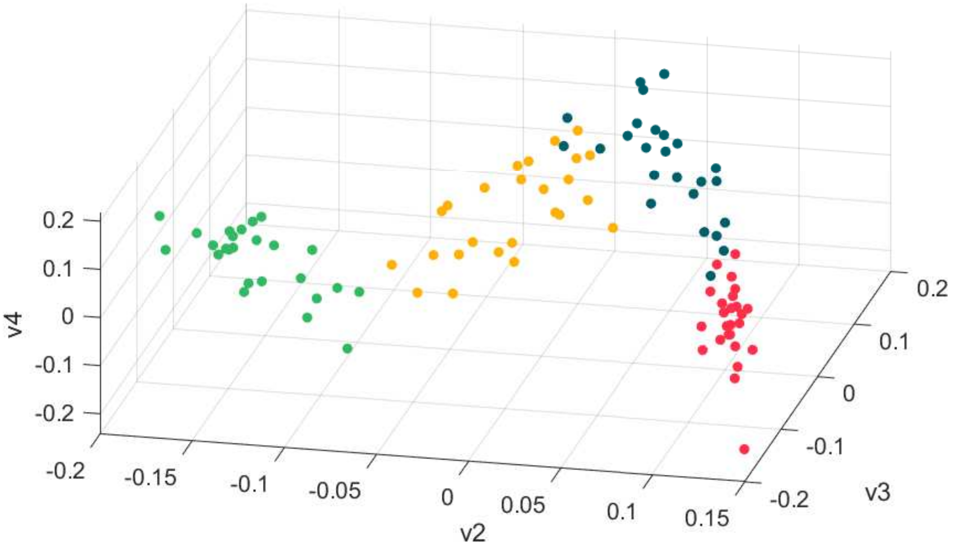


Figure 6.24: Eigenvector plot, cumulative flow network from $t^+ = 0$ to $t^+ = 15\,200$. Colour codes as in 6.21

Chapter 7

Conclusions

In this thesis, the capabilities of network-based methods to describe inhomogeneous turbulence have been successfully assessed. The Lagrangian viewpoint, in particular, has proven especially suitable for this approach, since moving tracers provide a proper starting point for the definition of network nodes. In the following of this chapter, the main results will be summarized, with a focus on the physical aspects emerging from the network description. The second section will deal with some possible future developments of the treated subject.

Main results

The clustering network has the ability to effectively describe interactions between close tracers, which are represented as graph edges. The spatial and temporal distribution of links is non trivial and possibly reflects the presence of structures in the underlying flow field, since it highlights contracting regions in an otherwise diffusive flow. For these reasons, while some metrics of the network (*i.e.* the mean degree, see figure 5.15 on page 45) reflect the shift in diffusion regimes at the already identified turning point of $t^+ \approx 4000$, the most interesting and effective feature of the clustering network lies in the ability to detect and describe coherent structures.

The analysis of both link strengths and durations yields useful information on the presence of complex motions in the channel flow. This is remarkable for two main reasons; the first is that this approach works with a relatively low number of tracers and provides quite significant results even when these are far one from the other due to diffusion. The second reason is the low computational cost, mainly due to the simplicity of the matrix representation of particles' motion. Nonetheless, a more detailed analysis of the dynamics of tracers involved in clustering links is needed, especially in the definition of the threshold distance d_t^+. The choice of a fixed thresholds seems more promising because of its consistent physical meaning, but the adoption of different values does not seem to radically alter the evolution of the network, rather only the number and intensity of its links. Moreover, the evolution of relative distance between tracers (for example figure 5.33) does not seem to reflect different behaviours according to the mutual distance value.

The flow map network reveals in a straightforward manner the motion of tracers between different levels in the non homogeneous direction of the channel, so it is a powerful tool in the representation of wall normal diffusion. The nature of the phenomenon, a transfer of elements between different locations in space, is perfectly described by graphs embedded in physical space. Because of this, many tools of network theory may be employed successfully, most notably the analysis of paths and communities.

Again, network properties shift at $t^+ \approx 4000$ and this change is evidenced by the statistics of degree k and strength s, which reach asymptotic values. The evolution of the degree and of the strength decomposed by y^+ highlights a coherent behaviour and the existence of definite regions inside the flow with different properties regarding the exchange of tracers. Most notably, the region closer to the wall ($y^+ < 475$) forms a reduced number of outgoing links in the initial phase of the evolution of the trajectories, so that tracers appear to be trapped. The analysis of link physical length reveals additional information on the behaviour of nodes; levels close to the walls form longer distance links, probably a signal that some tracers are *ejected* from the lower part of the boundary layer to higher y^+ coordinates. The cumulative network is successful in revealing repeated behaviours in the flow map, thus making most active links visible. Again, it results that the inner part of the boundary layer have strong particles exchanges inside itself, while not being very communicative with other zones. Also, the features of most active links are complex and may require a more accurate interpretation in terms of preferential paths inside the network. Both network formalism and spectal graph theory provide powerful tools to partition the channel in separate zones with little interchange between them. It should again be noted that all of these approaches have a moderate computational cost while still providing effective results.

Future developments

Starting from the present work, some additional research seems necessary, both to expand the subject and to verify some possible unclear points of these approaches. For what concerns clustering, the main working points are:

- a more accurate and reasoned choice of the threshold, which should be connected to physical properties of the underlying flow. This could be obtained employing the Eulerian integral length scales, which seem suitable for the purpose but were not provided in the Lagrangian dataset;
- while the number of tracers seems suitable to adequately describe the clustering problem (at least before diffusion move particles too far one from the other), their disposition at release may be changed. In particular, since tracers are released in a plane ($x^+ = 0$) the description of clustering along the x direction may be flawed; additionally, no tracers are released inside the viscous sublayer ($y^+ < 5$);
- since clustering is heavily influenced by particles' trajectories, a review of the interpolation scheme adopted to extract Lagrangian data seems necessary. The motion of a single particle is surely affected by the usage of a low order scheme and the extent of this may be rather large on clustering statistics.

- a more accurate review of graph theory should be performed, in order to find metrics and tools (other than the degree and strength distributions) fit to provide further insight on clustering;
- the physical behaviour of clustering should be investigated thoroughly, either by providing correlations with the underlying flow or by further analysing trajectories of clustering tracers, especially their velocities and accelerations.

The possible developments for the flow map network are:

- an increase of the number of tracers and a modification of their disposition at release. A greater number of tracers would not provide any difficulty to the method because of its computational lightness, but may provide a better definition to statistical network quantities. Moreover, tracers released inside the viscous sublayer could provide additional detail to that zone of the boundary layer. It may also be of great interest to release a large number of particles at several times from a single point in space and track their motion using the tools here developed;
- a further investigation of the properties of the cumulative network and of its relation with non cumulative ones at different times. In particular, the analysis of paths could provide more detail to the dynamics of tracers while still being computationally effective. To this purpose, the theory regarding temporal evolving networks seems promising [12] [28];
- a further characterization of the spectral representation of this network. Metrics to analytically identify and separate communities should be provided; also, the behaviour of the Laplacian eigenvectors should be investigated, in particular regarding their components' disposition.

Concerning both methods and also the preliminary data presented in chapter 4, it should be investigated why the transition to the Taylor dispersion regime was identified to start at $t^+ \approx 4000$, while in the work of Iacobello et al. [8], which used the same dataset, this was located at $t^+ \approx 5200$.

Bibliography

[1] S. Boccaletti et al. “Complex networks: Structure and dynamics”. In: *Physics Reports* 424.4 (2006), pp. 175–308. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/j.physrep.2005.10.009`.

[2] Béla Bollobás. *Modern graph theory.* Vol. 184. Springer Science & Business Media, 1998.

[3] Andries E. Brouwer and Willem H. Haemers. *Spectra of graphs.* Springer Science & Business Media, 2011.

[4] Jung-Il Choi, Kyongmin Yeo, and Changhoon Lee. “Lagrangian statistics in turbulent channel flow”. In: *Physics of fluids* 16.3 (2004), pp. 779–793.

[5] Luca Donetti and Miguel A. Munoz. “Detecting network communities: a new systematic and efficient algorithm”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.10 (2004), P10012.

[6] Francois Grey and Jørgen K Kjems. “Aggregates, broccoli and cauliflower”. In: *Physica D: Nonlinear Phenomena* 38.1-3 (1989), pp. 154–159.

[7] Giovanni Iacobello, Stefania Scarsoglio, and Luca Ridolfi. “Visibility graph analysis of wall turbulence time-series”. In: *Physics Letters A* 382.1 (2018), pp. 1–11.

[8] Giovanni Iacobello et al. “Lagrangian network analysis of turbulent mixing”. In: *Journal of Fluid Mechanics* 865 (2019), pp. 546–562.

[9] Giovanni Iacobello et al. “Spatial characterization of turbulent channel flow via complex networks”. In: *Physical Review E* 98.1 (2018), p. 013107.

[10] John Kim, Parviz Moin, and Robert Moser. “Turbulence statistics in fully developed channel flow at low Reynolds number”. In: *Journal of fluid mechanics* 177 (1987), pp. 133–166.

[11] Andrey Nikolaevich Kolmogorov. “The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers”. In: *Cr Acad. Sci. URSS* 30 (1941), pp. 301–305.

[12] Vassilis Kostakos. “Temporal graphs”. In: *Physica A: Statistical Mechanics and its Applications* 388.6 (2009), pp. 1007–1023.

[13] Johannes GM Kuerten and JJH Brouwers. “Lagrangian statistics of turbulent channel flow at Reτ= 950 calculated with direct numerical simulation and Langevin models”. In: *Physics of fluids* 25.10 (2013), p. 105108.

[14] Robert D Moser, John Kim, and Nagi N Mansour. "Direct numerical simulation of turbulent channel flow up to Reτ= 590". In: *Physics of fluids* 11.4 (1999), pp. 943–945.

[15] Stephen B. Pope. *Turbulent Flows*. Ed. by Cambridge University Press. 2000.

[16] Lewis F. Richardson. "Weather prediction by numerical process". In: *Quarterly Journal of the Royal Meteorological Society* 48.203 (1922), pp. 282–284. DOI: `10.1002/qj.49704820311`.

[17] Stefania Scarsoglio, Giovanni Iacobello, and Luca Ridolfi. "Complex Networks Unveiling Spatial Patterns in Turbulence". In: *International Journal of Bifurcation and Chaos* 26.13 (2016), p. 1650223. DOI: `10.1142/S0218127416502230`.

[18] Stefania Scarsoglio, Francesco Laio, and Luca Ridolfi. "Climate dynamics: a network-based approach for the analysis of global precipitation". In: *PLoS One* 8.8 (2013), e71129.

[19] Manfred Schroeder and M Herbich. "Fractals, chaos, power laws". In: *Pure and Applied Geophysics* 147.3 (1996), pp. 601–601.

[20] Enrico Ser-Giacomi et al. "Flow networks: A characterization of geophysical fluid transport". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.3 (2015), p. 036404.

[21] KR Sreenivasan and CJFM Meneveau. "The fractal facets of turbulence". In: *Journal of Fluid Mechanics* 173 (1986), pp. 357–386.

[22] Nickolas Stelzenmuller. "A Lagrangian study of inhomogeneous turbulence". PhD thesis. Université Grenoble Alpes, 2017.

[23] Nickolas Stelzenmuller et al. "Lagrangian acceleration statistics in a turbulent channel flow". In: *Physical Review Fluids* 2.5 (2017), p. 054602.

[24] Christian Szegedy et al. "Some Applications of the weighted combinatorial Laplacian". PhD thesis. Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn, 2005.

[25] Geoffrey I Taylor. "Diffusion by continuous movements". In: *Proceedings of the london mathematical society* 2.1 (1922), pp. 196–212.

[26] Federico Toschi and Eberhard Bodenschatz. "Lagrangian Properties of Particles in Turbulence". In: *Annual Review of Fluid Mechanics* 41.1 (2009), pp. 375–404. DOI: `10.1146/annurev.fluid.010908.165210`.

[27] MAT van Hinsberg et al. "Optimal interpolation schemes for particle tracking in turbulence". In: *Physical Review E* 87.4 (2013), p. 043307.

[28] Huanhuan Wu et al. "Path problems in temporal graphs". In: *Proceedings of the VLDB Endowment* 7.9 (2014), pp. 721–732.

[29] PK Yeung. "Lagrangian investigations of turbulence". In: *Annual review of fluid mechanics* 34.1 (2002), pp. 115–142.